

The phyloseq package and analysis of high throughput amplicon sequencing data

Paul McMurdie
Statistics Department
Stanford University

BioC 2012

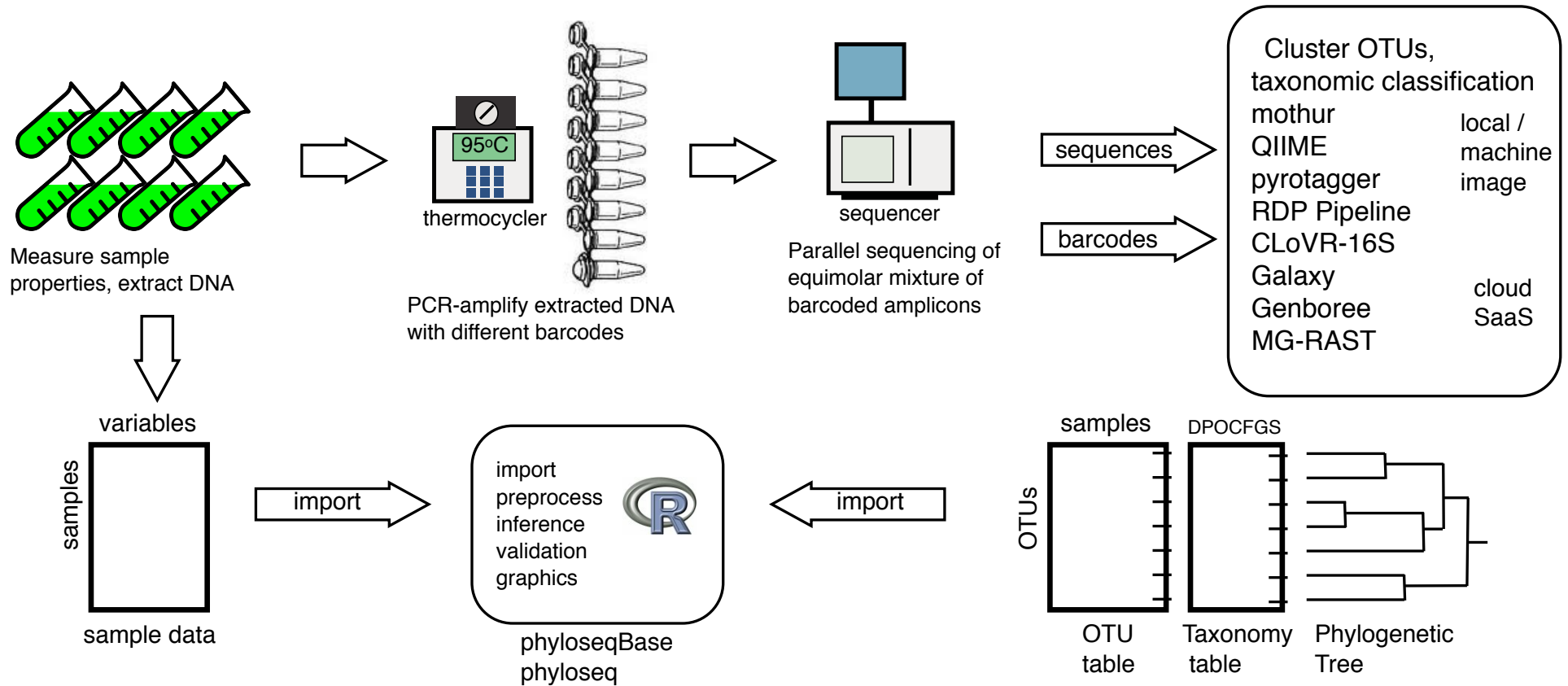
July 24-25, 2012 (Developer Day: July 23)

Fred Hutchinson Cancer Research Center - Seattle, WA



Amplicon Sequencing

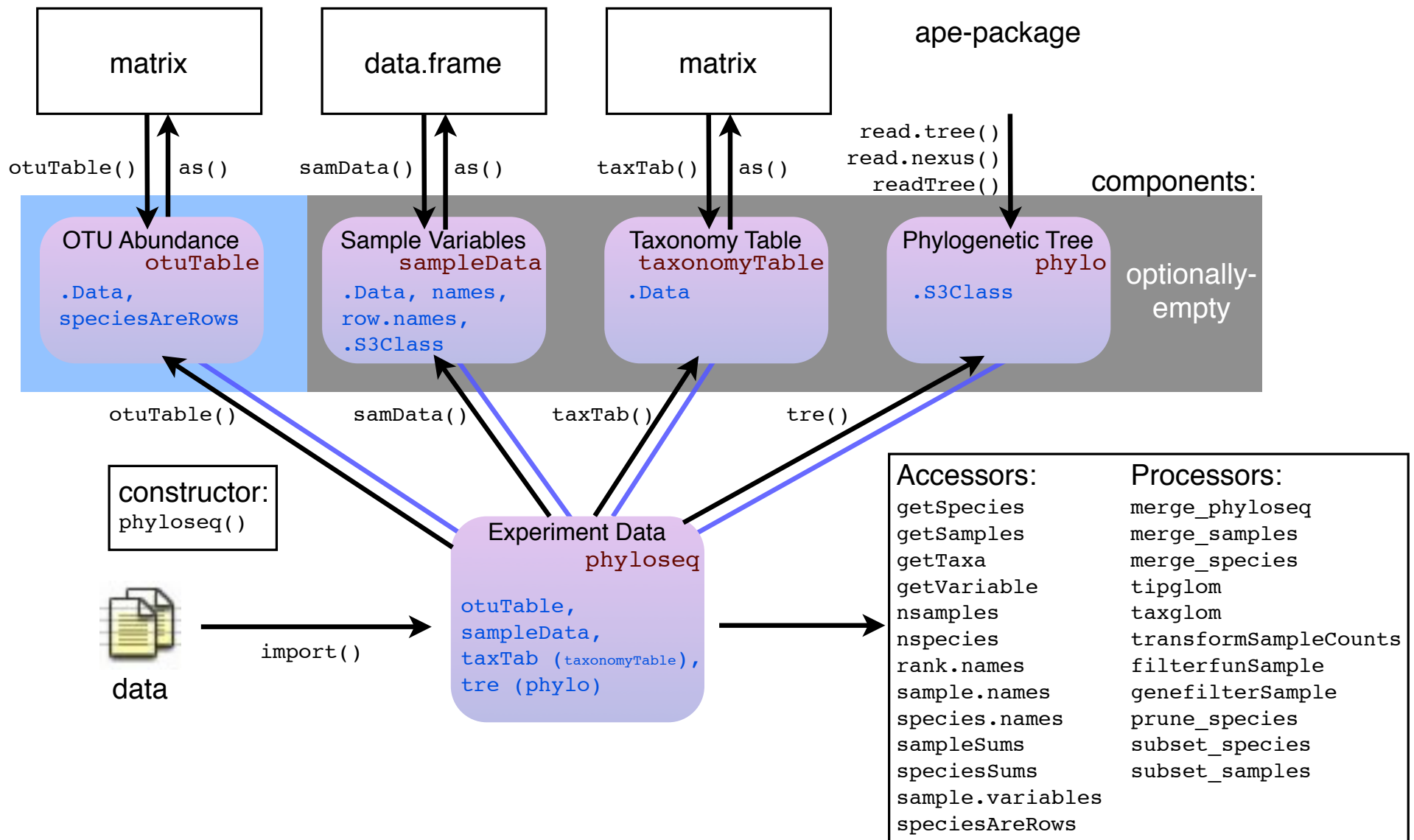
Overview of amplicon sequencing and analysis



phyloseq design and features

- A single, explicitly-defined S4 class that can store the different data types of a phylogenetic sequencing experiment in a single object.
- Importers all create this special “phyloseq” class
- Most phyloseq functions will act on this experiment-level object. It doesn't need to be diced-up to work. (Keep data together)
- Internal tools check validity and agreement among components of an experiment. Helps prevent mistakes.
- Plotting tools for creating quality graphics, built using [ggplot2](#).
- Example datasets from real published data, with references, documentation, and examples.
- Examples using other R tools after importing with phyloseq

phyloseq classes and “data infrastructure”



Overview of analysis using phyloseq

