**Creating a phyloseq object using the supporting data excel spreadsheets**

In order to create a phyloseq object in R, 3 different sets of data are needed: a matrix of counts, taxonomy information and metadata. Therefore, there are 15 supporting data spreadsheets - 3 each for the mock extraction, mock sequencing, within-run repeats, between-run repeats and stool/swab analysis.

#Showing an example for reading the mock DNA control spreadsheets into R and creating a phyloseq object for downstream analysis.

#Reading in the excel spreadsheets for mock controls

mock\_dna\_count <- read.xlsx(file="mock\_dna\_count.xlsx", header=TRUE, sheetName="sheet1", row.names = 1)

mock\_dna\_tax <- read.xlsx(file="mock\_dna\_tax.xlsx", header=TRUE, sheetName="sheet1", row.names = 1)

mock\_dna\_samples <- read.xlsx(file="mock\_dna\_samples.xlsx", header=TRUE, sheetName="sheet1", row.names = 1)

table(rownames(mock\_dna\_tax) == rownames(mock\_dna\_count))

mock\_dna\_tax\_matrix <- as.matrix(mock\_dna\_tax)

mock\_dna\_count\_matrix <- as.matrix(mock\_dna\_count)

#Creating a phyloseq object

library(phyloseq)

phy <- phyloseq(otu\_table(mock\_dna\_count\_matrix, taxa\_are\_rows = T),tax\_table(mock\_dna\_tax\_matrix), sample\_data(mock\_dna\_samples))

str(phy)

phy