Plotting Venn Diagrams for Microbiome studies

Gian MN Benucci

September 9th, 2016

This is an R Markdown document contains code to perform Venn Diagrams using the Bioconductor package *limma* starting from an object created using the *Pyloseq* package.

```
library(phyloseq)
data(GlobalPatterns)
GlobalPatterns
```

```
## phyloseq-class experiment-level object
## otu_table()   OTU Table:         [ 19216 taxa and 26 samples ]
## sample_data() Sample Data:       [ 26 samples by 7 sample variables ]
## tax_table()   Taxonomy Table:    [ 19216 taxa by 7 taxonomic ranks ]
## phy_tree()    Phylogenetic Tree: [ 19216 tips and 19215 internal nodes ]
```

First of all lets surce the required package and then merge samples according the metadata variable of our interest.

```
library(limma)

GB_st = merge_samples(GlobalPatterns, "SampleType")
sample_data(GB_st)
```

```
## Sample Data:        [9 samples by 7 sample variables]:
##                     X.SampleID Primer Final_Barcode
## Feces                     19.0   13.5          13.5
## Freshwater                15.0   11.5          11.5
## Freshwater (creek)         2.0   14.0          14.0
## Mock                       7.0   25.0          25.0
## Ocean                     18.0   17.0          17.0
## Sediment (estuary)        23.0   20.0          20.0
## Skin                      12.0    7.0           7.0
## Soil                      10.0    2.0           2.0
## Tongue                    14.5    9.5           9.5
##                     Barcode_truncated_plus_T Barcode_full_length SampleType
## Feces                               16.500000           13.750000          1
## Freshwater                          12.000000            4.500000          2
## Freshwater (creek)                  13.000000            6.666667          3
## Mock                                12.333333           16.000000          4
## Ocean                               13.666667           17.000000          5
## Sediment (estuary)                  15.000000           14.666667          6
## Skin                                 9.666667           14.666667          7
## Soil                                13.333333           11.333333          8
## Tongue                              15.000000           23.000000          9
##                     Description
## Feces                 18.500000
## Freshwater            15.500000
## Freshwater (creek)     2.000000
```

```
## Mock                   7.000000
## Ocean                 18.000000
## Sediment (estuary)    22.666667
## Skin                  12.000000
## Soil                   9.666667
## Tongue                14.500000
```

Now create the object to calculate the variable intersections and then plot the Venn Diagram.

```
table_GB_st <- t(otu_table(subset_samples(GB_st, SampleType%in%c(2,5,8))))
venn_counts <- vennCounts(table_GB_st)
venn_counts
```

```
##   Freshwater Ocean Soil Counts
## 1          0     0    0   5615
## 2          0     0    1   5706
## 3          0     1    0   2212
## 4          0     1    1    841
## 5          1     0    0    749
## 6          1     0    1   1477
## 7          1     1    0    669
## 8          1     1    1   1947
## attr(,"class")
## [1] "VennCounts"
```

It is possible to plot interesections of up to 5 group variables. Plots are high customizable, please run help(vennDiagram) for more details on the function.

```
vennDiagram(venn_counts, cex=c(1,1.2,0.8),
            names = c("Freshwater", "Ocean", "Soil"),
            circle.col = c("red","blue","green"))
```