The Maternal Newborn Oral Microbiome

Irene Yang

3/28/2018

## Note: The raw fastq files that support this project can be found within the secure Emory Box location at: <https://emory.app.box.com/folder/48209374654>

# Step 1: Processing fastq files to come up with OTU table.

### Load packages

library(dada2); packageVersion("dada2")

## [1] '1.6.0'

library(ShortRead); packageVersion("ShortRead")

## [1] '1.36.1'

library(phyloseq); packageVersion("phyloseq")

## [1] '1.22.3'

library(ggplot2); packageVersion("ggplot2")

## [1] '2.2.1'

### Change the path in the next chunk to where your files sit.

# Set the path to the data files  
  
path <- "~/Desktop/N741/2018Week7/AWHONN Fastq Files"  
fileNames <- list.files(path)  
fileNames

## [1] "filtered"   
## [2] "gg\_13\_8\_train\_set\_97.fa"   
## [3] "S0013-0001\_S97\_L001\_R1\_001.fastq"   
## [4] "S0013-0001\_S97\_L001\_R2\_001.fastq"   
## [5] "S0013-0002\_S98\_L001\_R1\_001.fastq"   
## [6] "S0013-0002\_S98\_L001\_R2\_001.fastq"   
## [7] "S0013-0003\_S99\_L001\_R1\_001.fastq"   
## [8] "S0013-0003\_S99\_L001\_R2\_001.fastq"   
## [9] "S0013-0004\_S100\_L001\_R1\_001.fastq"  
## [10] "S0013-0004\_S100\_L001\_R2\_001.fastq"  
## [11] "S0013-0005\_S101\_L001\_R1\_001.fastq"  
## [12] "S0013-0005\_S101\_L001\_R2\_001.fastq"  
## [13] "S0013-0006\_S102\_L001\_R1\_001.fastq"  
## [14] "S0013-0006\_S102\_L001\_R2\_001.fastq"  
## [15] "S0013-0007\_S103\_L001\_R1\_001.fastq"  
## [16] "S0013-0007\_S103\_L001\_R2\_001.fastq"  
## [17] "S0013-0008\_S104\_L001\_R1\_001.fastq"  
## [18] "S0013-0008\_S104\_L001\_R2\_001.fastq"  
## [19] "S0013-0009\_S105\_L001\_R1\_001.fastq"  
## [20] "S0013-0009\_S105\_L001\_R2\_001.fastq"  
## [21] "S0013-0010\_S106\_L001\_R1\_001.fastq"  
## [22] "S0013-0010\_S106\_L001\_R2\_001.fastq"  
## [23] "S0013-0011\_S107\_L001\_R1\_001.fastq"  
## [24] "S0013-0011\_S107\_L001\_R2\_001.fastq"  
## [25] "S0013-0012\_S108\_L001\_R1\_001.fastq"  
## [26] "S0013-0012\_S108\_L001\_R2\_001.fastq"  
## [27] "S0013-0013\_S109\_L001\_R1\_001.fastq"  
## [28] "S0013-0013\_S109\_L001\_R2\_001.fastq"  
## [29] "S0013-0014\_S110\_L001\_R1\_001.fastq"  
## [30] "S0013-0014\_S110\_L001\_R2\_001.fastq"  
## [31] "S0013-0015\_S111\_L001\_R1\_001.fastq"  
## [32] "S0013-0015\_S111\_L001\_R2\_001.fastq"  
## [33] "S0013-0016\_S112\_L001\_R1\_001.fastq"  
## [34] "S0013-0016\_S112\_L001\_R2\_001.fastq"  
## [35] "S0013-0017\_S113\_L001\_R1\_001.fastq"  
## [36] "S0013-0017\_S113\_L001\_R2\_001.fastq"  
## [37] "S0013-0018\_S114\_L001\_R1\_001.fastq"  
## [38] "S0013-0018\_S114\_L001\_R2\_001.fastq"  
## [39] "S0013-0019\_S115\_L001\_R1\_001.fastq"  
## [40] "S0013-0019\_S115\_L001\_R2\_001.fastq"  
## [41] "S0013-0020\_S116\_L001\_R1\_001.fastq"  
## [42] "S0013-0020\_S116\_L001\_R2\_001.fastq"  
## [43] "S0013-0021\_S117\_L001\_R1\_001.fastq"  
## [44] "S0013-0021\_S117\_L001\_R2\_001.fastq"  
## [45] "S0013-0022\_S118\_L001\_R1\_001.fastq"  
## [46] "S0013-0022\_S118\_L001\_R2\_001.fastq"  
## [47] "S0013-0023\_S119\_L001\_R1\_001.fastq"  
## [48] "S0013-0023\_S119\_L001\_R2\_001.fastq"  
## [49] "S0013-0024\_S120\_L001\_R1\_001.fastq"  
## [50] "S0013-0024\_S120\_L001\_R2\_001.fastq"  
## [51] "S0013-0025\_S121\_L001\_R1\_001.fastq"  
## [52] "S0013-0025\_S121\_L001\_R2\_001.fastq"  
## [53] "S0013-0026\_S122\_L001\_R1\_001.fastq"  
## [54] "S0013-0026\_S122\_L001\_R2\_001.fastq"  
## [55] "S0013-0027\_S123\_L001\_R1\_001.fastq"  
## [56] "S0013-0027\_S123\_L001\_R2\_001.fastq"  
## [57] "S0013-0028\_S124\_L001\_R1\_001.fastq"  
## [58] "S0013-0028\_S124\_L001\_R2\_001.fastq"  
## [59] "S0013-0029\_S125\_L001\_R1\_001.fastq"  
## [60] "S0013-0029\_S125\_L001\_R2\_001.fastq"  
## [61] "S0013-0030\_S126\_L001\_R1\_001.fastq"  
## [62] "S0013-0030\_S126\_L001\_R2\_001.fastq"  
## [63] "S0013-0031\_S127\_L001\_R1\_001.fastq"  
## [64] "S0013-0031\_S127\_L001\_R2\_001.fastq"  
## [65] "S0013-0032\_S128\_L001\_R1\_001.fastq"  
## [66] "S0013-0032\_S128\_L001\_R2\_001.fastq"  
## [67] "S0013-0033\_S129\_L001\_R1\_001.fastq"  
## [68] "S0013-0033\_S129\_L001\_R2\_001.fastq"  
## [69] "S0013-0034\_S130\_L001\_R1\_001.fastq"  
## [70] "S0013-0034\_S130\_L001\_R2\_001.fastq"  
## [71] "S0013-0035\_S131\_L001\_R1\_001.fastq"  
## [72] "S0013-0035\_S131\_L001\_R2\_001.fastq"  
## [73] "S0013-0036\_S132\_L001\_R1\_001.fastq"  
## [74] "S0013-0036\_S132\_L001\_R2\_001.fastq"  
## [75] "S0013-0037\_S133\_L001\_R1\_001.fastq"  
## [76] "S0013-0037\_S133\_L001\_R2\_001.fastq"  
## [77] "S0013-0038\_S134\_L001\_R1\_001.fastq"  
## [78] "S0013-0038\_S134\_L001\_R2\_001.fastq"  
## [79] "S0013-0039\_S135\_L001\_R1\_001.fastq"  
## [80] "S0013-0039\_S135\_L001\_R2\_001.fastq"  
## [81] "S0013-0040\_S136\_L001\_R1\_001.fastq"  
## [82] "S0013-0040\_S136\_L001\_R2\_001.fastq"  
## [83] "S0013-0041\_S137\_L001\_R1\_001.fastq"  
## [84] "S0013-0041\_S137\_L001\_R2\_001.fastq"  
## [85] "S0013-0042\_S138\_L001\_R1\_001.fastq"  
## [86] "S0013-0042\_S138\_L001\_R2\_001.fastq"  
## [87] "S0013-0043\_S139\_L001\_R1\_001.fastq"  
## [88] "S0013-0043\_S139\_L001\_R2\_001.fastq"  
## [89] "S0013-0044\_S140\_L001\_R1\_001.fastq"  
## [90] "S0013-0044\_S140\_L001\_R2\_001.fastq"  
## [91] "S0013-0045\_S141\_L001\_R1\_001.fastq"  
## [92] "S0013-0045\_S141\_L001\_R2\_001.fastq"  
## [93] "S0013-0046\_S142\_L001\_R1\_001.fastq"  
## [94] "S0013-0046\_S142\_L001\_R2\_001.fastq"  
## [95] "S0013-0047\_S143\_L001\_R1\_001.fastq"  
## [96] "S0013-0047\_S143\_L001\_R2\_001.fastq"  
## [97] "S0013-0048\_S144\_L001\_R1\_001.fastq"  
## [98] "S0013-0048\_S144\_L001\_R2\_001.fastq"  
## [99] "S0013-0049\_S145\_L001\_R1\_001.fastq"  
## [100] "S0013-0049\_S145\_L001\_R2\_001.fastq"  
## [101] "S0013-0050\_S146\_L001\_R1\_001.fastq"  
## [102] "S0013-0050\_S146\_L001\_R2\_001.fastq"  
## [103] "S0013-0051\_S147\_L001\_R1\_001.fastq"  
## [104] "S0013-0051\_S147\_L001\_R2\_001.fastq"  
## [105] "S0013-0052\_S148\_L001\_R1\_001.fastq"  
## [106] "S0013-0052\_S148\_L001\_R2\_001.fastq"  
## [107] "S0013-0053\_S149\_L001\_R1\_001.fastq"  
## [108] "S0013-0053\_S149\_L001\_R2\_001.fastq"  
## [109] "S0013-0054\_S150\_L001\_R1\_001.fastq"  
## [110] "S0013-0054\_S150\_L001\_R2\_001.fastq"  
## [111] "S0013-0055\_S151\_L001\_R1\_001.fastq"  
## [112] "S0013-0055\_S151\_L001\_R2\_001.fastq"  
## [113] "S0013-0056\_S152\_L001\_R1\_001.fastq"  
## [114] "S0013-0056\_S152\_L001\_R2\_001.fastq"

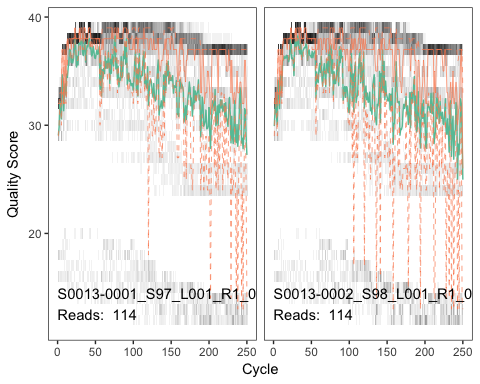
### Read in sample names

Using the dada2 pipeline, first read in the names of the .fastq files. Then manipulate those names as character variables, using regular expressions to create lists of the forward and reverse read .fastq files in *matched* order.

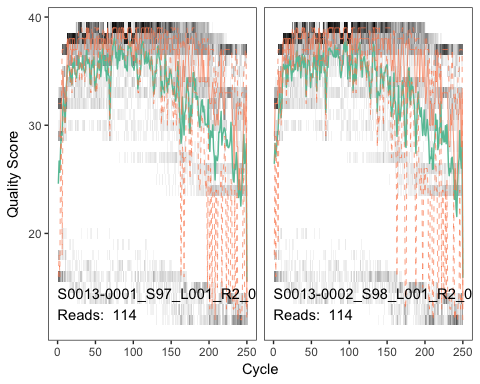
# Forward and reverse fastq filenames should have format: SAMPLENAME\_R1\_001.fastq and SAMPLENAME\_R2\_001.fastq  
  
# Start by reading in the names of the .fastq files  
  
fnFs <- sort(list.files(path, pattern="\_R1\_001.fastq", full.names=TRUE))  
fnRs <- sort(list.files(path, pattern="\_R2\_001.fastq", full.names=TRUE))  
  
# Extract sample names, assuming filenames have format: SAMPLENAME\_XXX.fastq  
  
sample.names <- sapply(strsplit(basename(fnFs), "\_"), `[`, 1)

### Generate Quality Profiles of the reads

# Visualize the quality profile of the first two files containing forward reads  
  
plotQualityProfile(fnFs[1:2])



# Visualize the quality profile of the first two files containing reverse reads  
  
plotQualityProfile(fnRs[1:2])



### Filter and Trim

Typical filtering parameters were used:  
- maxN = 0 – dada2 requires that there be no N’s in a sequence - truncQ = 2 – truncate reads at the first instance of a quality less than or equal to #. - maxEE = 2 – sets the maximum number of expected errors allowed in a read, which is a better filter than simply averaging quality scores.

Note: Decision made to trim forward reads at 180 and reverse reads at 160. Overlap between forward and reverse reads was ensured.

# Make a directory and filenames for the filtered fastqs  
   
# Place filtered files in a filtered/ subdirectory  
  
filt.path <- file.path(path, "filtered")  
if(!file\_test("-d", filt.path)) dir.create(filt.path)  
filtFs <- file.path(filt.path, paste0(sample.names, "\_F\_filt.fastq.gz"))  
filtRs <- file.path(filt.path, paste0(sample.names, "\_R\_file.fastq.gz"))  
  
# Filter the forward and reverse reads  
  
out <- filterAndTrim(fnFs, filtFs, fnRs, filtRs, truncLen = c(200, 190),  
 maxN=0, maxEE =c(2,2), truncQ = 2, rm.phix = TRUE,  
 compress=TRUE, multithread=TRUE)   
  
head(out)

## reads.in reads.out  
## S0013-0001\_S97\_L001\_R1\_001.fastq 114 99  
## S0013-0002\_S98\_L001\_R1\_001.fastq 114 94  
## S0013-0003\_S99\_L001\_R1\_001.fastq 114 103  
## S0013-0004\_S100\_L001\_R1\_001.fastq 114 93  
## S0013-0005\_S101\_L001\_R1\_001.fastq 114 99  
## S0013-0006\_S102\_L001\_R1\_001.fastq 114 95

### Learn the Error Rates

errF <- learnErrors(filtFs, multithread = TRUE)

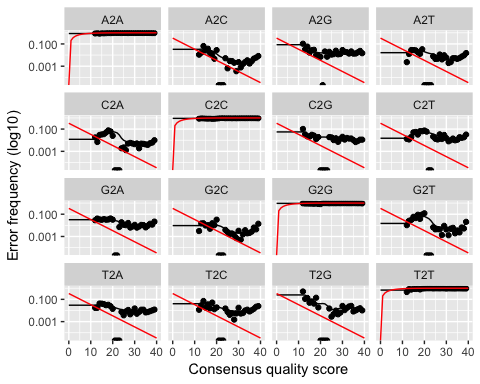
## Initializing error rates to maximum possible estimate.  
## Sample 1 - 99 reads in 57 unique sequences.  
## Sample 2 - 94 reads in 49 unique sequences.  
## Sample 3 - 103 reads in 66 unique sequences.  
## Sample 4 - 93 reads in 52 unique sequences.  
## Sample 5 - 99 reads in 35 unique sequences.  
## Sample 6 - 95 reads in 34 unique sequences.  
## Sample 7 - 105 reads in 63 unique sequences.  
## Sample 8 - 99 reads in 46 unique sequences.  
## Sample 9 - 96 reads in 39 unique sequences.  
## Sample 10 - 92 reads in 58 unique sequences.  
## Sample 11 - 97 reads in 58 unique sequences.  
## Sample 12 - 104 reads in 61 unique sequences.  
## Sample 13 - 102 reads in 56 unique sequences.  
## Sample 14 - 97 reads in 55 unique sequences.  
## Sample 15 - 101 reads in 55 unique sequences.  
## Sample 16 - 104 reads in 54 unique sequences.  
## Sample 17 - 91 reads in 47 unique sequences.  
## Sample 18 - 97 reads in 50 unique sequences.  
## Sample 19 - 105 reads in 61 unique sequences.  
## Sample 20 - 101 reads in 52 unique sequences.  
## Sample 21 - 100 reads in 65 unique sequences.  
## Sample 22 - 101 reads in 55 unique sequences.  
## Sample 23 - 98 reads in 37 unique sequences.  
## Sample 24 - 96 reads in 60 unique sequences.  
## Sample 25 - 95 reads in 54 unique sequences.  
## Sample 26 - 101 reads in 52 unique sequences.  
## Sample 27 - 91 reads in 60 unique sequences.  
## Sample 28 - 100 reads in 58 unique sequences.  
## Sample 29 - 98 reads in 57 unique sequences.  
## Sample 30 - 97 reads in 69 unique sequences.  
## Sample 31 - 89 reads in 58 unique sequences.  
## Sample 32 - 82 reads in 48 unique sequences.  
## Sample 33 - 104 reads in 67 unique sequences.  
## Sample 34 - 84 reads in 50 unique sequences.  
## Sample 35 - 110 reads in 70 unique sequences.  
## Sample 36 - 103 reads in 62 unique sequences.  
## Sample 37 - 108 reads in 77 unique sequences.  
## Sample 38 - 100 reads in 58 unique sequences.  
## Sample 39 - 101 reads in 64 unique sequences.  
## Sample 40 - 101 reads in 67 unique sequences.  
## Sample 41 - 96 reads in 49 unique sequences.  
## Sample 42 - 107 reads in 62 unique sequences.  
## Sample 43 - 99 reads in 69 unique sequences.  
## Sample 44 - 96 reads in 50 unique sequences.  
## Sample 45 - 106 reads in 68 unique sequences.  
## Sample 46 - 103 reads in 59 unique sequences.  
## Sample 47 - 89 reads in 45 unique sequences.  
## Sample 48 - 103 reads in 64 unique sequences.  
## Sample 49 - 97 reads in 59 unique sequences.  
## Sample 50 - 100 reads in 41 unique sequences.  
## Sample 51 - 108 reads in 49 unique sequences.  
## Sample 52 - 103 reads in 69 unique sequences.  
## Sample 53 - 98 reads in 62 unique sequences.  
## Sample 54 - 99 reads in 71 unique sequences.  
## Sample 55 - 90 reads in 65 unique sequences.  
## Sample 56 - 84 reads in 38 unique sequences.  
## selfConsist step 2   
## selfConsist step 3   
## selfConsist step 4   
## Convergence after 4 rounds.  
## Total reads used: 5511

errR <- learnErrors(filtRs, multithread = TRUE)

## Initializing error rates to maximum possible estimate.  
## Sample 1 - 99 reads in 61 unique sequences.  
## Sample 2 - 94 reads in 55 unique sequences.  
## Sample 3 - 103 reads in 74 unique sequences.  
## Sample 4 - 93 reads in 52 unique sequences.  
## Sample 5 - 99 reads in 49 unique sequences.  
## Sample 6 - 95 reads in 41 unique sequences.  
## Sample 7 - 105 reads in 89 unique sequences.  
## Sample 8 - 99 reads in 53 unique sequences.  
## Sample 9 - 96 reads in 46 unique sequences.  
## Sample 10 - 92 reads in 52 unique sequences.  
## Sample 11 - 97 reads in 51 unique sequences.  
## Sample 12 - 104 reads in 72 unique sequences.  
## Sample 13 - 102 reads in 54 unique sequences.  
## Sample 14 - 97 reads in 60 unique sequences.  
## Sample 15 - 101 reads in 75 unique sequences.  
## Sample 16 - 104 reads in 67 unique sequences.  
## Sample 17 - 91 reads in 36 unique sequences.  
## Sample 18 - 97 reads in 65 unique sequences.  
## Sample 19 - 105 reads in 66 unique sequences.  
## Sample 20 - 101 reads in 56 unique sequences.  
## Sample 21 - 100 reads in 77 unique sequences.  
## Sample 22 - 101 reads in 67 unique sequences.  
## Sample 23 - 98 reads in 52 unique sequences.  
## Sample 24 - 96 reads in 65 unique sequences.  
## Sample 25 - 95 reads in 50 unique sequences.  
## Sample 26 - 101 reads in 59 unique sequences.  
## Sample 27 - 91 reads in 63 unique sequences.  
## Sample 28 - 100 reads in 54 unique sequences.  
## Sample 29 - 98 reads in 40 unique sequences.  
## Sample 30 - 97 reads in 74 unique sequences.  
## Sample 31 - 89 reads in 76 unique sequences.  
## Sample 32 - 82 reads in 30 unique sequences.  
## Sample 33 - 104 reads in 73 unique sequences.  
## Sample 34 - 84 reads in 41 unique sequences.  
## Sample 35 - 110 reads in 78 unique sequences.  
## Sample 36 - 103 reads in 66 unique sequences.  
## Sample 37 - 108 reads in 79 unique sequences.  
## Sample 38 - 100 reads in 48 unique sequences.  
## Sample 39 - 101 reads in 83 unique sequences.  
## Sample 40 - 101 reads in 71 unique sequences.  
## Sample 41 - 96 reads in 40 unique sequences.  
## Sample 42 - 107 reads in 73 unique sequences.  
## Sample 43 - 99 reads in 71 unique sequences.  
## Sample 44 - 96 reads in 38 unique sequences.  
## Sample 45 - 106 reads in 74 unique sequences.  
## Sample 46 - 103 reads in 65 unique sequences.  
## Sample 47 - 89 reads in 51 unique sequences.  
## Sample 48 - 103 reads in 66 unique sequences.  
## Sample 49 - 97 reads in 66 unique sequences.  
## Sample 50 - 100 reads in 41 unique sequences.  
## Sample 51 - 108 reads in 73 unique sequences.  
## Sample 52 - 103 reads in 80 unique sequences.  
## Sample 53 - 98 reads in 78 unique sequences.  
## Sample 54 - 99 reads in 78 unique sequences.  
## Sample 55 - 90 reads in 62 unique sequences.  
## Sample 56 - 84 reads in 35 unique sequences.  
## selfConsist step 2   
## selfConsist step 3   
## Convergence after 3 rounds.  
## Total reads used: 5511

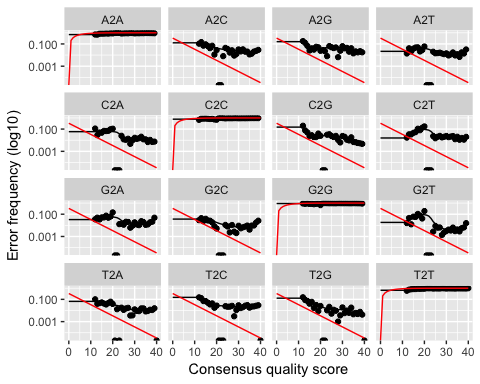
# Visualize the estimated error rates by plotting the forward and reverse reads  
  
plotErrors(errF, nominalQ=TRUE)

## Warning: Transformation introduced infinite values in continuous y-axis  
  
## Warning: Transformation introduced infinite values in continuous y-axis



plotErrors(errR, nominalQ = TRUE)

## Warning: Transformation introduced infinite values in continuous y-axis  
  
## Warning: Transformation introduced infinite values in continuous y-axis



### Dereplication

# Dereplicate  
  
derepFs <- derepFastq(filtFs, verbose=TRUE)

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0001\_F\_filt.fastq.gz

## Encountered 57 unique sequences from 99 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0002\_F\_filt.fastq.gz

## Encountered 49 unique sequences from 94 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0003\_F\_filt.fastq.gz

## Encountered 66 unique sequences from 103 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0004\_F\_filt.fastq.gz

## Encountered 52 unique sequences from 93 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0005\_F\_filt.fastq.gz

## Encountered 35 unique sequences from 99 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0006\_F\_filt.fastq.gz

## Encountered 34 unique sequences from 95 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0007\_F\_filt.fastq.gz

## Encountered 63 unique sequences from 105 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0008\_F\_filt.fastq.gz

## Encountered 46 unique sequences from 99 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0009\_F\_filt.fastq.gz

## Encountered 39 unique sequences from 96 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0010\_F\_filt.fastq.gz

## Encountered 58 unique sequences from 92 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0011\_F\_filt.fastq.gz

## Encountered 58 unique sequences from 97 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0012\_F\_filt.fastq.gz

## Encountered 61 unique sequences from 104 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0013\_F\_filt.fastq.gz

## Encountered 56 unique sequences from 102 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0014\_F\_filt.fastq.gz

## Encountered 55 unique sequences from 97 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0015\_F\_filt.fastq.gz

## Encountered 55 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0016\_F\_filt.fastq.gz

## Encountered 54 unique sequences from 104 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0017\_F\_filt.fastq.gz

## Encountered 47 unique sequences from 91 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0018\_F\_filt.fastq.gz

## Encountered 50 unique sequences from 97 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0019\_F\_filt.fastq.gz

## Encountered 61 unique sequences from 105 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0020\_F\_filt.fastq.gz

## Encountered 52 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0021\_F\_filt.fastq.gz

## Encountered 65 unique sequences from 100 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0022\_F\_filt.fastq.gz

## Encountered 55 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0023\_F\_filt.fastq.gz

## Encountered 37 unique sequences from 98 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0024\_F\_filt.fastq.gz

## Encountered 60 unique sequences from 96 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0025\_F\_filt.fastq.gz

## Encountered 54 unique sequences from 95 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0026\_F\_filt.fastq.gz

## Encountered 52 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0027\_F\_filt.fastq.gz

## Encountered 60 unique sequences from 91 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0028\_F\_filt.fastq.gz

## Encountered 58 unique sequences from 100 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0029\_F\_filt.fastq.gz

## Encountered 57 unique sequences from 98 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0030\_F\_filt.fastq.gz

## Encountered 69 unique sequences from 97 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0031\_F\_filt.fastq.gz

## Encountered 58 unique sequences from 89 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0032\_F\_filt.fastq.gz

## Encountered 48 unique sequences from 82 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0033\_F\_filt.fastq.gz

## Encountered 67 unique sequences from 104 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0034\_F\_filt.fastq.gz

## Encountered 50 unique sequences from 84 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0035\_F\_filt.fastq.gz

## Encountered 70 unique sequences from 110 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0036\_F\_filt.fastq.gz

## Encountered 62 unique sequences from 103 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0037\_F\_filt.fastq.gz

## Encountered 77 unique sequences from 108 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0038\_F\_filt.fastq.gz

## Encountered 58 unique sequences from 100 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0039\_F\_filt.fastq.gz

## Encountered 64 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0040\_F\_filt.fastq.gz

## Encountered 67 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0041\_F\_filt.fastq.gz

## Encountered 49 unique sequences from 96 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0042\_F\_filt.fastq.gz

## Encountered 62 unique sequences from 107 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0043\_F\_filt.fastq.gz

## Encountered 69 unique sequences from 99 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0044\_F\_filt.fastq.gz

## Encountered 50 unique sequences from 96 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0045\_F\_filt.fastq.gz

## Encountered 68 unique sequences from 106 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0046\_F\_filt.fastq.gz

## Encountered 59 unique sequences from 103 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0047\_F\_filt.fastq.gz

## Encountered 45 unique sequences from 89 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0048\_F\_filt.fastq.gz

## Encountered 64 unique sequences from 103 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0049\_F\_filt.fastq.gz

## Encountered 59 unique sequences from 97 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0050\_F\_filt.fastq.gz

## Encountered 41 unique sequences from 100 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0051\_F\_filt.fastq.gz

## Encountered 49 unique sequences from 108 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0052\_F\_filt.fastq.gz

## Encountered 69 unique sequences from 103 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0053\_F\_filt.fastq.gz

## Encountered 62 unique sequences from 98 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0054\_F\_filt.fastq.gz

## Encountered 71 unique sequences from 99 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0055\_F\_filt.fastq.gz

## Encountered 65 unique sequences from 90 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0056\_F\_filt.fastq.gz

## Encountered 38 unique sequences from 84 total sequences read.

derepRs <- derepFastq(filtRs, verbose=TRUE)

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0001\_R\_file.fastq.gz

## Encountered 61 unique sequences from 99 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0002\_R\_file.fastq.gz

## Encountered 55 unique sequences from 94 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0003\_R\_file.fastq.gz

## Encountered 74 unique sequences from 103 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0004\_R\_file.fastq.gz

## Encountered 52 unique sequences from 93 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0005\_R\_file.fastq.gz

## Encountered 49 unique sequences from 99 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0006\_R\_file.fastq.gz

## Encountered 41 unique sequences from 95 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0007\_R\_file.fastq.gz

## Encountered 89 unique sequences from 105 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0008\_R\_file.fastq.gz

## Encountered 53 unique sequences from 99 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0009\_R\_file.fastq.gz

## Encountered 46 unique sequences from 96 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0010\_R\_file.fastq.gz

## Encountered 52 unique sequences from 92 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0011\_R\_file.fastq.gz

## Encountered 51 unique sequences from 97 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0012\_R\_file.fastq.gz

## Encountered 72 unique sequences from 104 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0013\_R\_file.fastq.gz

## Encountered 54 unique sequences from 102 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0014\_R\_file.fastq.gz

## Encountered 60 unique sequences from 97 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0015\_R\_file.fastq.gz

## Encountered 75 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0016\_R\_file.fastq.gz

## Encountered 67 unique sequences from 104 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0017\_R\_file.fastq.gz

## Encountered 36 unique sequences from 91 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0018\_R\_file.fastq.gz

## Encountered 65 unique sequences from 97 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0019\_R\_file.fastq.gz

## Encountered 66 unique sequences from 105 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0020\_R\_file.fastq.gz

## Encountered 56 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0021\_R\_file.fastq.gz

## Encountered 77 unique sequences from 100 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0022\_R\_file.fastq.gz

## Encountered 67 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0023\_R\_file.fastq.gz

## Encountered 52 unique sequences from 98 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0024\_R\_file.fastq.gz

## Encountered 65 unique sequences from 96 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0025\_R\_file.fastq.gz

## Encountered 50 unique sequences from 95 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0026\_R\_file.fastq.gz

## Encountered 59 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0027\_R\_file.fastq.gz

## Encountered 63 unique sequences from 91 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0028\_R\_file.fastq.gz

## Encountered 54 unique sequences from 100 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0029\_R\_file.fastq.gz

## Encountered 40 unique sequences from 98 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0030\_R\_file.fastq.gz

## Encountered 74 unique sequences from 97 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0031\_R\_file.fastq.gz

## Encountered 76 unique sequences from 89 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0032\_R\_file.fastq.gz

## Encountered 30 unique sequences from 82 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0033\_R\_file.fastq.gz

## Encountered 73 unique sequences from 104 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0034\_R\_file.fastq.gz

## Encountered 41 unique sequences from 84 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0035\_R\_file.fastq.gz

## Encountered 78 unique sequences from 110 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0036\_R\_file.fastq.gz

## Encountered 66 unique sequences from 103 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0037\_R\_file.fastq.gz

## Encountered 79 unique sequences from 108 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0038\_R\_file.fastq.gz

## Encountered 48 unique sequences from 100 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0039\_R\_file.fastq.gz

## Encountered 83 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0040\_R\_file.fastq.gz

## Encountered 71 unique sequences from 101 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0041\_R\_file.fastq.gz

## Encountered 40 unique sequences from 96 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0042\_R\_file.fastq.gz

## Encountered 73 unique sequences from 107 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0043\_R\_file.fastq.gz

## Encountered 71 unique sequences from 99 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0044\_R\_file.fastq.gz

## Encountered 38 unique sequences from 96 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0045\_R\_file.fastq.gz

## Encountered 74 unique sequences from 106 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0046\_R\_file.fastq.gz

## Encountered 65 unique sequences from 103 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0047\_R\_file.fastq.gz

## Encountered 51 unique sequences from 89 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0048\_R\_file.fastq.gz

## Encountered 66 unique sequences from 103 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0049\_R\_file.fastq.gz

## Encountered 66 unique sequences from 97 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0050\_R\_file.fastq.gz

## Encountered 41 unique sequences from 100 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0051\_R\_file.fastq.gz

## Encountered 73 unique sequences from 108 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0052\_R\_file.fastq.gz

## Encountered 80 unique sequences from 103 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0053\_R\_file.fastq.gz

## Encountered 78 unique sequences from 98 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0054\_R\_file.fastq.gz

## Encountered 78 unique sequences from 99 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0055\_R\_file.fastq.gz

## Encountered 62 unique sequences from 90 total sequences read.

## Dereplicating sequence entries in Fastq file: ~/Desktop/N741/2018Week7/AWHONN Fastq Files/filtered/S0013-0056\_R\_file.fastq.gz

## Encountered 35 unique sequences from 84 total sequences read.

# Name the derep-class objects by the sample names  
  
names(derepFs) <- sample.names  
names(derepRs) <- sample.names

### Sample Inference

Infer the sequence variants in each sample (second dada pass)

# First with the Forward reads  
  
dadaFs <- dada(derepFs, err = errF, multithread = TRUE)

## Sample 1 - 99 reads in 57 unique sequences.  
## Sample 2 - 94 reads in 49 unique sequences.  
## Sample 3 - 103 reads in 66 unique sequences.  
## Sample 4 - 93 reads in 52 unique sequences.  
## Sample 5 - 99 reads in 35 unique sequences.  
## Sample 6 - 95 reads in 34 unique sequences.  
## Sample 7 - 105 reads in 63 unique sequences.  
## Sample 8 - 99 reads in 46 unique sequences.  
## Sample 9 - 96 reads in 39 unique sequences.  
## Sample 10 - 92 reads in 58 unique sequences.  
## Sample 11 - 97 reads in 58 unique sequences.  
## Sample 12 - 104 reads in 61 unique sequences.  
## Sample 13 - 102 reads in 56 unique sequences.  
## Sample 14 - 97 reads in 55 unique sequences.  
## Sample 15 - 101 reads in 55 unique sequences.  
## Sample 16 - 104 reads in 54 unique sequences.  
## Sample 17 - 91 reads in 47 unique sequences.  
## Sample 18 - 97 reads in 50 unique sequences.  
## Sample 19 - 105 reads in 61 unique sequences.  
## Sample 20 - 101 reads in 52 unique sequences.  
## Sample 21 - 100 reads in 65 unique sequences.  
## Sample 22 - 101 reads in 55 unique sequences.  
## Sample 23 - 98 reads in 37 unique sequences.  
## Sample 24 - 96 reads in 60 unique sequences.  
## Sample 25 - 95 reads in 54 unique sequences.  
## Sample 26 - 101 reads in 52 unique sequences.  
## Sample 27 - 91 reads in 60 unique sequences.  
## Sample 28 - 100 reads in 58 unique sequences.  
## Sample 29 - 98 reads in 57 unique sequences.  
## Sample 30 - 97 reads in 69 unique sequences.  
## Sample 31 - 89 reads in 58 unique sequences.  
## Sample 32 - 82 reads in 48 unique sequences.  
## Sample 33 - 104 reads in 67 unique sequences.  
## Sample 34 - 84 reads in 50 unique sequences.  
## Sample 35 - 110 reads in 70 unique sequences.  
## Sample 36 - 103 reads in 62 unique sequences.  
## Sample 37 - 108 reads in 77 unique sequences.  
## Sample 38 - 100 reads in 58 unique sequences.  
## Sample 39 - 101 reads in 64 unique sequences.  
## Sample 40 - 101 reads in 67 unique sequences.  
## Sample 41 - 96 reads in 49 unique sequences.  
## Sample 42 - 107 reads in 62 unique sequences.  
## Sample 43 - 99 reads in 69 unique sequences.  
## Sample 44 - 96 reads in 50 unique sequences.  
## Sample 45 - 106 reads in 68 unique sequences.  
## Sample 46 - 103 reads in 59 unique sequences.  
## Sample 47 - 89 reads in 45 unique sequences.  
## Sample 48 - 103 reads in 64 unique sequences.  
## Sample 49 - 97 reads in 59 unique sequences.  
## Sample 50 - 100 reads in 41 unique sequences.  
## Sample 51 - 108 reads in 49 unique sequences.  
## Sample 52 - 103 reads in 69 unique sequences.  
## Sample 53 - 98 reads in 62 unique sequences.  
## Sample 54 - 99 reads in 71 unique sequences.  
## Sample 55 - 90 reads in 65 unique sequences.  
## Sample 56 - 84 reads in 38 unique sequences.

# Then with the Reverse reads  
  
dadaRs <- dada(derepRs, err = errR, multithread = TRUE)

## Sample 1 - 99 reads in 61 unique sequences.  
## Sample 2 - 94 reads in 55 unique sequences.  
## Sample 3 - 103 reads in 74 unique sequences.  
## Sample 4 - 93 reads in 52 unique sequences.  
## Sample 5 - 99 reads in 49 unique sequences.  
## Sample 6 - 95 reads in 41 unique sequences.  
## Sample 7 - 105 reads in 89 unique sequences.  
## Sample 8 - 99 reads in 53 unique sequences.  
## Sample 9 - 96 reads in 46 unique sequences.  
## Sample 10 - 92 reads in 52 unique sequences.  
## Sample 11 - 97 reads in 51 unique sequences.  
## Sample 12 - 104 reads in 72 unique sequences.  
## Sample 13 - 102 reads in 54 unique sequences.  
## Sample 14 - 97 reads in 60 unique sequences.  
## Sample 15 - 101 reads in 75 unique sequences.  
## Sample 16 - 104 reads in 67 unique sequences.  
## Sample 17 - 91 reads in 36 unique sequences.  
## Sample 18 - 97 reads in 65 unique sequences.  
## Sample 19 - 105 reads in 66 unique sequences.  
## Sample 20 - 101 reads in 56 unique sequences.  
## Sample 21 - 100 reads in 77 unique sequences.  
## Sample 22 - 101 reads in 67 unique sequences.  
## Sample 23 - 98 reads in 52 unique sequences.  
## Sample 24 - 96 reads in 65 unique sequences.  
## Sample 25 - 95 reads in 50 unique sequences.  
## Sample 26 - 101 reads in 59 unique sequences.  
## Sample 27 - 91 reads in 63 unique sequences.  
## Sample 28 - 100 reads in 54 unique sequences.  
## Sample 29 - 98 reads in 40 unique sequences.  
## Sample 30 - 97 reads in 74 unique sequences.  
## Sample 31 - 89 reads in 76 unique sequences.  
## Sample 32 - 82 reads in 30 unique sequences.  
## Sample 33 - 104 reads in 73 unique sequences.  
## Sample 34 - 84 reads in 41 unique sequences.  
## Sample 35 - 110 reads in 78 unique sequences.  
## Sample 36 - 103 reads in 66 unique sequences.  
## Sample 37 - 108 reads in 79 unique sequences.  
## Sample 38 - 100 reads in 48 unique sequences.  
## Sample 39 - 101 reads in 83 unique sequences.  
## Sample 40 - 101 reads in 71 unique sequences.  
## Sample 41 - 96 reads in 40 unique sequences.  
## Sample 42 - 107 reads in 73 unique sequences.  
## Sample 43 - 99 reads in 71 unique sequences.  
## Sample 44 - 96 reads in 38 unique sequences.  
## Sample 45 - 106 reads in 74 unique sequences.  
## Sample 46 - 103 reads in 65 unique sequences.  
## Sample 47 - 89 reads in 51 unique sequences.  
## Sample 48 - 103 reads in 66 unique sequences.  
## Sample 49 - 97 reads in 66 unique sequences.  
## Sample 50 - 100 reads in 41 unique sequences.  
## Sample 51 - 108 reads in 73 unique sequences.  
## Sample 52 - 103 reads in 80 unique sequences.  
## Sample 53 - 98 reads in 78 unique sequences.  
## Sample 54 - 99 reads in 78 unique sequences.  
## Sample 55 - 90 reads in 62 unique sequences.  
## Sample 56 - 84 reads in 35 unique sequences.

# Inspect the dada-class objects returned by the dada function  
  
dadaFs[[1]]

## dada-class: object describing DADA2 denoising results  
## 6 sample sequences were inferred from 57 input unique sequences.  
## Key parameters: OMEGA\_A = 1e-40, BAND\_SIZE = 16, USE\_QUALS = TRUE

dadaRs[[1]]

## dada-class: object describing DADA2 denoising results  
## 9 sample sequences were inferred from 61 input unique sequences.  
## Key parameters: OMEGA\_A = 1e-40, BAND\_SIZE = 16, USE\_QUALS = TRUE

We can see that the algorithm has inferred 6 unique sequence variants from the forward reads and 8 from the reverse reads.

### Merge Paired Reads

We can eliminate further spurious sequence variants by merging overlapping reads. The core function is mergePairs and it depends on the forward and reverse reads being in matching order at the time they were dereplicated.

# Merge the denoised forward and reverse reads  
  
mergers <- mergePairs(dadaFs, derepFs, dadaRs, derepRs, verbose = TRUE )

## 74 paired-reads (in 6 unique pairings) successfully merged out of 99 (in 12 pairings) input.

## 78 paired-reads (in 3 unique pairings) successfully merged out of 94 (in 6 pairings) input.

## 62 paired-reads (in 4 unique pairings) successfully merged out of 103 (in 9 pairings) input.

## 74 paired-reads (in 5 unique pairings) successfully merged out of 93 (in 9 pairings) input.

## 99 paired-reads (in 5 unique pairings) successfully merged out of 99 (in 5 pairings) input.

## 94 paired-reads (in 5 unique pairings) successfully merged out of 95 (in 6 pairings) input.

## 48 paired-reads (in 4 unique pairings) successfully merged out of 105 (in 20 pairings) input.

## 95 paired-reads (in 7 unique pairings) successfully merged out of 99 (in 10 pairings) input.

## 94 paired-reads (in 6 unique pairings) successfully merged out of 96 (in 7 pairings) input.

## 86 paired-reads (in 5 unique pairings) successfully merged out of 92 (in 7 pairings) input.

## 83 paired-reads (in 8 unique pairings) successfully merged out of 97 (in 13 pairings) input.

## 85 paired-reads (in 7 unique pairings) successfully merged out of 104 (in 15 pairings) input.

## 96 paired-reads (in 4 unique pairings) successfully merged out of 102 (in 7 pairings) input.

## 75 paired-reads (in 7 unique pairings) successfully merged out of 97 (in 13 pairings) input.

## 88 paired-reads (in 5 unique pairings) successfully merged out of 101 (in 9 pairings) input.

## 85 paired-reads (in 10 unique pairings) successfully merged out of 104 (in 16 pairings) input.

## 87 paired-reads (in 6 unique pairings) successfully merged out of 91 (in 7 pairings) input.

## 74 paired-reads (in 5 unique pairings) successfully merged out of 97 (in 11 pairings) input.

## 91 paired-reads (in 8 unique pairings) successfully merged out of 105 (in 14 pairings) input.

## 101 paired-reads (in 4 unique pairings) successfully merged out of 101 (in 4 pairings) input.

## 73 paired-reads (in 7 unique pairings) successfully merged out of 100 (in 14 pairings) input.

## 72 paired-reads (in 6 unique pairings) successfully merged out of 101 (in 11 pairings) input.

## 89 paired-reads (in 3 unique pairings) successfully merged out of 98 (in 6 pairings) input.

## 77 paired-reads (in 10 unique pairings) successfully merged out of 96 (in 14 pairings) input.

## 86 paired-reads (in 8 unique pairings) successfully merged out of 95 (in 11 pairings) input.

## 75 paired-reads (in 5 unique pairings) successfully merged out of 101 (in 10 pairings) input.

## 51 paired-reads (in 7 unique pairings) successfully merged out of 91 (in 12 pairings) input.

## 97 paired-reads (in 6 unique pairings) successfully merged out of 100 (in 8 pairings) input.

## 94 paired-reads (in 3 unique pairings) successfully merged out of 98 (in 5 pairings) input.

## 70 paired-reads (in 7 unique pairings) successfully merged out of 97 (in 13 pairings) input.

## 49 paired-reads (in 4 unique pairings) successfully merged out of 89 (in 9 pairings) input.

## 79 paired-reads (in 2 unique pairings) successfully merged out of 82 (in 3 pairings) input.

## 62 paired-reads (in 8 unique pairings) successfully merged out of 104 (in 16 pairings) input.

## 74 paired-reads (in 3 unique pairings) successfully merged out of 84 (in 5 pairings) input.

## 87 paired-reads (in 9 unique pairings) successfully merged out of 110 (in 18 pairings) input.

## 75 paired-reads (in 4 unique pairings) successfully merged out of 103 (in 10 pairings) input.

## 91 paired-reads (in 9 unique pairings) successfully merged out of 108 (in 16 pairings) input.

## 94 paired-reads (in 3 unique pairings) successfully merged out of 100 (in 5 pairings) input.

## 67 paired-reads (in 5 unique pairings) successfully merged out of 101 (in 14 pairings) input.

## 69 paired-reads (in 8 unique pairings) successfully merged out of 101 (in 14 pairings) input.

## 88 paired-reads (in 2 unique pairings) successfully merged out of 96 (in 6 pairings) input.

## 100 paired-reads (in 11 unique pairings) successfully merged out of 107 (in 15 pairings) input.

## 67 paired-reads (in 9 unique pairings) successfully merged out of 99 (in 20 pairings) input.

## 88 paired-reads (in 5 unique pairings) successfully merged out of 96 (in 6 pairings) input.

## 76 paired-reads (in 6 unique pairings) successfully merged out of 106 (in 11 pairings) input.

## 33 paired-reads (in 3 unique pairings) successfully merged out of 103 (in 9 pairings) input.

## 78 paired-reads (in 4 unique pairings) successfully merged out of 89 (in 7 pairings) input.

## 77 paired-reads (in 6 unique pairings) successfully merged out of 103 (in 9 pairings) input.

## 82 paired-reads (in 6 unique pairings) successfully merged out of 97 (in 10 pairings) input.

## 98 paired-reads (in 3 unique pairings) successfully merged out of 100 (in 4 pairings) input.

## 98 paired-reads (in 10 unique pairings) successfully merged out of 108 (in 16 pairings) input.

## 79 paired-reads (in 8 unique pairings) successfully merged out of 103 (in 16 pairings) input.

## 56 paired-reads (in 6 unique pairings) successfully merged out of 98 (in 14 pairings) input.

## 86 paired-reads (in 8 unique pairings) successfully merged out of 99 (in 12 pairings) input.

## 75 paired-reads (in 3 unique pairings) successfully merged out of 90 (in 5 pairings) input.

## 78 paired-reads (in 2 unique pairings) successfully merged out of 84 (in 4 pairings) input.

# Inspect the merged data.frame from the first sample  
  
head(mergers[[1]])

## sequence  
## 1 TACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGATAGGTCAGTCTGTCTTAAAAGTTCGGGGCTTAACCCCGTGATGGGATGGAAACTGCCAATCTAGAGTATCGGAGAGGAAAGTGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAAGAACACCAGTGGCGAAGGCGACTTTCTGGACGAAAACTGACGCTGAGGCGCGAAAGCCAGGGGAGCGAACGGG  
## 3 TACGTAGGTCCCGAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTAGATAAGTCTGAAGTTAAAGGCTGTGGCTTAACCATAGTATGCTTTGGAAACTGTTTAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGATATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAAACAGG  
## 4 TACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCGTAAAGCGGGCGCAGACGGTTACTTAAGCAGGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCGTTCTGAACTGGGTAACTAGAGTGTGTCAGAGGGAGGTAGAATTCCACGTGTAGCAGTGAAATGCGTAGAGATGTGGAGGAATACCGATGGCGAAGGCAGCCTCCTGGGATAACACTGACGTTCATGCCCGAAAGCGTGGGTAGCAAACAGG  
## 5 TACGTAGGGTGCGAGCGTTGTCCGGAATTACTGGGCGTAAAGAGCTCGTAGGTGGTTTGTTGCGTCGTCTGTGAAATTCCGGGGCTTAACTTCGGGGTGGCAGGCGATACGGGCATAACTAGAGTGCTGTAGGGGAGACTGGAATTCCTGGTGTAGCGGTGGAATGCGCAGATATCAGGAGGAACACCGATGGCGAAGGCAGGTCTCTGGGCAGTAACTGACGCTGAGGAGCGAAAGCATGGGGAGCGAACAGG  
## 6 TACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGATCAGTCAGTCTGTCTTAAAAGTTCGGGGCTTAACCCCGTGATGGGATGGAAACTGCTGATCTAGAGTATCGGAGAGGAAAGTGGAATTCCTAGTGTAGCGGTGAAATGCGTAGATATTAGGAAGAACACCAGTGGCGAAGGCGACTTTCTGGACGAAAACTGACGCTGAGGCGCGAAAGCCAGGGGAGCGAACGGG  
## 7 TACGGAGGGTGCGAGCGTTAATCGGAATAACTGGGCGTAAAGGGCACGCAGGCGGTGACTTAAGTGAGGTGTGAAAGCCCCGGGCTTAACCTGGGAATTGCATTTCATACTGGGTCGCTAGAGTACTTTAGGGAGGGGTAGAATTCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAATACCGAAGGCGAAGGCAGCCCCTTGGGAATGTACTGACGCTCATGTGCGAAAGCGTGGGGAGCAAACAGG  
## abundance forward reverse nmatch nmismatch nindel prefer accept  
## 1 23 1 1 138 0 0 1 TRUE  
## 3 14 4 6 138 0 0 2 TRUE  
## 4 12 2 3 137 0 0 1 TRUE  
## 5 11 3 8 136 0 0 1 TRUE  
## 6 9 5 9 138 0 0 1 TRUE  
## 7 5 6 4 137 0 0 1 TRUE

### Sequence Table Construction

We will now construct the sequence table, this being analogous to the “OTU table” produced by other methods.

# Construct sequence table  
  
seqtab <- makeSequenceTable(mergers)

## The sequences being tabled vary in length.

# Consider the table  
  
dim(seqtab)

## [1] 56 135

class(seqtab)

## [1] "matrix"

# Inspect the distribution of sequence lengths  
  
table(nchar(getSequences(seqtab)))

##   
## 252 253 254   
## 39 93 3

### Remove Chimeras

# Remove chimeric sequences  
  
seqtab.nochim <- removeBimeraDenovo(seqtab, method = "consensus", multithread = TRUE, verbose=TRUE)

## Identified 0 bimeras out of 135 input sequences.

dim(seqtab.nochim)

## [1] 56 135

sum(seqtab.nochim)/sum(seqtab)

## [1] 1

### Track Reads through the Pipeline

getN <- function(x) sum(getUniques(x))  
pctSurv <- rowSums(seqtab.nochim)\*100/out[,1]  
track <- cbind(out, sapply(dadaFs, getN), sapply(mergers, getN), rowSums(seqtab), rowSums(seqtab.nochim), pctSurv)  
colnames(track) <- c("input", "filtered", "denoised", "merged", "tabled", "nonchimeric", "% passing")  
rownames(track) <- sample.names  
head(track)

## input filtered denoised merged tabled nonchimeric % passing  
## S0013-0001 114 99 99 74 74 74 64.91228  
## S0013-0002 114 94 94 78 78 78 68.42105  
## S0013-0003 114 103 103 62 62 62 54.38596  
## S0013-0004 114 93 93 74 74 74 64.91228  
## S0013-0005 114 99 99 99 99 99 86.84211  
## S0013-0006 114 95 95 94 94 94 82.45614

### Assign Taxonomy

GreenGenes 13\_8 reference will be used.

# Assign taxonomy  
  
# First initialize random number generator for reproducibility  
  
set.seed(100)  
getwd()

## [1] "/Users/ireneyang/Desktop/N741/Final-Project/Milestone-2"

path

## [1] "~/Desktop/N741/2018Week7/AWHONN Fastq Files"

list.files(path)

## [1] "filtered"   
## [2] "gg\_13\_8\_train\_set\_97.fa"   
## [3] "S0013-0001\_S97\_L001\_R1\_001.fastq"   
## [4] "S0013-0001\_S97\_L001\_R2\_001.fastq"   
## [5] "S0013-0002\_S98\_L001\_R1\_001.fastq"   
## [6] "S0013-0002\_S98\_L001\_R2\_001.fastq"   
## [7] "S0013-0003\_S99\_L001\_R1\_001.fastq"   
## [8] "S0013-0003\_S99\_L001\_R2\_001.fastq"   
## [9] "S0013-0004\_S100\_L001\_R1\_001.fastq"  
## [10] "S0013-0004\_S100\_L001\_R2\_001.fastq"  
## [11] "S0013-0005\_S101\_L001\_R1\_001.fastq"  
## [12] "S0013-0005\_S101\_L001\_R2\_001.fastq"  
## [13] "S0013-0006\_S102\_L001\_R1\_001.fastq"  
## [14] "S0013-0006\_S102\_L001\_R2\_001.fastq"  
## [15] "S0013-0007\_S103\_L001\_R1\_001.fastq"  
## [16] "S0013-0007\_S103\_L001\_R2\_001.fastq"  
## [17] "S0013-0008\_S104\_L001\_R1\_001.fastq"  
## [18] "S0013-0008\_S104\_L001\_R2\_001.fastq"  
## [19] "S0013-0009\_S105\_L001\_R1\_001.fastq"  
## [20] "S0013-0009\_S105\_L001\_R2\_001.fastq"  
## [21] "S0013-0010\_S106\_L001\_R1\_001.fastq"  
## [22] "S0013-0010\_S106\_L001\_R2\_001.fastq"  
## [23] "S0013-0011\_S107\_L001\_R1\_001.fastq"  
## [24] "S0013-0011\_S107\_L001\_R2\_001.fastq"  
## [25] "S0013-0012\_S108\_L001\_R1\_001.fastq"  
## [26] "S0013-0012\_S108\_L001\_R2\_001.fastq"  
## [27] "S0013-0013\_S109\_L001\_R1\_001.fastq"  
## [28] "S0013-0013\_S109\_L001\_R2\_001.fastq"  
## [29] "S0013-0014\_S110\_L001\_R1\_001.fastq"  
## [30] "S0013-0014\_S110\_L001\_R2\_001.fastq"  
## [31] "S0013-0015\_S111\_L001\_R1\_001.fastq"  
## [32] "S0013-0015\_S111\_L001\_R2\_001.fastq"  
## [33] "S0013-0016\_S112\_L001\_R1\_001.fastq"  
## [34] "S0013-0016\_S112\_L001\_R2\_001.fastq"  
## [35] "S0013-0017\_S113\_L001\_R1\_001.fastq"  
## [36] "S0013-0017\_S113\_L001\_R2\_001.fastq"  
## [37] "S0013-0018\_S114\_L001\_R1\_001.fastq"  
## [38] "S0013-0018\_S114\_L001\_R2\_001.fastq"  
## [39] "S0013-0019\_S115\_L001\_R1\_001.fastq"  
## [40] "S0013-0019\_S115\_L001\_R2\_001.fastq"  
## [41] "S0013-0020\_S116\_L001\_R1\_001.fastq"  
## [42] "S0013-0020\_S116\_L001\_R2\_001.fastq"  
## [43] "S0013-0021\_S117\_L001\_R1\_001.fastq"  
## [44] "S0013-0021\_S117\_L001\_R2\_001.fastq"  
## [45] "S0013-0022\_S118\_L001\_R1\_001.fastq"  
## [46] "S0013-0022\_S118\_L001\_R2\_001.fastq"  
## [47] "S0013-0023\_S119\_L001\_R1\_001.fastq"  
## [48] "S0013-0023\_S119\_L001\_R2\_001.fastq"  
## [49] "S0013-0024\_S120\_L001\_R1\_001.fastq"  
## [50] "S0013-0024\_S120\_L001\_R2\_001.fastq"  
## [51] "S0013-0025\_S121\_L001\_R1\_001.fastq"  
## [52] "S0013-0025\_S121\_L001\_R2\_001.fastq"  
## [53] "S0013-0026\_S122\_L001\_R1\_001.fastq"  
## [54] "S0013-0026\_S122\_L001\_R2\_001.fastq"  
## [55] "S0013-0027\_S123\_L001\_R1\_001.fastq"  
## [56] "S0013-0027\_S123\_L001\_R2\_001.fastq"  
## [57] "S0013-0028\_S124\_L001\_R1\_001.fastq"  
## [58] "S0013-0028\_S124\_L001\_R2\_001.fastq"  
## [59] "S0013-0029\_S125\_L001\_R1\_001.fastq"  
## [60] "S0013-0029\_S125\_L001\_R2\_001.fastq"  
## [61] "S0013-0030\_S126\_L001\_R1\_001.fastq"  
## [62] "S0013-0030\_S126\_L001\_R2\_001.fastq"  
## [63] "S0013-0031\_S127\_L001\_R1\_001.fastq"  
## [64] "S0013-0031\_S127\_L001\_R2\_001.fastq"  
## [65] "S0013-0032\_S128\_L001\_R1\_001.fastq"  
## [66] "S0013-0032\_S128\_L001\_R2\_001.fastq"  
## [67] "S0013-0033\_S129\_L001\_R1\_001.fastq"  
## [68] "S0013-0033\_S129\_L001\_R2\_001.fastq"  
## [69] "S0013-0034\_S130\_L001\_R1\_001.fastq"  
## [70] "S0013-0034\_S130\_L001\_R2\_001.fastq"  
## [71] "S0013-0035\_S131\_L001\_R1\_001.fastq"  
## [72] "S0013-0035\_S131\_L001\_R2\_001.fastq"  
## [73] "S0013-0036\_S132\_L001\_R1\_001.fastq"  
## [74] "S0013-0036\_S132\_L001\_R2\_001.fastq"  
## [75] "S0013-0037\_S133\_L001\_R1\_001.fastq"  
## [76] "S0013-0037\_S133\_L001\_R2\_001.fastq"  
## [77] "S0013-0038\_S134\_L001\_R1\_001.fastq"  
## [78] "S0013-0038\_S134\_L001\_R2\_001.fastq"  
## [79] "S0013-0039\_S135\_L001\_R1\_001.fastq"  
## [80] "S0013-0039\_S135\_L001\_R2\_001.fastq"  
## [81] "S0013-0040\_S136\_L001\_R1\_001.fastq"  
## [82] "S0013-0040\_S136\_L001\_R2\_001.fastq"  
## [83] "S0013-0041\_S137\_L001\_R1\_001.fastq"  
## [84] "S0013-0041\_S137\_L001\_R2\_001.fastq"  
## [85] "S0013-0042\_S138\_L001\_R1\_001.fastq"  
## [86] "S0013-0042\_S138\_L001\_R2\_001.fastq"  
## [87] "S0013-0043\_S139\_L001\_R1\_001.fastq"  
## [88] "S0013-0043\_S139\_L001\_R2\_001.fastq"  
## [89] "S0013-0044\_S140\_L001\_R1\_001.fastq"  
## [90] "S0013-0044\_S140\_L001\_R2\_001.fastq"  
## [91] "S0013-0045\_S141\_L001\_R1\_001.fastq"  
## [92] "S0013-0045\_S141\_L001\_R2\_001.fastq"  
## [93] "S0013-0046\_S142\_L001\_R1\_001.fastq"  
## [94] "S0013-0046\_S142\_L001\_R2\_001.fastq"  
## [95] "S0013-0047\_S143\_L001\_R1\_001.fastq"  
## [96] "S0013-0047\_S143\_L001\_R2\_001.fastq"  
## [97] "S0013-0048\_S144\_L001\_R1\_001.fastq"  
## [98] "S0013-0048\_S144\_L001\_R2\_001.fastq"  
## [99] "S0013-0049\_S145\_L001\_R1\_001.fastq"  
## [100] "S0013-0049\_S145\_L001\_R2\_001.fastq"  
## [101] "S0013-0050\_S146\_L001\_R1\_001.fastq"  
## [102] "S0013-0050\_S146\_L001\_R2\_001.fastq"  
## [103] "S0013-0051\_S147\_L001\_R1\_001.fastq"  
## [104] "S0013-0051\_S147\_L001\_R2\_001.fastq"  
## [105] "S0013-0052\_S148\_L001\_R1\_001.fastq"  
## [106] "S0013-0052\_S148\_L001\_R2\_001.fastq"  
## [107] "S0013-0053\_S149\_L001\_R1\_001.fastq"  
## [108] "S0013-0053\_S149\_L001\_R2\_001.fastq"  
## [109] "S0013-0054\_S150\_L001\_R1\_001.fastq"  
## [110] "S0013-0054\_S150\_L001\_R2\_001.fastq"  
## [111] "S0013-0055\_S151\_L001\_R1\_001.fastq"  
## [112] "S0013-0055\_S151\_L001\_R2\_001.fastq"  
## [113] "S0013-0056\_S152\_L001\_R1\_001.fastq"  
## [114] "S0013-0056\_S152\_L001\_R2\_001.fastq"

taxa <- assignTaxonomy(seqtab.nochim, "~/Desktop/N741/2018Week7/AWHONN Fastq Files/gg\_13\_8\_train\_set\_97.fa", multithread = TRUE)  
unname(head(taxa))

## [,1] [,2] [,3]   
## [1,] "k\_\_Bacteria" "p\_\_Firmicutes" "c\_\_Bacilli"   
## [2,] "k\_\_Bacteria" "p\_\_Firmicutes" "c\_\_Bacilli"   
## [3,] "k\_\_Bacteria" "p\_\_Firmicutes" "c\_\_Clostridia"   
## [4,] "k\_\_Bacteria" "p\_\_Actinobacteria" "c\_\_Actinobacteria"  
## [5,] "k\_\_Bacteria" "p\_\_Fusobacteria" "c\_\_Fusobacteriia"   
## [6,] "k\_\_Bacteria" "p\_\_Fusobacteria" "c\_\_Fusobacteriia"   
## [,4] [,5] [,6]   
## [1,] "o\_\_Lactobacillales" "f\_\_Streptococcaceae" "g\_\_Streptococcus"   
## [2,] "o\_\_Lactobacillales" "f\_\_Streptococcaceae" "g\_\_Streptococcus"   
## [3,] "o\_\_Clostridiales" "f\_\_Veillonellaceae" "g\_\_Veillonella"   
## [4,] "o\_\_Actinomycetales" "f\_\_Corynebacteriaceae" "g\_\_Corynebacterium"  
## [5,] "o\_\_Fusobacteriales" "f\_\_Fusobacteriaceae" "g\_\_Fusobacterium"   
## [6,] "o\_\_Fusobacteriales" "f\_\_Fusobacteriaceae" "g\_\_Fusobacterium"   
## [,7]   
## [1,] "s\_\_"   
## [2,] "s\_\_"   
## [3,] "s\_\_dispar"  
## [4,] "s\_\_"   
## [5,] "s\_\_"   
## [6,] "s\_\_"

Inspect the taxonomic assignments:

taxa.print <- taxa #Removing sequence rownames for display only  
rownames (taxa.print) <- NULL  
head(taxa.print)

## Kingdom Phylum Class   
## [1,] "k\_\_Bacteria" "p\_\_Firmicutes" "c\_\_Bacilli"   
## [2,] "k\_\_Bacteria" "p\_\_Firmicutes" "c\_\_Bacilli"   
## [3,] "k\_\_Bacteria" "p\_\_Firmicutes" "c\_\_Clostridia"   
## [4,] "k\_\_Bacteria" "p\_\_Actinobacteria" "c\_\_Actinobacteria"  
## [5,] "k\_\_Bacteria" "p\_\_Fusobacteria" "c\_\_Fusobacteriia"   
## [6,] "k\_\_Bacteria" "p\_\_Fusobacteria" "c\_\_Fusobacteriia"   
## Order Family Genus   
## [1,] "o\_\_Lactobacillales" "f\_\_Streptococcaceae" "g\_\_Streptococcus"   
## [2,] "o\_\_Lactobacillales" "f\_\_Streptococcaceae" "g\_\_Streptococcus"   
## [3,] "o\_\_Clostridiales" "f\_\_Veillonellaceae" "g\_\_Veillonella"   
## [4,] "o\_\_Actinomycetales" "f\_\_Corynebacteriaceae" "g\_\_Corynebacterium"  
## [5,] "o\_\_Fusobacteriales" "f\_\_Fusobacteriaceae" "g\_\_Fusobacterium"   
## [6,] "o\_\_Fusobacteriales" "f\_\_Fusobacteriaceae" "g\_\_Fusobacterium"   
## Species   
## [1,] "s\_\_"   
## [2,] "s\_\_"   
## [3,] "s\_\_dispar"  
## [4,] "s\_\_"   
## [5,] "s\_\_"   
## [6,] "s\_\_"

### Construct a Phylogenetic Tree

library(DECIPHER)

## Loading required package: RSQLite

seqs <- getSequences(seqtab.nochim)  
  
# This next command will allow propagation of sequence names to the tip labels of the tree  
names(seqs) <- seqs  
alignment <- AlignSeqs(DNAStringSet(seqs), anchor=NA)

## Determining distance matrix based on shared 7-mers:  
##   
 |   
 | | 0%  
 |   
 |= | 1%  
 |   
 |= | 2%  
 |   
 |=== | 4%  
 |   
 |=== | 5%  
 |   
 |===== | 7%  
 |   
 |===== | 8%  
 |   
 |====== | 10%  
 |   
 |======= | 11%  
 |   
 |======== | 12%  
 |   
 |========= | 14%  
 |   
 |========== | 15%  
 |   
 |=========== | 17%  
 |   
 |============ | 18%  
 |   
 |============ | 19%  
 |   
 |============== | 21%  
 |   
 |============== | 22%  
 |   
 |=============== | 23%  
 |   
 |================ | 24%  
 |   
 |================= | 26%  
 |   
 |================== | 27%  
 |   
 |================== | 28%  
 |   
 |==================== | 30%  
 |   
 |==================== | 31%  
 |   
 |===================== | 32%  
 |   
 |===================== | 33%  
 |   
 |====================== | 34%  
 |   
 |======================= | 36%  
 |   
 |======================== | 37%  
 |   
 |========================= | 38%  
 |   
 |========================= | 39%  
 |   
 |========================== | 40%  
 |   
 |=========================== | 41%  
 |   
 |============================ | 43%  
 |   
 |============================= | 44%  
 |   
 |============================= | 45%  
 |   
 |============================== | 46%  
 |   
 |=============================== | 47%  
 |   
 |=============================== | 48%  
 |   
 |================================ | 49%  
 |   
 |================================ | 50%  
 |   
 |================================= | 51%  
 |   
 |================================== | 52%  
 |   
 |================================== | 53%  
 |   
 |=================================== | 54%  
 |   
 |==================================== | 55%  
 |   
 |==================================== | 56%  
 |   
 |===================================== | 57%  
 |   
 |====================================== | 58%  
 |   
 |====================================== | 59%  
 |   
 |======================================= | 60%  
 |   
 |======================================== | 61%  
 |   
 |======================================== | 62%  
 |   
 |========================================= | 63%  
 |   
 |========================================== | 64%  
 |   
 |========================================== | 65%  
 |   
 |=========================================== | 66%  
 |   
 |============================================ | 67%  
 |   
 |============================================ | 68%  
 |   
 |============================================= | 69%  
 |   
 |============================================== | 70%  
 |   
 |============================================== | 71%  
 |   
 |=============================================== | 72%  
 |   
 |=============================================== | 73%  
 |   
 |================================================ | 74%  
 |   
 |================================================= | 75%  
 |   
 |================================================= | 76%  
 |   
 |================================================== | 77%  
 |   
 |=================================================== | 78%  
 |   
 |=================================================== | 79%  
 |   
 |==================================================== | 80%  
 |   
 |===================================================== | 81%  
 |   
 |===================================================== | 82%  
 |   
 |====================================================== | 83%  
 |   
 |======================================================= | 84%  
 |   
 |======================================================= | 85%  
 |   
 |======================================================== | 86%  
 |   
 |========================================================= | 87%  
 |   
 |========================================================= | 88%  
 |   
 |========================================================== | 89%  
 |   
 |========================================================== | 90%  
 |   
 |=========================================================== | 91%  
 |   
 |============================================================ | 92%  
 |   
 |============================================================ | 93%  
 |   
 |============================================================= | 94%  
 |   
 |============================================================== | 95%  
 |   
 |============================================================== | 96%  
 |   
 |=============================================================== | 97%  
 |   
 |================================================================ | 98%  
 |   
 |================================================================ | 99%  
 |   
 |=================================================================| 100%  
##   
## Time difference of 0.38 secs  
##   
## Clustering into groups by similarity:  
##   
 |   
 | | 0%  
 |   
 |= | 1%  
 |   
 |= | 2%  
 |   
 |=== | 4%  
 |   
 |=== | 5%  
 |   
 |===== | 7%  
 |   
 |===== | 8%  
 |   
 |====== | 10%  
 |   
 |======= | 11%  
 |   
 |======== | 12%  
 |   
 |========= | 14%  
 |   
 |========== | 15%  
 |   
 |=========== | 17%  
 |   
 |============ | 18%  
 |   
 |============ | 19%  
 |   
 |============== | 21%  
 |   
 |============== | 22%  
 |   
 |=============== | 23%  
 |   
 |================ | 24%  
 |   
 |================= | 26%  
 |   
 |================== | 27%  
 |   
 |================== | 28%  
 |   
 |==================== | 30%  
 |   
 |==================== | 31%  
 |   
 |===================== | 32%  
 |   
 |===================== | 33%  
 |   
 |====================== | 34%  
 |   
 |======================= | 36%  
 |   
 |======================== | 37%  
 |   
 |========================= | 38%  
 |   
 |========================= | 39%  
 |   
 |========================== | 40%  
 |   
 |=========================== | 41%  
 |   
 |============================ | 43%  
 |   
 |============================= | 44%  
 |   
 |============================= | 45%  
 |   
 |============================== | 46%  
 |   
 |=============================== | 47%  
 |   
 |=============================== | 48%  
 |   
 |================================ | 49%  
 |   
 |================================ | 50%  
 |   
 |================================= | 51%  
 |   
 |================================== | 52%  
 |   
 |================================== | 53%  
 |   
 |=================================== | 54%  
 |   
 |==================================== | 55%  
 |   
 |==================================== | 56%  
 |   
 |===================================== | 57%  
 |   
 |====================================== | 58%  
 |   
 |====================================== | 59%  
 |   
 |======================================= | 60%  
 |   
 |======================================== | 61%  
 |   
 |======================================== | 62%  
 |   
 |========================================= | 63%  
 |   
 |========================================== | 64%  
 |   
 |========================================== | 65%  
 |   
 |=========================================== | 66%  
 |   
 |============================================ | 67%  
 |   
 |============================================ | 68%  
 |   
 |============================================= | 69%  
 |   
 |============================================== | 70%  
 |   
 |============================================== | 71%  
 |   
 |=============================================== | 72%  
 |   
 |=============================================== | 73%  
 |   
 |================================================ | 74%  
 |   
 |================================================= | 75%  
 |   
 |================================================= | 76%  
 |   
 |================================================== | 77%  
 |   
 |=================================================== | 78%  
 |   
 |=================================================== | 79%  
 |   
 |==================================================== | 80%  
 |   
 |===================================================== | 81%  
 |   
 |===================================================== | 82%  
 |   
 |====================================================== | 83%  
 |   
 |======================================================= | 84%  
 |   
 |======================================================= | 85%  
 |   
 |======================================================== | 86%  
 |   
 |========================================================= | 87%  
 |   
 |========================================================= | 88%  
 |   
 |========================================================== | 89%  
 |   
 |========================================================== | 90%  
 |   
 |=========================================================== | 91%  
 |   
 |============================================================ | 92%  
 |   
 |============================================================ | 93%  
 |   
 |============================================================= | 94%  
 |   
 |============================================================== | 95%  
 |   
 |============================================================== | 96%  
 |   
 |=============================================================== | 97%  
 |   
 |================================================================ | 98%  
 |   
 |================================================================ | 99%  
 |   
 |=================================================================| 100%  
##   
## Time difference of 0.17 secs  
##   
## Aligning Sequences:  
##   
 |   
 | | 0%  
 |   
 |= | 1%  
 |   
 |= | 2%  
 |   
 |== | 3%  
 |   
 |=== | 4%  
 |   
 |=== | 5%  
 |   
 |==== | 6%  
 |   
 |===== | 7%  
 |   
 |===== | 8%  
 |   
 |====== | 9%  
 |   
 |====== | 10%  
 |   
 |======= | 11%  
 |   
 |======== | 12%  
 |   
 |======== | 13%  
 |   
 |========= | 14%  
 |   
 |========== | 15%  
 |   
 |========== | 16%  
 |   
 |=========== | 17%  
 |   
 |============ | 18%  
 |   
 |============ | 19%  
 |   
 |============= | 20%  
 |   
 |============== | 21%  
 |   
 |============== | 22%  
 |   
 |=============== | 23%  
 |   
 |================ | 24%  
 |   
 |================ | 25%  
 |   
 |================= | 26%  
 |   
 |================== | 27%  
 |   
 |================== | 28%  
 |   
 |=================== | 29%  
 |   
 |==================== | 30%  
 |   
 |==================== | 31%  
 |   
 |===================== | 32%  
 |   
 |===================== | 33%  
 |   
 |====================== | 34%  
 |   
 |======================= | 35%  
 |   
 |======================= | 36%  
 |   
 |======================== | 37%  
 |   
 |========================= | 38%  
 |   
 |========================= | 39%  
 |   
 |========================== | 40%  
 |   
 |=========================== | 41%  
 |   
 |=========================== | 42%  
 |   
 |============================ | 43%  
 |   
 |============================= | 44%  
 |   
 |============================= | 45%  
 |   
 |============================== | 46%  
 |   
 |=============================== | 47%  
 |   
 |=============================== | 48%  
 |   
 |================================ | 49%  
 |   
 |================================ | 50%  
 |   
 |================================= | 51%  
 |   
 |================================== | 52%  
 |   
 |================================== | 53%  
 |   
 |=================================== | 54%  
 |   
 |==================================== | 55%  
 |   
 |==================================== | 56%  
 |   
 |===================================== | 57%  
 |   
 |====================================== | 58%  
 |   
 |====================================== | 59%  
 |   
 |======================================= | 60%  
 |   
 |======================================== | 61%  
 |   
 |======================================== | 62%  
 |   
 |========================================= | 63%  
 |   
 |========================================== | 64%  
 |   
 |========================================== | 65%  
 |   
 |=========================================== | 66%  
 |   
 |============================================ | 67%  
 |   
 |============================================ | 68%  
 |   
 |============================================= | 69%  
 |   
 |============================================== | 70%  
 |   
 |============================================== | 71%  
 |   
 |=============================================== | 72%  
 |   
 |=============================================== | 73%  
 |   
 |================================================ | 74%  
 |   
 |================================================= | 75%  
 |   
 |================================================= | 76%  
 |   
 |================================================== | 77%  
 |   
 |=================================================== | 78%  
 |   
 |=================================================== | 79%  
 |   
 |==================================================== | 80%  
 |   
 |===================================================== | 81%  
 |   
 |===================================================== | 82%  
 |   
 |====================================================== | 83%  
 |   
 |======================================================= | 84%  
 |   
 |======================================================= | 85%  
 |   
 |======================================================== | 86%  
 |   
 |========================================================= | 87%  
 |   
 |========================================================= | 88%  
 |   
 |========================================================== | 89%  
 |   
 |========================================================== | 90%  
 |   
 |=========================================================== | 91%  
 |   
 |============================================================ | 92%  
 |   
 |============================================================ | 93%  
 |   
 |============================================================= | 94%  
 |   
 |============================================================== | 95%  
 |   
 |============================================================== | 96%  
 |   
 |=============================================================== | 97%  
 |   
 |================================================================ | 98%  
 |   
 |================================================================ | 99%  
 |   
 |=================================================================| 100%  
##   
## Time difference of 2.7 secs  
##   
## Iteration 1 of 2:  
##   
## Determining distance matrix based on alignment:  
##   
 |   
 | | 0%  
 |   
 |= | 1%  
 |   
 |= | 2%  
 |   
 |=== | 4%  
 |   
 |=== | 5%  
 |   
 |===== | 7%  
 |   
 |===== | 8%  
 |   
 |====== | 10%  
 |   
 |======= | 11%  
 |   
 |======== | 12%  
 |   
 |========= | 14%  
 |   
 |========== | 15%  
 |   
 |=========== | 17%  
 |   
 |============ | 18%  
 |   
 |============ | 19%  
 |   
 |============== | 21%  
 |   
 |============== | 22%  
 |   
 |=============== | 23%  
 |   
 |================ | 24%  
 |   
 |================= | 26%  
 |   
 |================== | 27%  
 |   
 |================== | 28%  
 |   
 |==================== | 30%  
 |   
 |==================== | 31%  
 |   
 |===================== | 32%  
 |   
 |===================== | 33%  
 |   
 |====================== | 34%  
 |   
 |======================= | 36%  
 |   
 |======================== | 37%  
 |   
 |========================= | 38%  
 |   
 |========================= | 39%  
 |   
 |========================== | 40%  
 |   
 |=========================== | 41%  
 |   
 |============================ | 43%  
 |   
 |============================= | 44%  
 |   
 |============================= | 45%  
 |   
 |============================== | 46%  
 |   
 |=============================== | 47%  
 |   
 |=============================== | 48%  
 |   
 |================================ | 49%  
 |   
 |================================ | 50%  
 |   
 |================================= | 51%  
 |   
 |================================== | 52%  
 |   
 |================================== | 53%  
 |   
 |=================================== | 54%  
 |   
 |==================================== | 55%  
 |   
 |==================================== | 56%  
 |   
 |===================================== | 57%  
 |   
 |====================================== | 58%  
 |   
 |====================================== | 59%  
 |   
 |======================================= | 60%  
 |   
 |======================================== | 61%  
 |   
 |======================================== | 62%  
 |   
 |========================================= | 63%  
 |   
 |========================================== | 64%  
 |   
 |========================================== | 65%  
 |   
 |=========================================== | 66%  
 |   
 |============================================ | 67%  
 |   
 |============================================ | 68%  
 |   
 |============================================= | 69%  
 |   
 |============================================== | 70%  
 |   
 |============================================== | 71%  
 |   
 |=============================================== | 72%  
 |   
 |=============================================== | 73%  
 |   
 |================================================ | 74%  
 |   
 |================================================= | 75%  
 |   
 |================================================= | 76%  
 |   
 |================================================== | 77%  
 |   
 |=================================================== | 78%  
 |   
 |=================================================== | 79%  
 |   
 |==================================================== | 80%  
 |   
 |===================================================== | 81%  
 |   
 |===================================================== | 82%  
 |   
 |====================================================== | 83%  
 |   
 |======================================================= | 84%  
 |   
 |======================================================= | 85%  
 |   
 |======================================================== | 86%  
 |   
 |========================================================= | 87%  
 |   
 |========================================================= | 88%  
 |   
 |========================================================== | 89%  
 |   
 |========================================================== | 90%  
 |   
 |=========================================================== | 91%  
 |   
 |============================================================ | 92%  
 |   
 |============================================================ | 93%  
 |   
 |============================================================= | 94%  
 |   
 |============================================================== | 95%  
 |   
 |============================================================== | 96%  
 |   
 |=============================================================== | 97%  
 |   
 |================================================================ | 98%  
 |   
 |================================================================ | 99%  
 |   
 |=================================================================| 100%  
##   
## Time difference of 0.04 secs  
##   
## Reclustering into groups by similarity:  
##   
 |   
 | | 0%  
 |   
 |= | 1%  
 |   
 |= | 2%  
 |   
 |=== | 4%  
 |   
 |=== | 5%  
 |   
 |===== | 7%  
 |   
 |===== | 8%  
 |   
 |====== | 10%  
 |   
 |======= | 11%  
 |   
 |======== | 12%  
 |   
 |========= | 14%  
 |   
 |========== | 15%  
 |   
 |=========== | 17%  
 |   
 |============ | 18%  
 |   
 |============ | 19%  
 |   
 |============== | 21%  
 |   
 |============== | 22%  
 |   
 |=============== | 23%  
 |   
 |================ | 24%  
 |   
 |================= | 26%  
 |   
 |================== | 27%  
 |   
 |================== | 28%  
 |   
 |==================== | 30%  
 |   
 |==================== | 31%  
 |   
 |===================== | 32%  
 |   
 |===================== | 33%  
 |   
 |====================== | 34%  
 |   
 |======================= | 36%  
 |   
 |======================== | 37%  
 |   
 |========================= | 38%  
 |   
 |========================= | 39%  
 |   
 |========================== | 40%  
 |   
 |=========================== | 41%  
 |   
 |============================ | 43%  
 |   
 |============================= | 44%  
 |   
 |============================= | 45%  
 |   
 |============================== | 46%  
 |   
 |=============================== | 47%  
 |   
 |=============================== | 48%  
 |   
 |================================ | 49%  
 |   
 |================================ | 50%  
 |   
 |================================= | 51%  
 |   
 |================================== | 52%  
 |   
 |================================== | 53%  
 |   
 |=================================== | 54%  
 |   
 |==================================== | 55%  
 |   
 |==================================== | 56%  
 |   
 |===================================== | 57%  
 |   
 |====================================== | 58%  
 |   
 |====================================== | 59%  
 |   
 |======================================= | 60%  
 |   
 |======================================== | 61%  
 |   
 |======================================== | 62%  
 |   
 |========================================= | 63%  
 |   
 |========================================== | 64%  
 |   
 |========================================== | 65%  
 |   
 |=========================================== | 66%  
 |   
 |============================================ | 67%  
 |   
 |============================================ | 68%  
 |   
 |============================================= | 69%  
 |   
 |============================================== | 70%  
 |   
 |============================================== | 71%  
 |   
 |=============================================== | 72%  
 |   
 |=============================================== | 73%  
 |   
 |================================================ | 74%  
 |   
 |================================================= | 75%  
 |   
 |================================================= | 76%  
 |   
 |================================================== | 77%  
 |   
 |=================================================== | 78%  
 |   
 |=================================================== | 79%  
 |   
 |==================================================== | 80%  
 |   
 |===================================================== | 81%  
 |   
 |===================================================== | 82%  
 |   
 |====================================================== | 83%  
 |   
 |======================================================= | 84%  
 |   
 |======================================================= | 85%  
 |   
 |======================================================== | 86%  
 |   
 |========================================================= | 87%  
 |   
 |========================================================= | 88%  
 |   
 |========================================================== | 89%  
 |   
 |========================================================== | 90%  
 |   
 |=========================================================== | 91%  
 |   
 |============================================================ | 92%  
 |   
 |============================================================ | 93%  
 |   
 |============================================================= | 94%  
 |   
 |============================================================== | 95%  
 |   
 |============================================================== | 96%  
 |   
 |=============================================================== | 97%  
 |   
 |================================================================ | 98%  
 |   
 |================================================================ | 99%  
 |   
 |=================================================================| 100%  
##   
## Time difference of 0.24 secs  
##   
## Realigning Sequences:  
##   
 |   
 | | 0%  
 |   
 |= | 1%  
 |   
 |= | 2%  
 |   
 |== | 3%  
 |   
 |=== | 5%  
 |   
 |==== | 6%  
 |   
 |===== | 7%  
 |   
 |===== | 8%  
 |   
 |====== | 10%  
 |   
 |======= | 11%  
 |   
 |======== | 12%  
 |   
 |========= | 14%  
 |   
 |========== | 15%  
 |   
 |========== | 16%  
 |   
 |=========== | 17%  
 |   
 |============ | 19%  
 |   
 |============= | 20%  
 |   
 |============== | 21%  
 |   
 |=============== | 23%  
 |   
 |================ | 24%  
 |   
 |================ | 25%  
 |   
 |================= | 26%  
 |   
 |================== | 28%  
 |   
 |=================== | 29%  
 |   
 |==================== | 30%  
 |   
 |===================== | 32%  
 |   
 |===================== | 33%  
 |   
 |====================== | 34%  
 |   
 |======================= | 35%  
 |   
 |======================== | 37%  
 |   
 |========================= | 38%  
 |   
 |========================= | 39%  
 |   
 |=========================== | 41%  
 |   
 |=========================== | 42%  
 |   
 |============================ | 43%  
 |   
 |============================= | 44%  
 |   
 |============================== | 46%  
 |   
 |=============================== | 47%  
 |   
 |=============================== | 48%  
 |   
 |================================ | 50%  
 |   
 |================================= | 51%  
 |   
 |================================== | 52%  
 |   
 |================================== | 53%  
 |   
 |==================================== | 55%  
 |   
 |==================================== | 56%  
 |   
 |===================================== | 57%  
 |   
 |====================================== | 58%  
 |   
 |======================================= | 60%  
 |   
 |======================================== | 61%  
 |   
 |======================================== | 62%  
 |   
 |========================================== | 64%  
 |   
 |========================================== | 65%  
 |   
 |=========================================== | 66%  
 |   
 |============================================ | 67%  
 |   
 |============================================= | 69%  
 |   
 |============================================== | 70%  
 |   
 |============================================== | 71%  
 |   
 |=============================================== | 73%  
 |   
 |================================================ | 74%  
 |   
 |================================================= | 75%  
 |   
 |================================================= | 76%  
 |   
 |=================================================== | 78%  
 |   
 |=================================================== | 79%  
 |   
 |==================================================== | 80%  
 |   
 |===================================================== | 82%  
 |   
 |====================================================== | 83%  
 |   
 |======================================================= | 84%  
 |   
 |======================================================= | 85%  
 |   
 |========================================================= | 87%  
 |   
 |========================================================= | 88%  
 |   
 |========================================================== | 89%  
 |   
 |=========================================================== | 91%  
 |   
 |============================================================ | 92%  
 |   
 |============================================================ | 93%  
 |   
 |============================================================= | 94%  
 |   
 |============================================================== | 96%  
 |   
 |=================================================================| 100%  
##   
## Time difference of 1.09 secs  
##   
## Iteration 2 of 2:  
##   
## Determining distance matrix based on alignment:  
##   
 |   
 | | 0%  
 |   
 |= | 1%  
 |   
 |= | 2%  
 |   
 |=== | 4%  
 |   
 |=== | 5%  
 |   
 |===== | 7%  
 |   
 |===== | 8%  
 |   
 |====== | 10%  
 |   
 |======= | 11%  
 |   
 |======== | 12%  
 |   
 |========= | 14%  
 |   
 |========== | 15%  
 |   
 |=========== | 17%  
 |   
 |============ | 18%  
 |   
 |============ | 19%  
 |   
 |============== | 21%  
 |   
 |============== | 22%  
 |   
 |=============== | 23%  
 |   
 |================ | 24%  
 |   
 |================= | 26%  
 |   
 |================== | 27%  
 |   
 |================== | 28%  
 |   
 |==================== | 30%  
 |   
 |==================== | 31%  
 |   
 |===================== | 32%  
 |   
 |===================== | 33%  
 |   
 |====================== | 34%  
 |   
 |======================= | 36%  
 |   
 |======================== | 37%  
 |   
 |========================= | 38%  
 |   
 |========================= | 39%  
 |   
 |========================== | 40%  
 |   
 |=========================== | 41%  
 |   
 |============================ | 43%  
 |   
 |============================= | 44%  
 |   
 |============================= | 45%  
 |   
 |============================== | 46%  
 |   
 |=============================== | 47%  
 |   
 |=============================== | 48%  
 |   
 |================================ | 49%  
 |   
 |================================ | 50%  
 |   
 |================================= | 51%  
 |   
 |================================== | 52%  
 |   
 |================================== | 53%  
 |   
 |=================================== | 54%  
 |   
 |==================================== | 55%  
 |   
 |==================================== | 56%  
 |   
 |===================================== | 57%  
 |   
 |====================================== | 58%  
 |   
 |====================================== | 59%  
 |   
 |======================================= | 60%  
 |   
 |======================================== | 61%  
 |   
 |======================================== | 62%  
 |   
 |========================================= | 63%  
 |   
 |========================================== | 64%  
 |   
 |========================================== | 65%  
 |   
 |=========================================== | 66%  
 |   
 |============================================ | 67%  
 |   
 |============================================ | 68%  
 |   
 |============================================= | 69%  
 |   
 |============================================== | 70%  
 |   
 |============================================== | 71%  
 |   
 |=============================================== | 72%  
 |   
 |=============================================== | 73%  
 |   
 |================================================ | 74%  
 |   
 |================================================= | 75%  
 |   
 |================================================= | 76%  
 |   
 |================================================== | 77%  
 |   
 |=================================================== | 78%  
 |   
 |=================================================== | 79%  
 |   
 |==================================================== | 80%  
 |   
 |===================================================== | 81%  
 |   
 |===================================================== | 82%  
 |   
 |====================================================== | 83%  
 |   
 |======================================================= | 84%  
 |   
 |======================================================= | 85%  
 |   
 |======================================================== | 86%  
 |   
 |========================================================= | 87%  
 |   
 |========================================================= | 88%  
 |   
 |========================================================== | 89%  
 |   
 |========================================================== | 90%  
 |   
 |=========================================================== | 91%  
 |   
 |============================================================ | 92%  
 |   
 |============================================================ | 93%  
 |   
 |============================================================= | 94%  
 |   
 |============================================================== | 95%  
 |   
 |============================================================== | 96%  
 |   
 |=============================================================== | 97%  
 |   
 |================================================================ | 98%  
 |   
 |================================================================ | 99%  
 |   
 |=================================================================| 100%  
##   
## Time difference of 0.03 secs  
##   
## Reclustering into groups by similarity:  
##   
 |   
 | | 0%  
 |   
 |= | 1%  
 |   
 |= | 2%  
 |   
 |=== | 4%  
 |   
 |=== | 5%  
 |   
 |===== | 7%  
 |   
 |===== | 8%  
 |   
 |====== | 10%  
 |   
 |======= | 11%  
 |   
 |======== | 12%  
 |   
 |========= | 14%  
 |   
 |========== | 15%  
 |   
 |=========== | 17%  
 |   
 |============ | 18%  
 |   
 |============ | 19%  
 |   
 |============== | 21%  
 |   
 |============== | 22%  
 |   
 |=============== | 23%  
 |   
 |================ | 24%  
 |   
 |================= | 26%  
 |   
 |================== | 27%  
 |   
 |================== | 28%  
 |   
 |==================== | 30%  
 |   
 |==================== | 31%  
 |   
 |===================== | 32%  
 |   
 |===================== | 33%  
 |   
 |====================== | 34%  
 |   
 |======================= | 36%  
 |   
 |======================== | 37%  
 |   
 |========================= | 38%  
 |   
 |========================= | 39%  
 |   
 |========================== | 40%  
 |   
 |=========================== | 41%  
 |   
 |============================ | 43%  
 |   
 |============================= | 44%  
 |   
 |============================= | 45%  
 |   
 |============================== | 46%  
 |   
 |=============================== | 47%  
 |   
 |=============================== | 48%  
 |   
 |================================ | 49%  
 |   
 |================================ | 50%  
 |   
 |================================= | 51%  
 |   
 |================================== | 52%  
 |   
 |================================== | 53%  
 |   
 |=================================== | 54%  
 |   
 |==================================== | 55%  
 |   
 |==================================== | 56%  
 |   
 |===================================== | 57%  
 |   
 |====================================== | 58%  
 |   
 |====================================== | 59%  
 |   
 |======================================= | 60%  
 |   
 |======================================== | 61%  
 |   
 |======================================== | 62%  
 |   
 |========================================= | 63%  
 |   
 |========================================== | 64%  
 |   
 |========================================== | 65%  
 |   
 |=========================================== | 66%  
 |   
 |============================================ | 67%  
 |   
 |============================================ | 68%  
 |   
 |============================================= | 69%  
 |   
 |============================================== | 70%  
 |   
 |============================================== | 71%  
 |   
 |=============================================== | 72%  
 |   
 |=============================================== | 73%  
 |   
 |================================================ | 74%  
 |   
 |================================================= | 75%  
 |   
 |================================================= | 76%  
 |   
 |================================================== | 77%  
 |   
 |=================================================== | 78%  
 |   
 |=================================================== | 79%  
 |   
 |==================================================== | 80%  
 |   
 |===================================================== | 81%  
 |   
 |===================================================== | 82%  
 |   
 |====================================================== | 83%  
 |   
 |======================================================= | 84%  
 |   
 |======================================================= | 85%  
 |   
 |======================================================== | 86%  
 |   
 |========================================================= | 87%  
 |   
 |========================================================= | 88%  
 |   
 |========================================================== | 89%  
 |   
 |========================================================== | 90%  
 |   
 |=========================================================== | 91%  
 |   
 |============================================================ | 92%  
 |   
 |============================================================ | 93%  
 |   
 |============================================================= | 94%  
 |   
 |============================================================== | 95%  
 |   
 |============================================================== | 96%  
 |   
 |=============================================================== | 97%  
 |   
 |================================================================ | 98%  
 |   
 |================================================================ | 99%  
 |   
 |=================================================================| 100%  
##   
## Time difference of 0.16 secs  
##   
## Realigning Sequences:  
##   
 |   
 | | 0%  
 |   
 |================ | 25%  
 |   
 |=================================================================| 100%  
##   
## Time difference of 0.14 secs

# Construct tree  
  
library(phangorn)

## Loading required package: ape

##   
## Attaching package: 'ape'

## The following object is masked from 'package:ShortRead':  
##   
## zoom

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

phang.align <- phyDat(as(alignment, "matrix"), type="DNA")  
dm <- dist.ml(phang.align)  
treeNJ <- NJ(dm) # Tip order will not equal sequence order  
fit <- pml(treeNJ, data=phang.align)

## negative edges length changed to 0!

## negative edges length changed to 0.  
  
fitGTR <- update(fit, k=4, inv=0.2)  
fitGTR <- optim.pml(fitGTR, model="GTR", optInv=TRUE, optGamma=TRUE,   
 rearrangement = "stochastic", control=pml.control(trace=0))  
detach("package:phangorn", unload=TRUE)

### Handoff to phyloseq

Our next activity will be to hand off the data to the phyloseq package for analysis. This package requires three items: the “OTUtable,” the taxonomy table, and data about the samples. The first two items are directly available at the end of your dada2run, and you can import the latter as a .csv file.

# Import metadata file.  
  
samdf <- read.csv("~/Desktop/N741/2018Week7/Metadata.csv",header=TRUE)  
  
rownames(samdf) <- samdf$Sample\_ID  
  
rownames(samdf)

## [1] "S0013-0001" "S0013-0002" "S0013-0003" "S0013-0004" "S0013-0005"  
## [6] "S0013-0006" "S0013-0007" "S0013-0008" "S0013-0009" "S0013-0010"  
## [11] "S0013-0011" "S0013-0012" "S0013-0013" "S0013-0014" "S0013-0015"  
## [16] "S0013-0016" "S0013-0017" "S0013-0018" "S0013-0019" "S0013-0020"  
## [21] "S0013-0021" "S0013-0022" "S0013-0023" "S0013-0024" "S0013-0025"  
## [26] "S0013-0026" "S0013-0027" "S0013-0028" "S0013-0029" "S0013-0030"  
## [31] "S0013-0031" "S0013-0032" "S0013-0033" "S0013-0034" "S0013-0035"  
## [36] "S0013-0036" "S0013-0037" "S0013-0038" "S0013-0039" "S0013-0040"  
## [41] "S0013-0041" "S0013-0042" "S0013-0043" "S0013-0044" "S0013-0045"  
## [46] "S0013-0046" "S0013-0047" "S0013-0048" "S0013-0049" "S0013-0050"  
## [51] "S0013-0051" "S0013-0052" "S0013-0053" "S0013-0054" "S0013-0055"  
## [56] "S0013-0056"

rownames(seqtab.nochim)

## [1] "S0013-0001" "S0013-0002" "S0013-0003" "S0013-0004" "S0013-0005"  
## [6] "S0013-0006" "S0013-0007" "S0013-0008" "S0013-0009" "S0013-0010"  
## [11] "S0013-0011" "S0013-0012" "S0013-0013" "S0013-0014" "S0013-0015"  
## [16] "S0013-0016" "S0013-0017" "S0013-0018" "S0013-0019" "S0013-0020"  
## [21] "S0013-0021" "S0013-0022" "S0013-0023" "S0013-0024" "S0013-0025"  
## [26] "S0013-0026" "S0013-0027" "S0013-0028" "S0013-0029" "S0013-0030"  
## [31] "S0013-0031" "S0013-0032" "S0013-0033" "S0013-0034" "S0013-0035"  
## [36] "S0013-0036" "S0013-0037" "S0013-0038" "S0013-0039" "S0013-0040"  
## [41] "S0013-0041" "S0013-0042" "S0013-0043" "S0013-0044" "S0013-0045"  
## [46] "S0013-0046" "S0013-0047" "S0013-0048" "S0013-0049" "S0013-0050"  
## [51] "S0013-0051" "S0013-0052" "S0013-0053" "S0013-0054" "S0013-0055"  
## [56] "S0013-0056"

Create the phyloseq object.

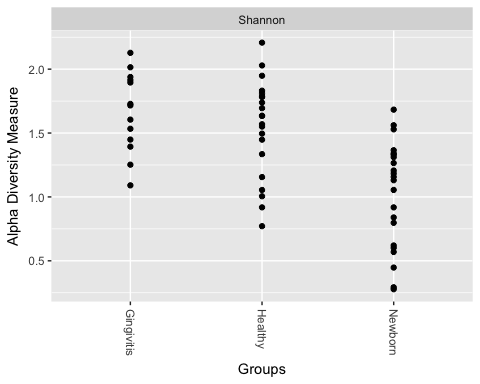
library(phyloseq)  
  
# Create phyloseq object  
  
ps <- phyloseq(otu\_table(seqtab.nochim, taxa\_are\_rows=FALSE),   
 sample\_data(samdf),  
 tax\_table(taxa),  
 phy\_tree(fitGTR$tree))  
  
# Describe it  
  
ps

## phyloseq-class experiment-level object  
## otu\_table() OTU Table: [ 135 taxa and 56 samples ]  
## sample\_data() Sample Data: [ 56 samples by 6 sample variables ]  
## tax\_table() Taxonomy Table: [ 135 taxa by 7 taxonomic ranks ]  
## phy\_tree() Phylogenetic Tree: [ 135 tips and 133 internal nodes ]

### Diversity in Microbial Ecology

# Plot alpha-diversity  
  
plot\_richness(ps, x="Groups", measures = c("Shannon"))

## Warning in estimate\_richness(physeq, split = TRUE, measures = measures): The data you have provided does not have  
## any singletons. This is highly suspicious. Results of richness  
## estimates (for example) are probably unreliable, or wrong, if you have already  
## trimmed low-abundance taxa from the data.  
##   
## We recommended that you find the un-trimmed data and retry.



theme\_bw()

## List of 57  
## $ line :List of 6  
## ..$ colour : chr "black"  
## ..$ size : num 0.5  
## ..$ linetype : num 1  
## ..$ lineend : chr "butt"  
## ..$ arrow : logi FALSE  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_line" "element"  
## $ rect :List of 5  
## ..$ fill : chr "white"  
## ..$ colour : chr "black"  
## ..$ size : num 0.5  
## ..$ linetype : num 1  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ text :List of 11  
## ..$ family : chr ""  
## ..$ face : chr "plain"  
## ..$ colour : chr "black"  
## ..$ size : num 11  
## ..$ hjust : num 0.5  
## ..$ vjust : num 0.5  
## ..$ angle : num 0  
## ..$ lineheight : num 0.9  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 0 0 0 0  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : logi FALSE  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.title.x :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 1  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 5.5 0 0 0  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.title.x.top :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 0  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 0 0 5.5 0  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.title.y :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 1  
## ..$ angle : num 90  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 0 5.5 0 0  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.title.y.right :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 0  
## ..$ angle : num -90  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 0 0 0 5.5  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.text :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : chr "grey30"  
## ..$ size :Class 'rel' num 0.8  
## ..$ hjust : NULL  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : NULL  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.text.x :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 1  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 2.2 0 0 0  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.text.x.top :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : num 0  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 0 0 2.2 0  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.text.y :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : num 1  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 0 2.2 0 0  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.text.y.right :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : num 0  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 0 0 0 2.2  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ axis.ticks :List of 6  
## ..$ colour : chr "grey20"  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ lineend : NULL  
## ..$ arrow : logi FALSE  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_line" "element"  
## $ axis.ticks.length :Class 'unit' atomic [1:1] 2.75  
## .. ..- attr(\*, "valid.unit")= int 8  
## .. ..- attr(\*, "unit")= chr "pt"  
## $ axis.line : list()  
## ..- attr(\*, "class")= chr [1:2] "element\_blank" "element"  
## $ axis.line.x : NULL  
## $ axis.line.y : NULL  
## $ legend.background :List of 5  
## ..$ fill : NULL  
## ..$ colour : logi NA  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ legend.margin :Classes 'margin', 'unit' atomic [1:4] 0.2 0.2 0.2 0.2  
## .. ..- attr(\*, "valid.unit")= int 1  
## .. ..- attr(\*, "unit")= chr "cm"  
## $ legend.spacing :Class 'unit' atomic [1:1] 0.4  
## .. ..- attr(\*, "valid.unit")= int 1  
## .. ..- attr(\*, "unit")= chr "cm"  
## $ legend.spacing.x : NULL  
## $ legend.spacing.y : NULL  
## $ legend.key :List of 5  
## ..$ fill : chr "white"  
## ..$ colour : logi NA  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ legend.key.size :Class 'unit' atomic [1:1] 1.2  
## .. ..- attr(\*, "valid.unit")= int 3  
## .. ..- attr(\*, "unit")= chr "lines"  
## $ legend.key.height : NULL  
## $ legend.key.width : NULL  
## $ legend.text :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size :Class 'rel' num 0.8  
## ..$ hjust : NULL  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : NULL  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ legend.text.align : NULL  
## $ legend.title :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : num 0  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : NULL  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ legend.title.align : NULL  
## $ legend.position : chr "right"  
## $ legend.direction : NULL  
## $ legend.justification : chr "center"  
## $ legend.box : NULL  
## $ legend.box.margin :Classes 'margin', 'unit' atomic [1:4] 0 0 0 0  
## .. ..- attr(\*, "valid.unit")= int 1  
## .. ..- attr(\*, "unit")= chr "cm"  
## $ legend.box.background: list()  
## ..- attr(\*, "class")= chr [1:2] "element\_blank" "element"  
## $ legend.box.spacing :Class 'unit' atomic [1:1] 0.4  
## .. ..- attr(\*, "valid.unit")= int 1  
## .. ..- attr(\*, "unit")= chr "cm"  
## $ panel.background :List of 5  
## ..$ fill : chr "white"  
## ..$ colour : logi NA  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ panel.border :List of 5  
## ..$ fill : logi NA  
## ..$ colour : chr "grey20"  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ panel.spacing :Class 'unit' atomic [1:1] 5.5  
## .. ..- attr(\*, "valid.unit")= int 8  
## .. ..- attr(\*, "unit")= chr "pt"  
## $ panel.spacing.x : NULL  
## $ panel.spacing.y : NULL  
## $ panel.grid.major :List of 6  
## ..$ colour : chr "grey92"  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ lineend : NULL  
## ..$ arrow : logi FALSE  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_line" "element"  
## $ panel.grid.minor :List of 6  
## ..$ colour : chr "grey92"  
## ..$ size : num 0.25  
## ..$ linetype : NULL  
## ..$ lineend : NULL  
## ..$ arrow : logi FALSE  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_line" "element"  
## $ panel.ontop : logi FALSE  
## $ plot.background :List of 5  
## ..$ fill : NULL  
## ..$ colour : chr "white"  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ plot.title :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size :Class 'rel' num 1.2  
## ..$ hjust : num 0  
## ..$ vjust : num 1  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 0 0 6.6 0  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ plot.subtitle :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size :Class 'rel' num 0.9  
## ..$ hjust : num 0  
## ..$ vjust : num 1  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 0 0 4.95 0  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ plot.caption :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size :Class 'rel' num 0.9  
## ..$ hjust : num 1  
## ..$ vjust : num 1  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 4.95 0 0 0  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ plot.margin :Classes 'margin', 'unit' atomic [1:4] 5.5 5.5 5.5 5.5  
## .. ..- attr(\*, "valid.unit")= int 8  
## .. ..- attr(\*, "unit")= chr "pt"  
## $ strip.background :List of 5  
## ..$ fill : chr "grey85"  
## ..$ colour : chr "grey20"  
## ..$ size : NULL  
## ..$ linetype : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_rect" "element"  
## $ strip.placement : chr "inside"  
## $ strip.text :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : chr "grey10"  
## ..$ size :Class 'rel' num 0.8  
## ..$ hjust : NULL  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin : NULL  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ strip.text.x :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : NULL  
## ..$ angle : NULL  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 5.5 0 5.5 0  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ strip.text.y :List of 11  
## ..$ family : NULL  
## ..$ face : NULL  
## ..$ colour : NULL  
## ..$ size : NULL  
## ..$ hjust : NULL  
## ..$ vjust : NULL  
## ..$ angle : num -90  
## ..$ lineheight : NULL  
## ..$ margin :Classes 'margin', 'unit' atomic [1:4] 0 5.5 0 5.5  
## .. .. ..- attr(\*, "valid.unit")= int 8  
## .. .. ..- attr(\*, "unit")= chr "pt"  
## ..$ debug : NULL  
## ..$ inherit.blank: logi TRUE  
## ..- attr(\*, "class")= chr [1:2] "element\_text" "element"  
## $ strip.switch.pad.grid:Class 'unit' atomic [1:1] 0.1  
## .. ..- attr(\*, "valid.unit")= int 1  
## .. ..- attr(\*, "unit")= chr "cm"  
## $ strip.switch.pad.wrap:Class 'unit' atomic [1:1] 0.1  
## .. ..- attr(\*, "valid.unit")= int 1  
## .. ..- attr(\*, "unit")= chr "cm"  
## - attr(\*, "class")= chr [1:2] "theme" "gg"  
## - attr(\*, "complete")= logi TRUE  
## - attr(\*, "validate")= logi TRUE

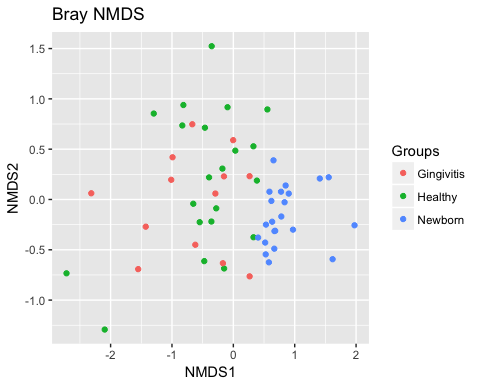
### Ordinate

Using the Bray-Curtis dissimilarity index.

# Ordinate with Bray-Curtis  
  
ord.nmds.bray <- ordinate(ps, method="NMDS", distance="bray")

## Square root transformation  
## Wisconsin double standardization  
## Run 0 stress 0.1356063   
## Run 1 stress 0.1347283   
## ... New best solution  
## ... Procrustes: rmse 0.08529231 max resid 0.3072408   
## Run 2 stress 0.1358782   
## Run 3 stress 0.1384943   
## Run 4 stress 0.1346348   
## ... New best solution  
## ... Procrustes: rmse 0.06332175 max resid 0.2678842   
## Run 5 stress 0.1354997   
## Run 6 stress 0.1391081   
## Run 7 stress 0.1397812   
## Run 8 stress 0.1337538   
## ... New best solution  
## ... Procrustes: rmse 0.05795525 max resid 0.2486489   
## Run 9 stress 0.1421732   
## Run 10 stress 0.1334337   
## ... New best solution  
## ... Procrustes: rmse 0.01949492 max resid 0.1075307   
## Run 11 stress 0.134563   
## Run 12 stress 0.1330346   
## ... New best solution  
## ... Procrustes: rmse 0.05368703 max resid 0.2401102   
## Run 13 stress 0.1351115   
## Run 14 stress 0.143663   
## Run 15 stress 0.1350092   
## Run 16 stress 0.1341607   
## Run 17 stress 0.14216   
## Run 18 stress 0.1392832   
## Run 19 stress 0.1336896   
## Run 20 stress 0.1370797   
## \*\*\* No convergence -- monoMDS stopping criteria:  
## 1: no. of iterations >= maxit  
## 19: stress ratio > sratmax

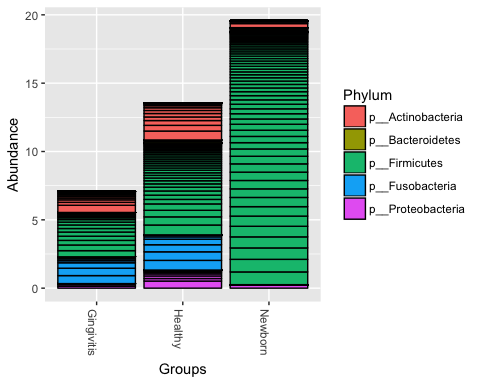
plot\_ordination(ps, ord.nmds.bray, color="Groups", title="Bray NMDS")



We see that ordination picks out a separation between maternal and newborn samples.

### Bar Plots

# Create bar plots for top 20 OTUs  
  
top20 <- names(sort(taxa\_sums(ps), decreasing = TRUE))[1:20]  
ps.top20 <- transform\_sample\_counts(ps, function(OTU) OTU/sum(OTU))  
ps.top20 <- prune\_taxa(top20, ps.top20)  
plot\_bar(ps.top20, x="Groups", fill="Phylum")



# Plot richness  
  
plot\_richness(ps, "Groups", "Sample\_Type")

## Warning in estimate\_richness(physeq, split = TRUE, measures = measures): The data you have provided does not have  
## any singletons. This is highly suspicious. Results of richness  
## estimates (for example) are probably unreliable, or wrong, if you have already  
## trimmed low-abundance taxa from the data.  
##   
## We recommended that you find the un-trimmed data and retry.

## Warning: Removed 286 rows containing missing values (geom\_errorbar).

