Homework3

Madelyn Houser

February 27, 2019

Package Installation

Install packages dada2, `phyloseq`, `DECIPHER`, `ShortRead`, and `ggplot2`. Then verify their installation by confirming the package version.

```
library(phyloseq)
packageVersion("phyloseq")
```

```
## [1] '1.26.1'
```

```
library(DECIPHER)
```

```
## Loading required package: Biostrings

## 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, basename, cbind,
##     colMeans, colnames, colSums, dirname, 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: S4Vectors
```

```
## Loading required package: stats4

##
## Attaching package: 'S4Vectors'

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

## Loading required package: IRanges

##
## Attaching package: 'IRanges'

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

## The following object is masked from 'package:grDevices':
##
##     windows

## Loading required package: XVector

##
## Attaching package: 'Biostrings'

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

## Loading required package: RSQLite
```

```r
packageVersion("DECIPHER")
```

```
## [1] '2.10.2'
```

```r
library(ShortRead)
```

```
## Loading required package: BiocParallel

## 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")'.

##
## Attaching package: 'Biobase'

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

## 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 objects are masked from 'package:base':
##
##     aperm, apply
```

```r
packageVersion("ShortRead")
```

```
## [1] '1.40.0'
```

```r
library(dada2)
```

```
## Loading required package: Rcpp
```

```r
packageVersion("dada2")
```

```
## [1] '1.10.1'
```

```r
library(ggplot2)
packageVersion("ggplot2")
```

```
## [1] '3.1.0'
```

File Acquisition

Download and unzip the fastq files from the Mothur MiSeq SOP. Save the Silva 132 training set and Silva 132 species assignment zip files and Silva SSU r132 RData object in the same folder.

```
fns <-
list.files("C:/Users/Acer/Documents/MadelynClasses/DataAnalytics/dada2_lesson
/MiSeqSOPData/MiSeq_SOP")

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] "HMP_MOCK.v35.fasta"
## [40] "Mock_S280_L001_R1_001.fastq"
## [41] "Mock_S280_L001_R2_001.fastq"
## [42] "mouse.dpw.metadata"
## [43] "mouse.time.design"
## [44] "silva_nr_v132_train_set.fa.gz"
## [45] "silva_species_assignment_v132.fa.gz"
## [46] "SILVA_SSU_r132_March2018.RData"
## [47] "stability.batch"
## [48] "stability.files"
```