Homework 3 - Answer Key

Melinda Higgins & Vicki Hertzberg

February 21, 2018

## Problem 1 - Document that you have installed dada2 and phyloseq, as well as shortread and ggplot2.

Yes, I successfull installed these packages. Here is proof that dada2 is working.

NOTE, the following versions should have been installed:

* dada2 version 1.6.x - see <https://www.bioconductor.org/packages/release/bioc/html/dada2.html>
* ggplot2 version 2.2.x - see <https://cran.r-project.org/web/packages/ggplot2/index.html>
* ShortRead version 1.36.x - see <https://www.bioconductor.org/packages/release/bioc/html/ShortRead.html>
* phyloseq version 1.22.x - see <https://www.bioconductor.org/packages/release/bioc/html/phyloseq.html>

library(dada2); packageVersion("dada2")

## Loading required package: Rcpp

## [1] '1.6.0'

library(ggplot2); packageVersion("ggplot2")

## [1] '2.2.1'

library(ShortRead); packageVersion("ShortRead")

## Loading required package: BiocGenerics

## Loading required package: parallel

##   
## Attaching package: 'BiocGenerics'

## The following objects are masked from 'package:parallel':  
##   
## clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,  
## clusterExport, clusterMap, parApply, parCapply, parLapply,  
## parLapplyLB, parRapply, parSapply, parSapplyLB

## The following objects are masked from 'package:stats':  
##   
## IQR, mad, sd, var, xtabs

## The following objects are masked from 'package:base':  
##   
## anyDuplicated, append, as.data.frame, cbind, colMeans,  
## colnames, colSums, do.call, duplicated, eval, evalq, Filter,  
## Find, get, grep, grepl, intersect, is.unsorted, lapply,  
## lengths, Map, mapply, match, mget, order, paste, pmax,  
## pmax.int, pmin, pmin.int, Position, rank, rbind, Reduce,  
## rowMeans, rownames, rowSums, sapply, setdiff, sort, table,  
## tapply, union, unique, unsplit, which, which.max, which.min

## Loading required package: BiocParallel

## Loading required package: Biostrings

## Loading required package: S4Vectors

## Loading required package: stats4

##   
## Attaching package: 'S4Vectors'

## The following object is masked from 'package:base':  
##   
## expand.grid

## Loading required package: IRanges

## Loading required package: XVector

##   
## Attaching package: 'Biostrings'

## The following object is masked from 'package:base':  
##   
## strsplit

## Loading required package: Rsamtools

## Loading required package: GenomeInfoDb

## Loading required package: GenomicRanges

## Loading required package: GenomicAlignments

## Loading required package: SummarizedExperiment

## Loading required package: Biobase

## Welcome to Bioconductor  
##   
## Vignettes contain introductory material; view with  
## 'browseVignettes()'. To cite Bioconductor, see  
## 'citation("Biobase")', and for packages 'citation("pkgname")'.

## Loading required package: DelayedArray

## Loading required package: matrixStats

##   
## Attaching package: 'matrixStats'

## The following objects are masked from 'package:Biobase':  
##   
## anyMissing, rowMedians

##   
## Attaching package: 'DelayedArray'

## The following objects are masked from 'package:matrixStats':  
##   
## colMaxs, colMins, colRanges, rowMaxs, rowMins, rowRanges

## The following object is masked from 'package:Biostrings':  
##   
## type

## The following object is masked from 'package:base':  
##   
## apply

## [1] '1.36.0'

library(phyloseq); packageVersion("phyloseq")

##   
## Attaching package: 'phyloseq'

## The following object is masked from 'package:SummarizedExperiment':  
##   
## distance

## The following object is masked from 'package:Biobase':  
##   
## sampleNames

## The following object is masked from 'package:GenomicRanges':  
##   
## distance

## The following object is masked from 'package:IRanges':  
##   
## distance

## [1] '1.22.3'

## Problem 2 - Please document that you have downloaded the THREE additional files:

* the fastq files from the Mothur MiseqSOP walkthrough
* silva\_nr\_v132\_train\_set.fa.gz
* silva\_species\_assignment\_v132.fa.gz

# the ~ should work on a Mac  
# path <- "~/MiSeq\_SOP"  
  
# if you are on a Windows OS computer (like me), use   
# . instead of ~  
path <- "./MiSeq\_SOP"  
  
# list the files in the  
# /MiSeq\_SOP directory  
fns <- list.files(path)  
fns

## [1] "F3D0\_S188\_L001\_R1\_001.fastq"   
## [2] "F3D0\_S188\_L001\_R2\_001.fastq"   
## [3] "F3D1\_S189\_L001\_R1\_001.fastq"   
## [4] "F3D1\_S189\_L001\_R2\_001.fastq"   
## [5] "F3D141\_S207\_L001\_R1\_001.fastq"   
## [6] "F3D141\_S207\_L001\_R2\_001.fastq"   
## [7] "F3D142\_S208\_L001\_R1\_001.fastq"   
## [8] "F3D142\_S208\_L001\_R2\_001.fastq"   
## [9] "F3D143\_S209\_L001\_R1\_001.fastq"   
## [10] "F3D143\_S209\_L001\_R2\_001.fastq"   
## [11] "F3D144\_S210\_L001\_R1\_001.fastq"   
## [12] "F3D144\_S210\_L001\_R2\_001.fastq"   
## [13] "F3D145\_S211\_L001\_R1\_001.fastq"   
## [14] "F3D145\_S211\_L001\_R2\_001.fastq"   
## [15] "F3D146\_S212\_L001\_R1\_001.fastq"   
## [16] "F3D146\_S212\_L001\_R2\_001.fastq"   
## [17] "F3D147\_S213\_L001\_R1\_001.fastq"   
## [18] "F3D147\_S213\_L001\_R2\_001.fastq"   
## [19] "F3D148\_S214\_L001\_R1\_001.fastq"   
## [20] "F3D148\_S214\_L001\_R2\_001.fastq"   
## [21] "F3D149\_S215\_L001\_R1\_001.fastq"   
## [22] "F3D149\_S215\_L001\_R2\_001.fastq"   
## [23] "F3D150\_S216\_L001\_R1\_001.fastq"   
## [24] "F3D150\_S216\_L001\_R2\_001.fastq"   
## [25] "F3D2\_S190\_L001\_R1\_001.fastq"   
## [26] "F3D2\_S190\_L001\_R2\_001.fastq"   
## [27] "F3D3\_S191\_L001\_R1\_001.fastq"   
## [28] "F3D3\_S191\_L001\_R2\_001.fastq"   
## [29] "F3D5\_S193\_L001\_R1\_001.fastq"   
## [30] "F3D5\_S193\_L001\_R2\_001.fastq"   
## [31] "F3D6\_S194\_L001\_R1\_001.fastq"   
## [32] "F3D6\_S194\_L001\_R2\_001.fastq"   
## [33] "F3D7\_S195\_L001\_R1\_001.fastq"   
## [34] "F3D7\_S195\_L001\_R2\_001.fastq"   
## [35] "F3D8\_S196\_L001\_R1\_001.fastq"   
## [36] "F3D8\_S196\_L001\_R2\_001.fastq"   
## [37] "F3D9\_S197\_L001\_R1\_001.fastq"   
## [38] "F3D9\_S197\_L001\_R2\_001.fastq"   
## [39] "filtered"   
## [40] "HMP\_MOCK.v35.fasta"   
## [41] "Mock\_S280\_L001\_R1\_001.fastq"   
## [42] "Mock\_S280\_L001\_R2\_001.fastq"   
## [43] "mouse.dpw.metadata"   
## [44] "mouse.time.design"   
## [45] "silva\_nr\_v132\_train\_set.fa.gz"   
## [46] "silva\_species\_assignment\_v132.fa.gz"  
## [47] "stability.batch"   
## [48] "stability.files"