Mar 10, 2016

Script R8: Plotting Bacterial Taxonomy from MetaPhlan

F1000Research

In 1 collection

DOI
dx.doi.org/10.17504/protocols.io.ejbbcin

HANNIGAN GD[1], GRICE EA[1], ET AL.[1]

[1]DEPARTMENT OF DERMATOLOGY UNIVERSITY OF PENNSYLVANIA

VERVE Net

Club Grice

Geoffrey D Hannigan
Merck Research Labs

OPEN ACCESS

DOI: dx.doi.org/10.17504/protocols.io.ejbbcin

External link: http://mbio.asm.org/content/6/5/e01578-15.full

Protocol Citation: HANNIGAN GD, GRICE EA, ET AL. 2016. Script R8: Plotting Bacterial Taxonomy from MetaPhlan. **protocols.io https://dx.doi.org/10.17504/protocols.io.ejbbcin**

Manuscript citation:
Kindler L, Stoliartchouk A, Teytelman L, Hurwitz BL, Method-centered digital communities on protocols.io for fast-paced scientific innovation. F1000Research doi: 10.12688/f1000research.9453.2

Protocol status: Working

Created: February 16, 2016

Last Modified: March 11, 2018

Protocol Integer ID: 2371

Abstract

This protocol outlines the analysis used to plot MetaPhlAn taxonomic assignments. Based on methods from the following publication:

Hannigan, Geoffrey D., et al. "The Human Skin Double-Stranded DNA Virome: Topographical and Temporal Diversity, Genetic Enrichment, and Dynamic Associations with the Host Microbiome." *mBio* 6.5 (2015): e01578-15.

Guidelines

sessionInfo()

```
## R version 3.2.0 (2015-04-16)
## Platform: x86_64-apple-darwin13.4.0 (64-bit)
## Running under: OS X 10.10.4 (Yosemite)
## ## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats graphics grDevices utils datasets methods base
##
## loaded via a namespace (and not attached):
## [1] magrittr_1.5    formatR_1.2    tools_3.2.0    htmltools_0.2.6
## [5] yaml_2.1.13    stringi_0.4-1    rmarkdown_0.7    knitr_1.10.5
## [9] stringr_1.0.0    digest_0.6.8    evaluate_0.7
```

Before start

Supplemental information available at:

https://figshare.com/articles/The_Human_Skin_dsDNA_Virome_Topographical_and_Temporal_Diversity_Genetic_Enrichment_and_Dynamic_Associations_with_the_Host_Microbiome/1281248

1 Load the libraries needed for analysis.

Command

```
library(ggplot2)
packageVersion(
```

Expected result

```
## [1] '1.0.1'
```

Expected result

```
## [1] '1.4.1'
```

Expected result

```
## [1] '1.8.2'
```

Expected result

```
## [1] '1.1.2'
```

2 Read in the metadata and format it so that it matches the samples we are working with.

Command

```
skinmet_metadata<-read.delim(
```

3 We only want to look at a certain number of taxa to make the data more visually informative.

Command

```
topTaxa<- function(data, numTaxa){ if(nrow(data)>numTaxa){
data$RowSum<-rowSums(data) data<-data[order(-data$RowSum),] tmp<-
data[numTaxa:nrow(data),] data<-data[-c(numTaxa:nrow(data)),] other<-
colSums(tmp) data<-rbind(data,other) row.names(data)[nrow(data)]<-
```

4 Now we are ready to read in the MetaPhlAn merged output at the genus level and format it for plotting.

Command

```
skinmet_data<-read.delim(
```

5 Format sample IDs.

Command

```
skinmet_data$ID<-gsub(x=skinmet_data$ID,pattern=
```

Expected result

	##	MG100128	MG100129	MG100130	MG100131
	## Abiotrophia	0	0	0	0
	## Acetobacteraceae_unclassified	0	0	0	0
	Achromobacter	0	0	0	0
	Acidaminococcaceae_unclassified	0	0	0	0
	Acidovorax	0	0	0	0
	Acinetobacter	0	0	0	0

6 Look at top 10 taxa.

Command

```
skinmet_data<-topTaxa(skinmet_data,10)
taxa_order<-as.vector(row.names(skinmet_data))
skinmet_datat<-as.data.frame(t(skinmet_data))
skinmet_data2<-merge(skinmet_datat,skinmet_metadata,by.x=
```

7 Plot by site symbol and site categories.

Command

```
ggplot(skinmet_datam, aes(x=factor(SampleID), y=value, fill=variable,
order=variable))+theme_bw()+geom_bar(stat =
```

Expected result