

Outline for today - 23 May 2012

1. Some technical details of amplicon sequencing
2. Overview of analysis process
3. Overview of phyloseq
4. Live demo “hackalong” of phyloseq in R

Stats 366

Joey McMurdie guest lecture

Bioconductor workshop in July:

<https://secure.bioconductor.org/BioC2012/>



COURSE REGISTRATION

BioC 2012

July 24-25, 2012 (Developer Day: July 23)
Fred Hutchinson Cancer Research Center - Seattle, WA

Overview

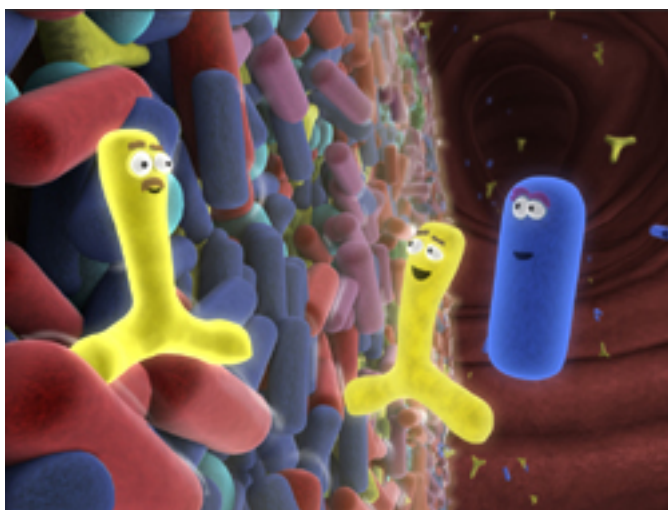
This conference highlights current developments within and beyond [Bioconductor](#), an international open source and open development software project for the analysis and comprehension of high-throughput genomic data.

[Submit](#) a BioC2012 **travel scholarship** application (due June 15).

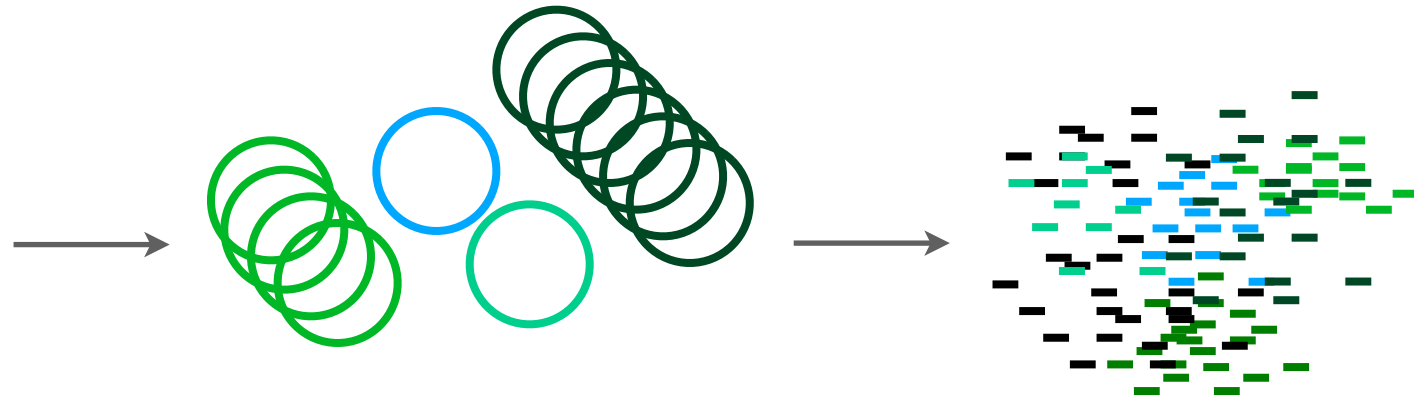
[Register](#)

Amplicon Sequencing

Goal: Infer original abundance of different types of target gene



biological sample
e.g. bacterial community

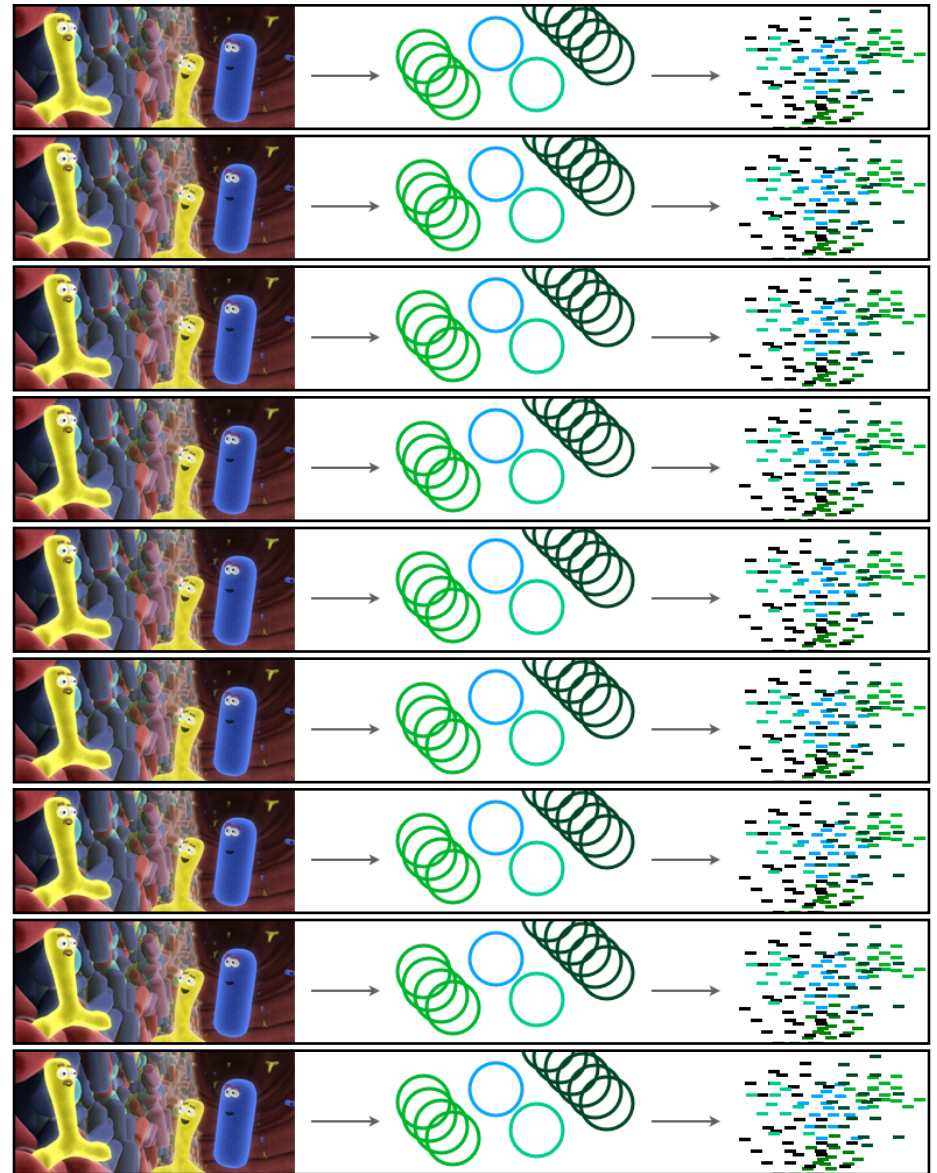
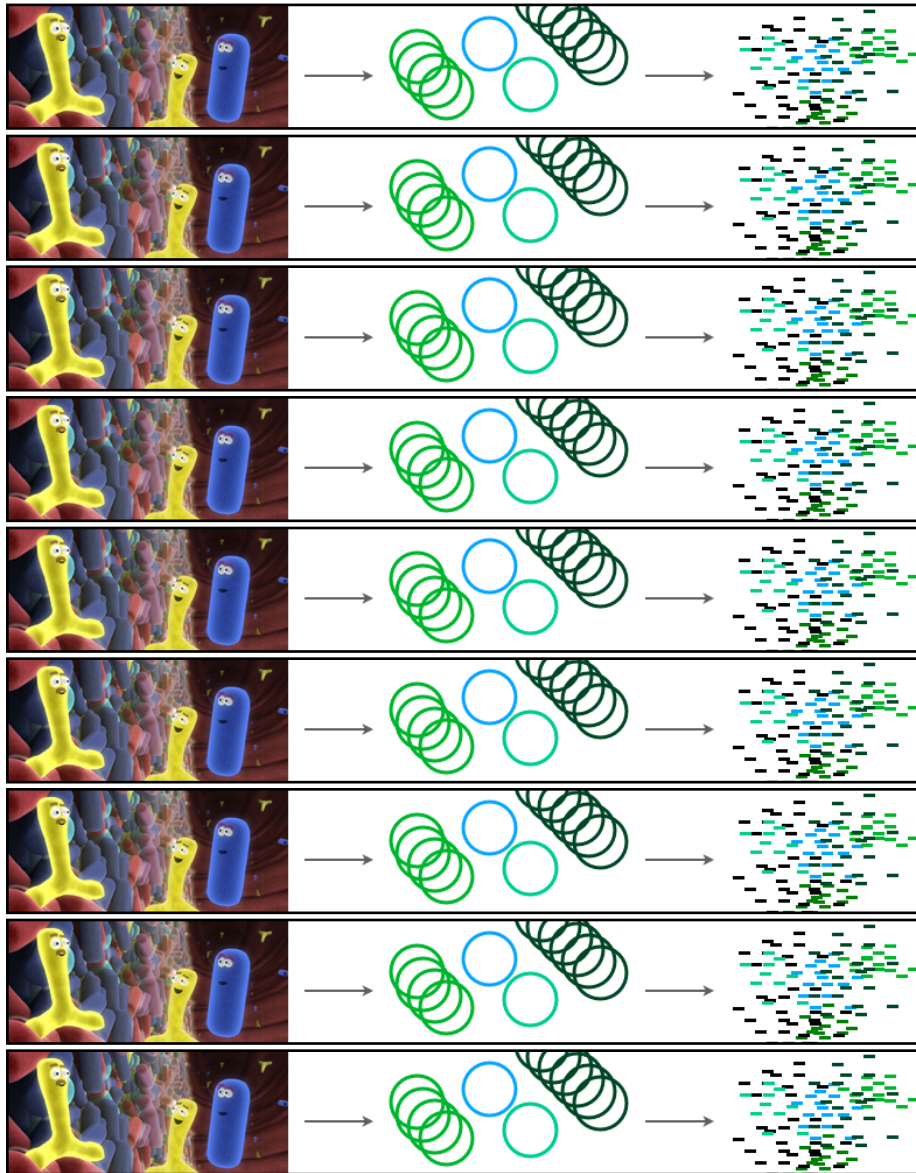


Extract DNA
(mixture)

Amplify single gene
of interest.
Sequence products

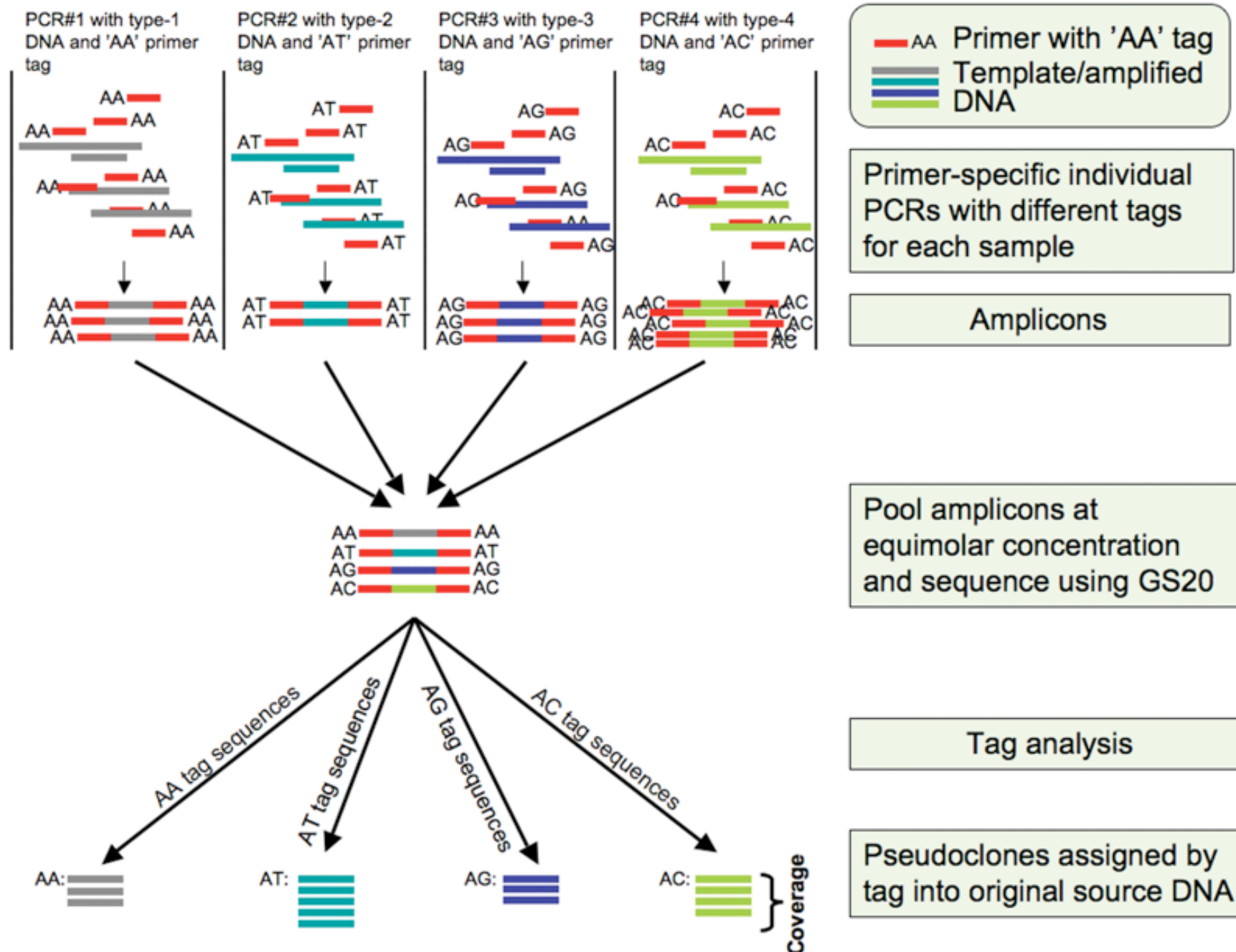
Amplicon Sequencing

Repeat many times with different samples/replicates



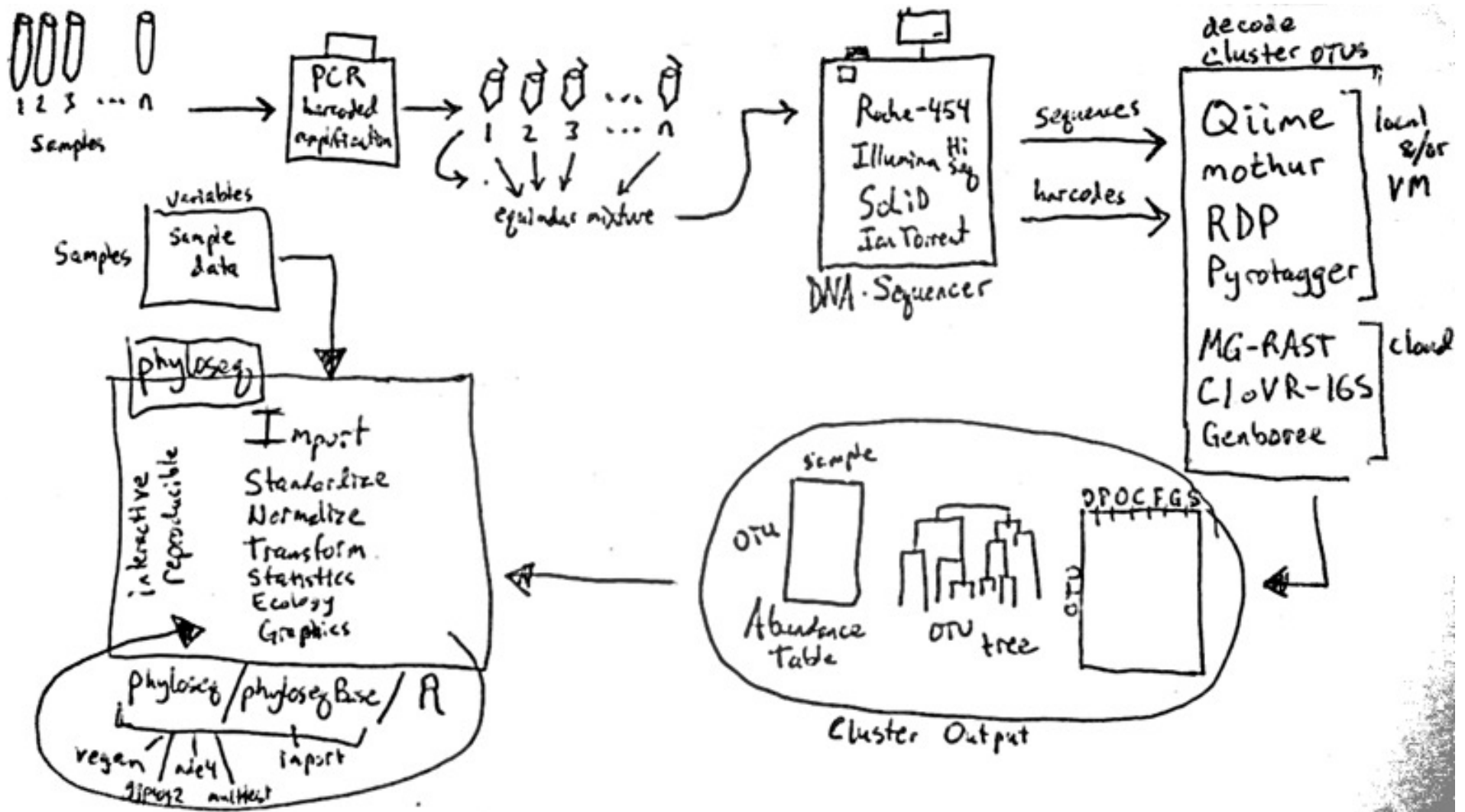
Amplicon Sequencing

parallel tagged sequencing
“bar-coded” sequences

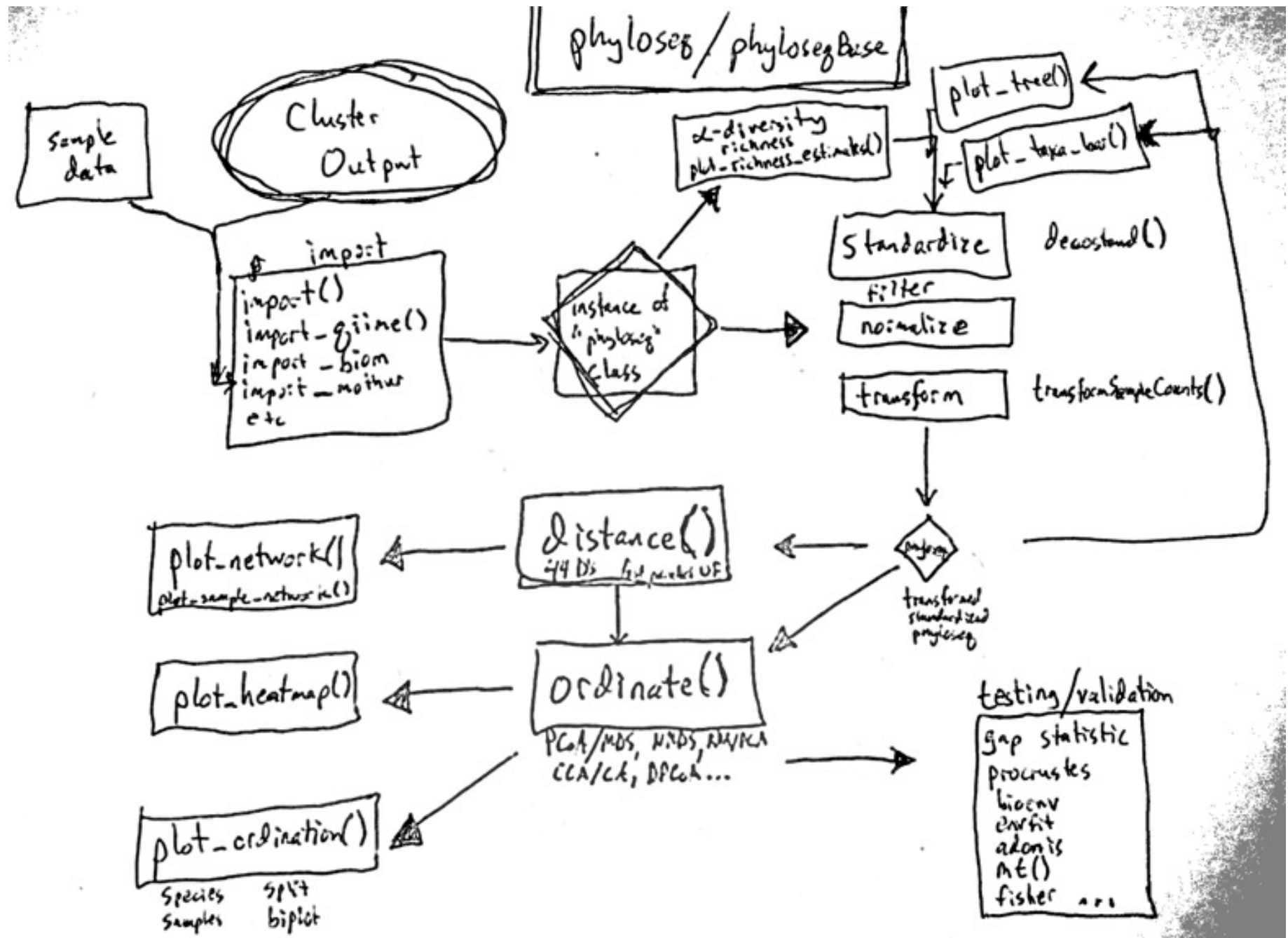


Amplicon Sequencing

Overview of amplicon sequencing and analysis

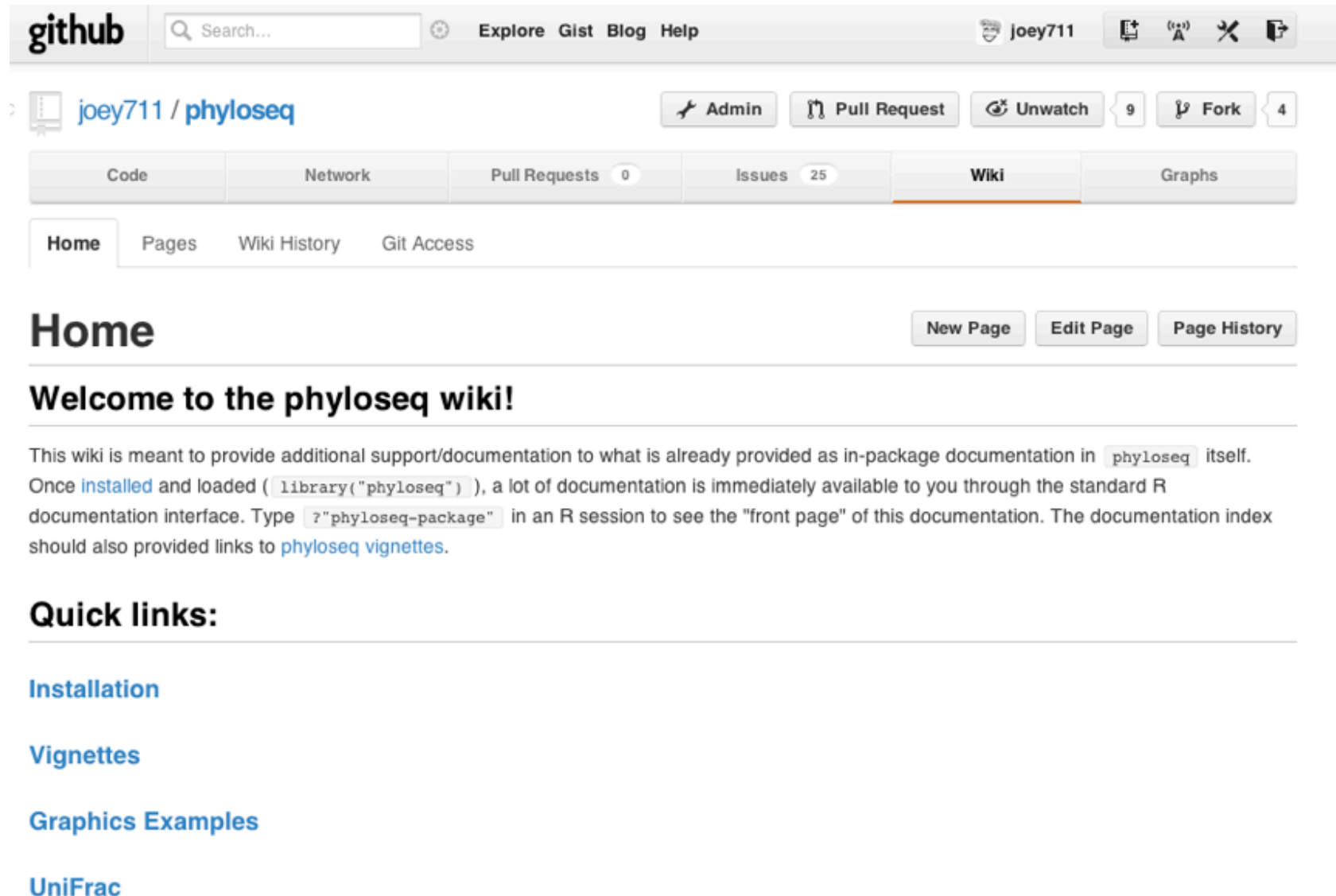


Overview of phyloseq



phyloseq wiki

<https://github.com/joey711/phyloseq/wiki>



The screenshot shows the GitHub interface for the `joey711 / phyloseq` repository. The top navigation bar includes the GitHub logo, a search bar, and links to Explore, Gist, Blog, and Help. The repository name `joey711 / phyloseq` is displayed, along with buttons for Admin, Pull Request, Unwatch, Fork, and a counter for 4 forks. Below this, a tabbed interface shows Code, Network, Pull Requests (0), Issues (25), Wiki (selected), and Graphs. The Wiki tab has sub-tabs for Home, Pages, Wiki History, and Git Access. The **Home** sub-tab is active, showing a heading **Welcome to the phyloseq wiki!** and three buttons: New Page, Edit Page, and Page History. The main content area contains a paragraph explaining the purpose of the wiki and providing instructions on how to use the documentation in an R session. Below this, there is a section titled **Quick links:** with a list of links: Installation, Vignettes, Graphics Examples, and UniFrac.

github Search... Explore Gist Blog Help Joey711

joey711 / phyloseq Admin Pull Request Unwatch 9 Fork 4

Code Network Pull Requests 0 Issues 25 Wiki Graphs

Home Pages Wiki History Git Access

Home

New Page Edit Page Page History

Welcome to the phyloseq wiki!

This wiki is meant to provide additional support/documentation to what is already provided as in-package documentation in `phyloseq` itself. Once [installed](#) and loaded (`library("phyloseq")`), a lot of documentation is immediately available to you through the standard R documentation interface. Type `? "phyloseq-package"` in an R session to see the "front page" of this documentation. The documentation index should also provided links to [phyloseq vignettes](#).

Quick links:

- [Installation](#)
- [Vignettes](#)
- [Graphics Examples](#)
- [UniFrac](#)

phyloseq wiki

<https://github.com/joey711/phyloseq/wiki/Graphics-Examples>

Graphics Examples

[New Page](#)[Edit Page](#)[Page History](#)

This page links you to a "show and tell" of wiki-pages describing different graphics options supported in the phyloseq-package, documented with example code that for making the plot. If you have data of your own to try, it is strongly recommended that you attempt some of the examples on your own dataset as well, and see if you like the results. Many of the default settings are modifiable within the function arguments directly, and virtually everything about these plots can be further modified via the layers interface of [ggplot2](#).

Some key graphics-producing functions:

[plot_heatmap](#)

[plot_tree](#)

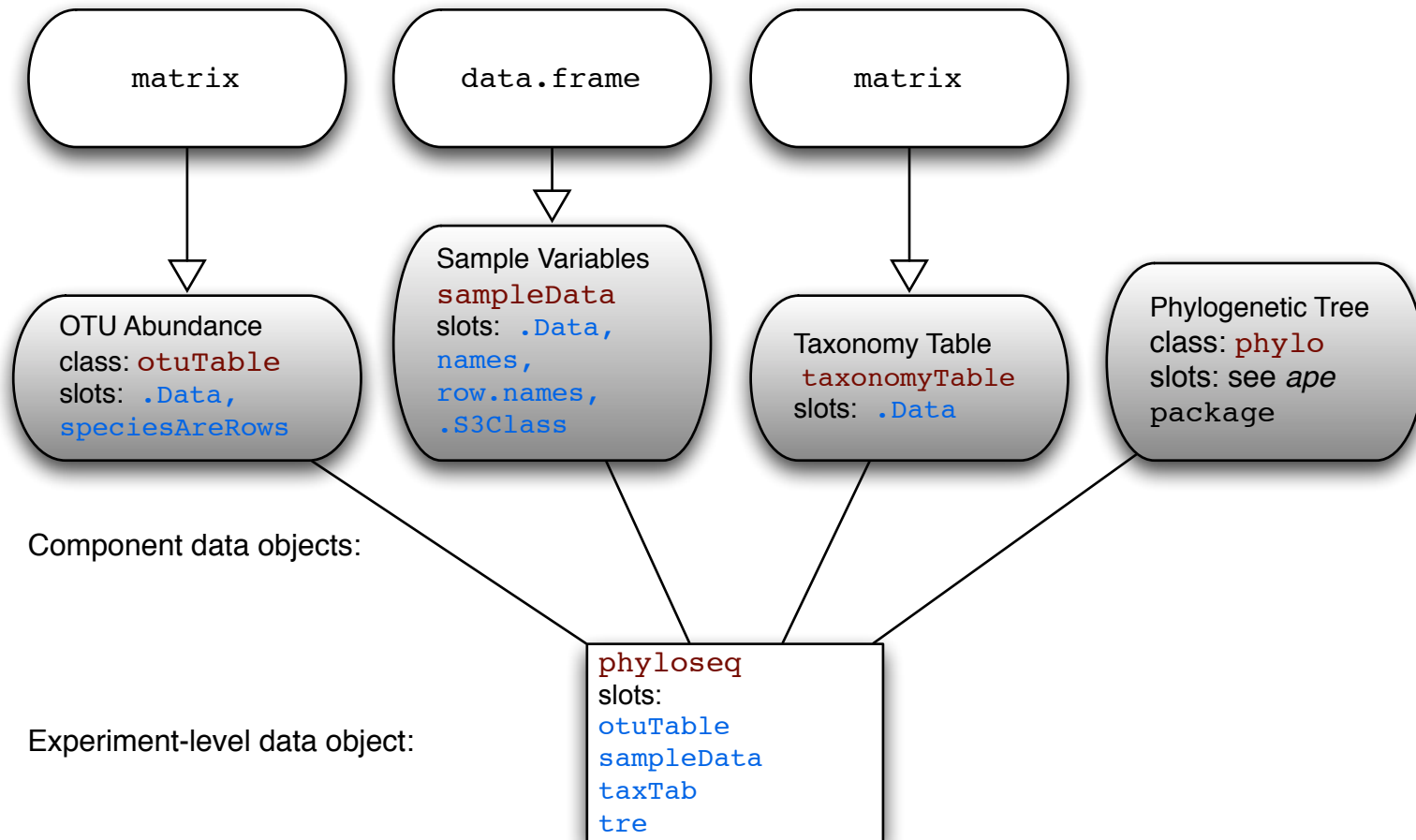
[plot_ordination](#)

[plot_sample_network](#)

[plot_richness_estimates](#)

[plot_taxa_bar](#)

phyloseq classes



phyloseq accessors

Function	Description
[Standard extraction operator. works on otuTable, sampleData, and taxonomyTable
access	General slot accessor function for phyloseq-package
getslots.phyloseq	Return the slot names of phyloseq objects
getSpecies	Returns the abundance values of sample 'i' for all species in 'x'
getSamples	Returns the abundance values of species 'i' for all samples in 'x'
getTaxa	Get a unique vector of the observed taxa at a particular taxonomic rank
getVariable	Returns an individual sample variable vector/factor
nsamples	Get the number of samples described by an object
nspecies	Get the number of species (taxa) described by an object
otuTable	Build or access otuTable objects
rank.names	Get the names of the available taxonomic ranks
sampleData	Build or access sampleData objects
sample.names	Return the names of the samples described by an object
species.names	Return the names of the species described by an object
sampleSums	Returns the total number of individuals observed from each sample
sample.variables	Returns the names of sample variables in an object
speciesSums	Returns the total number of individuals observed from each species
speciesAreRows	Returns the orientation of the abundance table
taxTab	Build or access taxTab objects
tre	Access the tree contained in a phyloseq object

phyloseq constructors

Functions for building component data objects

Function	Input Class	Output Description
<code>otuTable</code>	numeric matrix	<code>otuTable</code> object storing taxa abundance
<code>otuTable</code>	data.frame	<code>otuTable</code> object storing taxa abundance
<code>sampleData</code>	data.frame	<code>sampleData</code> object storing sample variables
<code>taxTab</code>	character string	<code>taxonomyTable</code> object storing taxonomic identities
<code>tre</code>	file path char	phylo4-class tree, read from file
<code>tre</code>	phylo-class tree	phylo4-class tree, converted from argument
<code>read.table</code>	table file path	A matrix or data.frame (Std Rcore function)
<code>read.tree</code>	Newick file path	phylo-class tree object (ape)
<code>read.nexus</code>	Nexus file path	phylo-class tree object (ape)
<code>readNexus</code>	Nexus file path	phylo4-class tree object (phylobase)

Functions for building complex data objects

Function	Input Class	Output Description
<code>phyloseq</code>	2 or more component objects	phyloseq-class, “experiment-level” object
<code>merge_phyloseq</code>	2 or more component or phyloseq-class objects	Combined instance of phyloseq-class

Importing Data

```
import()  
import_biom()  
import_qiime()  
import_mothur()  
import_pyrotagger_tab()  
import_RDP_cluster()  
import_RDP_otu()
```

Soon:

```
import_mgrast()  
import_clovr16s()  
import_genboree()
```

Example Data

(Begin live demo)