Setting up new Phyloseq Objects

Minuka Hewapathirana

2024-07-19

Contents

1 Loading Packages

A comprehensive list of packages required to perform metagenomics analysis for *ION torrent* data

```r
library("knitr")
library("qiime2R") # devtools::install_github("jbisanz/qiime2R")
library("phyloseq")
library("readxl")
library("tibble")
library("vegan")
library("DESeq2") #BiocManager::install("DESeq2")
library("speedyseq") # remotes::install_github("mikemc/speedyseq")
library("ape")
library("ggstar")
library("forcats")
library("patchwork")
library("ggpubr")
library("plotROC")
library("viridis")
library("cowplot")
library("ggplot2")
library("microbiome") # BiocManager::install("microbiome")
library("microbiomeutilities")

library("ggtree") # BiocManager::install("ggtree")
library("ggtreeExtra") #install.packages("ggExtra")
library('MicrobiotaProcess') # BiocManager::install("MicrobiotaProcess")
#library("tidytree")
library("file2meco")
library("microeco")
```

2   Converting Qiime Artifacts to Phyloseq Objects

Each **Phyloseq** object requires - features (table.qza) - tree (rooted-tree.qza) - taxonomy (taxonomy.qza) - metadata ("sample.tsv") : The metadata file shoudl be in .tsv format

```
#Creating a phyloseq object for the ITS1 Reverse Reads
PS_ITS1R_Global<-qza_to_phyloseq(
  features="D:\\Grad_School\\Grad_Project\\Qiime_Outputs\\ITS2_NEW\\table.qza",
  tree="D:\\Grad_School\\Grad_Project\\Qiime_Outputs\\ITS2_NEW\\rooted-tree.qza",
  taxonomy="D:\\Grad_School\\Grad_Project\\Qiime_Outputs\\ITS2_NEW\\taxonomy.qza",
  metadata = "D:\\Grad_School\\Grad_Project\\Qiime_Outputs\\ITS2_NEW\\ITS2_Global.tsv"
)


#Checking if objects exist
PS_ITS1R_Global
```

phyloseq-class experiment-level object otu_table() OTU Table: [11005 taxa and 251 samples]: sample_data() Sample Data: [251 samples by 15 sample variables]: tax_table() Taxonomy Table: [11005 taxa by 7 taxonomic ranks]: phy_tree() Phylogenetic Tree: [11005 tips and 10923 internal nodes]: taxa are rows

```
# Creating a phyloseq object for the 16S Reverse Reads
PS_16S_Global<-qza_to_phyloseq(
  features="D:\\Grad_School\\Grad_Project\\Qiime_Outputs\\New_Silva_Global\\table.qza",
  tree="D:\\Grad_School\\Grad_Project\\Qiime_Outputs\\New_Silva_Global\\rooted-tree.qza",
  taxonomy="D:\\Grad_School\\Grad_Project\\Qiime_Outputs\\New_Silva_Global\\taxonomy.qza",
  metadata = "D:\\Grad_School\\Grad_Project\\Qiime_Outputs\\New_Silva_Global\\Global.tsv"
)


PS_16S_Global
```

phyloseq-class experiment-level object otu_table() OTU Table: [11535 taxa and 254 samples]: sample_data() Sample Data: [254 samples by 14 sample variables]: tax_table() Taxonomy Table: [11535 taxa by 7 taxonomic ranks]: phy_tree() Phylogenetic Tree: [11535 tips and 11352 internal nodes]: taxa are rows

3   Obtaining StoneFruit ITS1 Datasets

Goal : Split ITS1 PS Object into two smaller PS objects based on StoneFruit and Blueberry in BC and ON

- First remove Alberta from ITS1R and 16S objects

- Then split based on Stonefruits and Blueberries respectively

```
#removal of AB province
PS_ITS1R_BC_ON <- subset_samples(PS_ITS1R_Global, Province !="AB")

#removal of Wild forage
PS_ITS1R_BC_ON_SF <- subset_samples(PS_ITS1R_BC_ON, Plant !="Wild forage")
#removal of Blueberry
PS_ITS1R_BC_ON_SF <- subset_samples(PS_ITS1R_BC_ON_SF, Plant !="Blueberry")
#PS Object with only Stonefruit in BC and ON
PS_ITS1R_BC_ON_SF
```

phyloseq-class experiment-level object otu_table() OTU Table: [11005 taxa and 138 samples]: sample_data() Sample Data: [138 samples by 15 sample variables]: tax_table() Taxonomy Table: [11005 taxa by 7 taxonomic ranks]: phy_tree() Phylogenetic Tree: [11005 tips and 10923 internal nodes]: taxa are rows

```
#Check metadata file for Stonefruit
metadata_PS_ITS1R_BC_ON_SF <- meta(PS_ITS1R_BC_ON_SF)

#For this example I showcase the first few entries of the 4,5,6,7 columns due to spacing issues
head(metadata_PS_ITS1R_BC_ON_SF[,c(4,5,6,7)])
```

	Description	RunNumber	SampleName	Number
A01	ITS2_A_B01	BCC_R1	BCCV1-AP-1B	1
A02	ITS2_A_B02	BCC_R1	BCCV1-AP-2B	2
A03	ITS2_A_B03	BCC_R1	BCCV1-AP-3B	3
A04	ITS2_A_B04	BCC_R1	BCCV2-AP-1B	4
A05	ITS2_A_B05	BCC_R1	BCCV2-AP-2B	5
A06	ITS2_A_B06	BCC_R1	BCCV2-AP-3B	6

```
#Remove other plant types to obtain Wild Forage and Bluberry subset
#Cherry, Apple , Peach , Apricot
PS_ITS1R_BC_ON_BB <-subset_samples(PS_ITS1R_BC_ON, Plant !="Cherry")
PS_ITS1R_BC_ON_BB <-subset_samples(PS_ITS1R_BC_ON_BB, Plant !="Apple")
PS_ITS1R_BC_ON_BB <- subset_samples(PS_ITS1R_BC_ON_BB, Plant !="Peach")
PS_ITS1R_BC_ON_BB <- subset_samples(PS_ITS1R_BC_ON_BB, Plant !="Apricot")

PS_ITS1R_BC_ON_BB
```

phyloseq-class experiment-level object otu_table() OTU Table: [11005 taxa and 82 samples]: sample_data() Sample Data: [82 samples by 15 sample variables]: tax_table() Taxonomy Table: [11005 taxa by 7 taxonomic ranks]: phy_tree() Phylogenetic Tree: [11005 tips and 10923 internal nodes]: taxa are rows

```
#Check metadatafile for blueberry
metadata_ITS1R_BC_ON_BB <- meta(PS_ITS1R_BC_ON_BB)

head(metadata_ITS1R_BC_ON_BB[,c(4,5,6,7)])
```

	Description	RunNumber	SampleName	Number
A10	ITS2_A_B10	BCC_R1	BCFV1-BB-1B	10
A11	ITS2_A_B11	BCC_R1	BCFV1-BB-2B	11
A12	ITS2_A_B12	BCC_R1	BCFV1-BB-3B	12
A13	ITS2_A_B13	BCC_R1	BCFV1-BB-4B	13
A14	ITS2_A_B14	BCC_R1	BCFV2-BB-1B	14
A15	ITS2_A_B15	BCC_R1	BCFV2-BB-2B	15

4   Obtaining Blueberry 16S Datasets

Goal : Repeat Step 3 for 16S

```
#removal of AB province
PS_16S_BC_ON <- subset_samples(PS_16S_Global, Province !="AB")

#meta(PS_16S_BC_ON)

#Stone fruit subset
PS_16S_BC_ON_SF <- subset_samples(PS_16S_BC_ON, Plant !="Wild forage")
PS_16S_BC_ON_SF <- subset_samples(PS_16S_BC_ON_SF, Plant !="Blueberry")
```

```
PS_16S_BC_ON_SF <- subset_samples(PS_16S_BC_ON_SF, Plant !="Kit")

PS_16S_BC_ON_SF
```

phyloseq-class experiment-level object otu_table() OTU Table: [11535 taxa and 140 samples]: sample_data() Sample Data: [140 samples by 14 sample variables]: tax_table() Taxonomy Table: [11535 taxa by 7 taxonomic ranks]: phy_tree() Phylogenetic Tree: [11535 tips and 11352 internal nodes]: taxa are rows

```
#meta(PS_16S_BC_ON_SF)

#Remove these to obtain Wild Forage and Bluberry subset
# Cherry, Apple , Peach , Apricot
PS_16S_BC_ON_BB <-subset_samples(PS_16S_BC_ON, Plant !="Cherry")
PS_16S_BC_ON_BB <-subset_samples(PS_16S_BC_ON_BB, Plant !="Apple")
PS_16S_BC_ON_BB <- subset_samples(PS_16S_BC_ON_BB, Plant !="Peach")
PS_16S_BC_ON_BB <- subset_samples(PS_16S_BC_ON_BB, Plant !="Apricot")
PS_16S_BC_ON_BB <- subset_samples(PS_16S_BC_ON_BB, Plant !="Kit")

PS_16S_BC_ON_BB
```

phyloseq-class experiment-level object otu_table() OTU Table: [11535 taxa and 82 samples]: sample_data() Sample Data: [82 samples by 14 sample variables]: tax_table() Taxonomy Table: [11535 taxa by 7 taxonomic ranks]: phy_tree() Phylogenetic Tree: [11535 tips and 11352 internal nodes]: taxa are rows

```
#meta(PS_16S_BC_ON_BB)
```

We have obtained **4 PS objects** that will be used in the following chapters to perorm metagenomics based visualizations and develop insights

UP NEXT - Chapter 2 : Metagenomics Analysis of Fungal (ITS1) and bacterial (16S) data to investigate agricultural and invasive pathogen detections in **Stonefruit Systems**