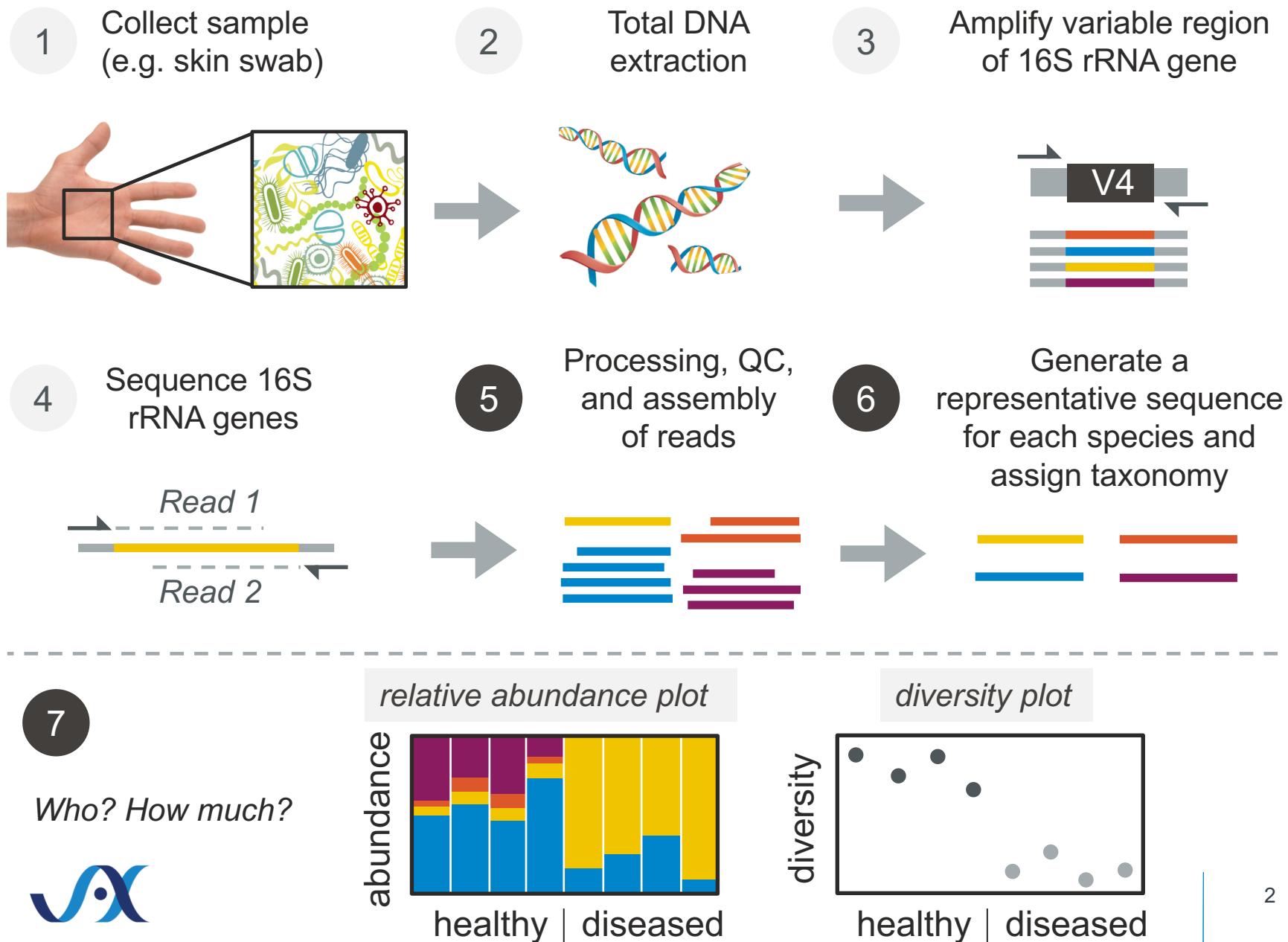


Synthesis: hands-on exercise



A workflow for 16S analysis



A microbiome module by design

Conceptual goals for students

- why do we sequence the 16S gene?
- how do we decide which 16S sequences belong to which organism?
- what can relative abundance and diversity plots tell us?
- develop hypothesis and test it

Data science goals for students

- file manipulation from the command line
- writing and executing "for loops"
- running commands in terminal
- data visualization in R
- writing R notebooks



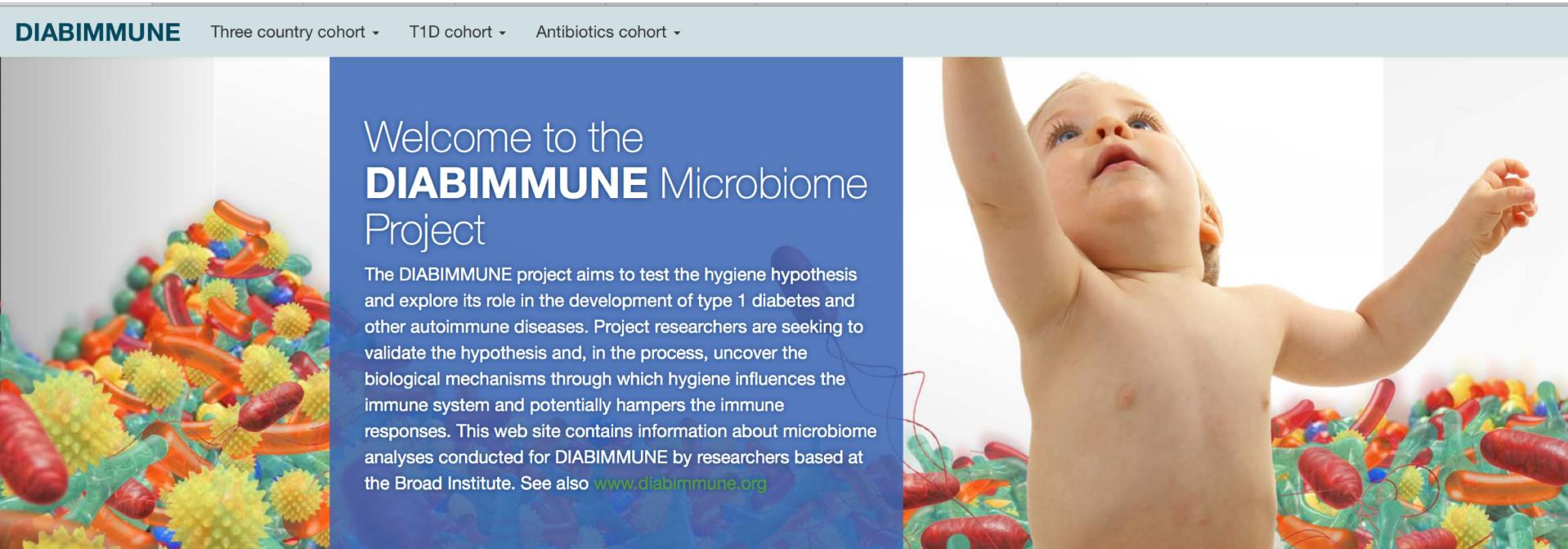
A microbiome module by design

Teaching implementation goals

- allows for individual student curiosity & open-endedness
- real-world medical relevance
- easily sourced raw data
- do not have to wait for code to run
- avoid painful coding “typos”
- avoid version/compatibility issues



The exercise: an overview



The screenshot shows the DIABIMMUNE project website. At the top, there is a navigation bar with links for "DIABIMMUNE", "Three country cohort", "T1D cohort", and "Antibiotics cohort". The main content area features a blue background with a collage of colorful, 3D-represented bacteria and viruses. On the left, a large image of a baby's torso is shown from a low angle, looking up. The central text reads: "Welcome to the **DIABIMMUNE** Microbiome Project". Below this, a detailed paragraph explains the project's aim: "The DIABIMMUNE project aims to test the hygiene hypothesis and explore its role in the development of type 1 diabetes and other autoimmune diseases. Project researchers are seeking to validate the hypothesis and, in the process, uncover the biological mechanisms through which hygiene influences the immune system and potentially hampers the immune responses. This web site contains information about microbiome analyses conducted for DIABIMMUNE by researchers based at the Broad Institute. See also www.diabimmune.org".

“What is the role of the microbiome and the “hygiene hypothesis” in the development of type 1 diabetes and other auto-immune diseases?”

<https://pubs.broadinstitute.org/diabimmune>

The exercise: an overview



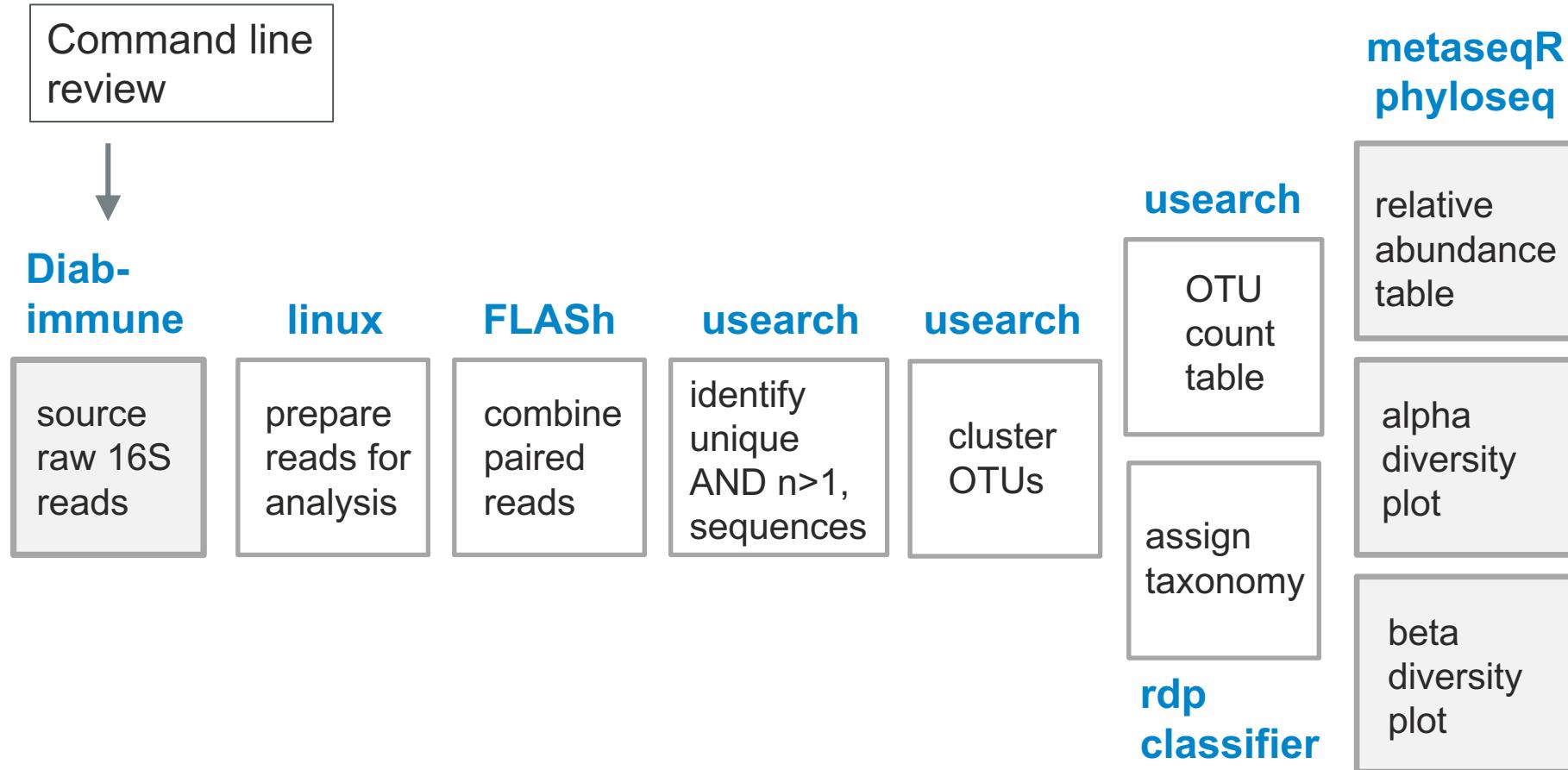
DIABIMMUNE Three country cohort ▾ T1D cohort ▾ Antibiotics cohort ▾

Welcome to the
DIABIMMUNE Microbiome
Project

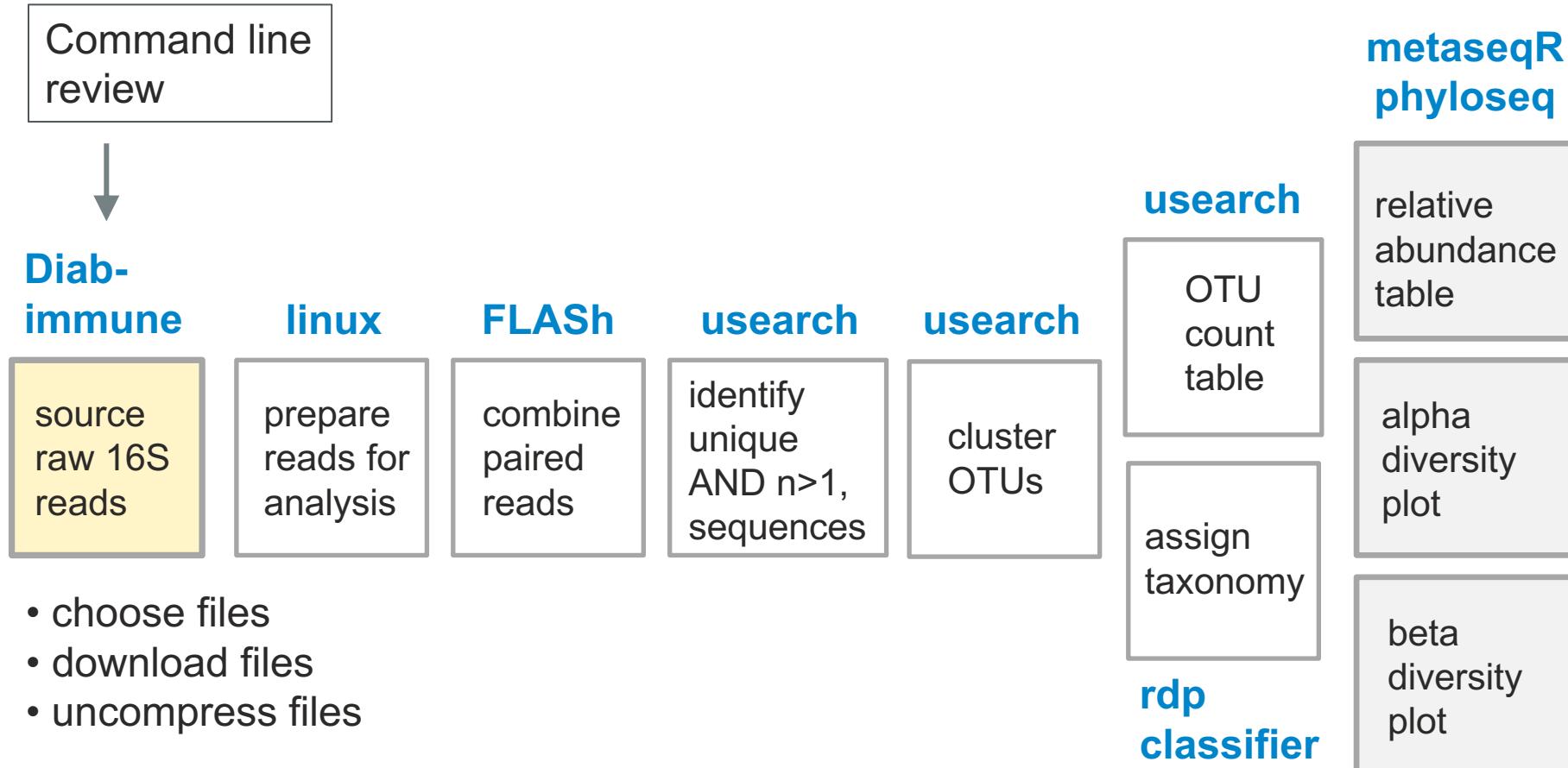
The DIABIMMUNE project aims to test the hygiene hypothesis and explore its role in the development of type 1 diabetes and other autoimmune diseases. Project researchers are seeking to validate the hypothesis and, in the process, uncover the biological mechanisms through which hygiene influences the immune system and potentially hampers the immune responses. This web site contains information about microbiome analyses conducted for DIABIMMUNE by researchers based at the Broad Institute. See also www.diabimmune.org

“The goal of this cohort is to compare microbiome in infants who have developed type 1 diabetes (T1D) or serum autoantibodies (markers predicting the onset of T1D) with healthy controls in the same area.”

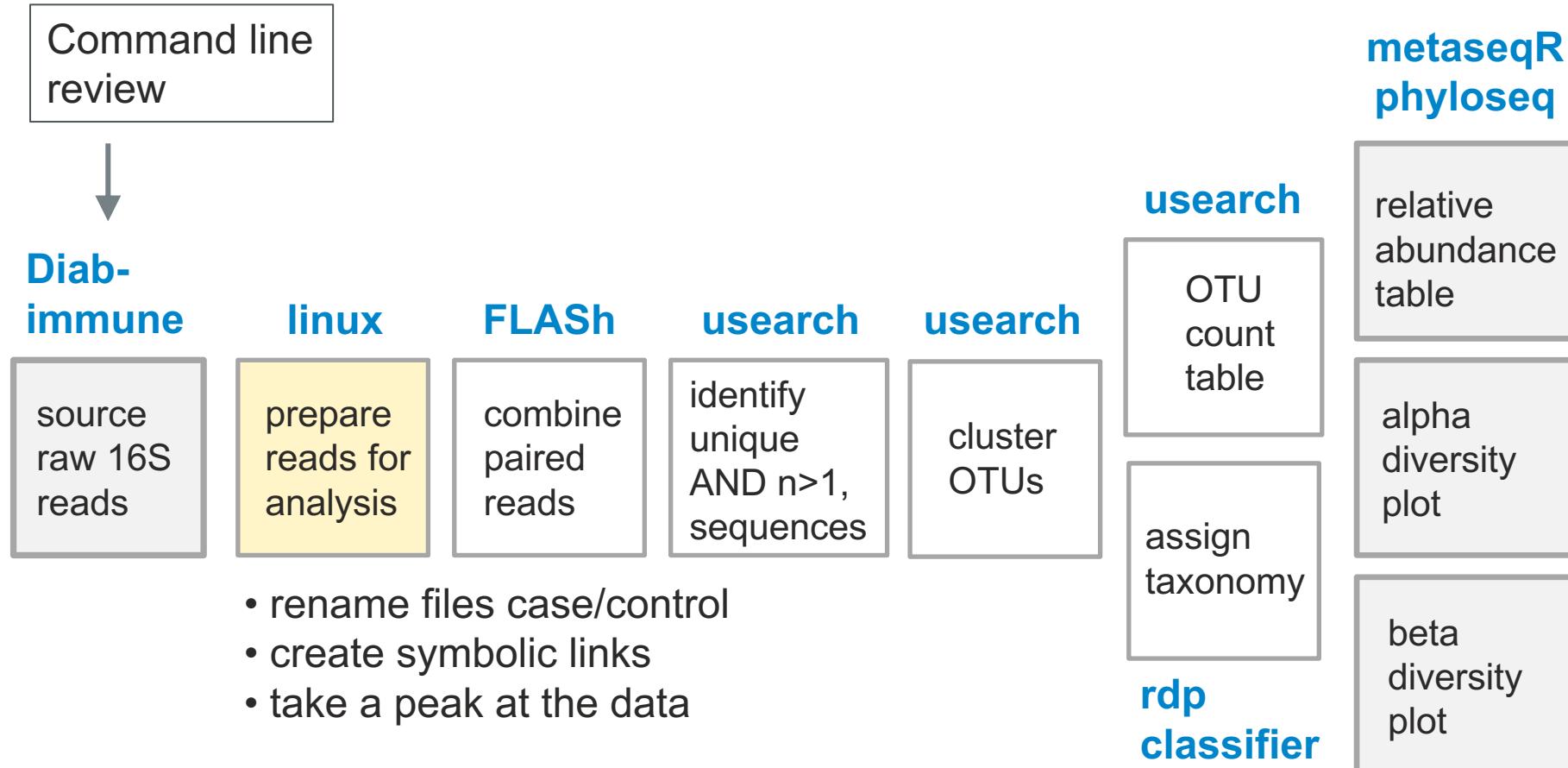
The exercise: an overview



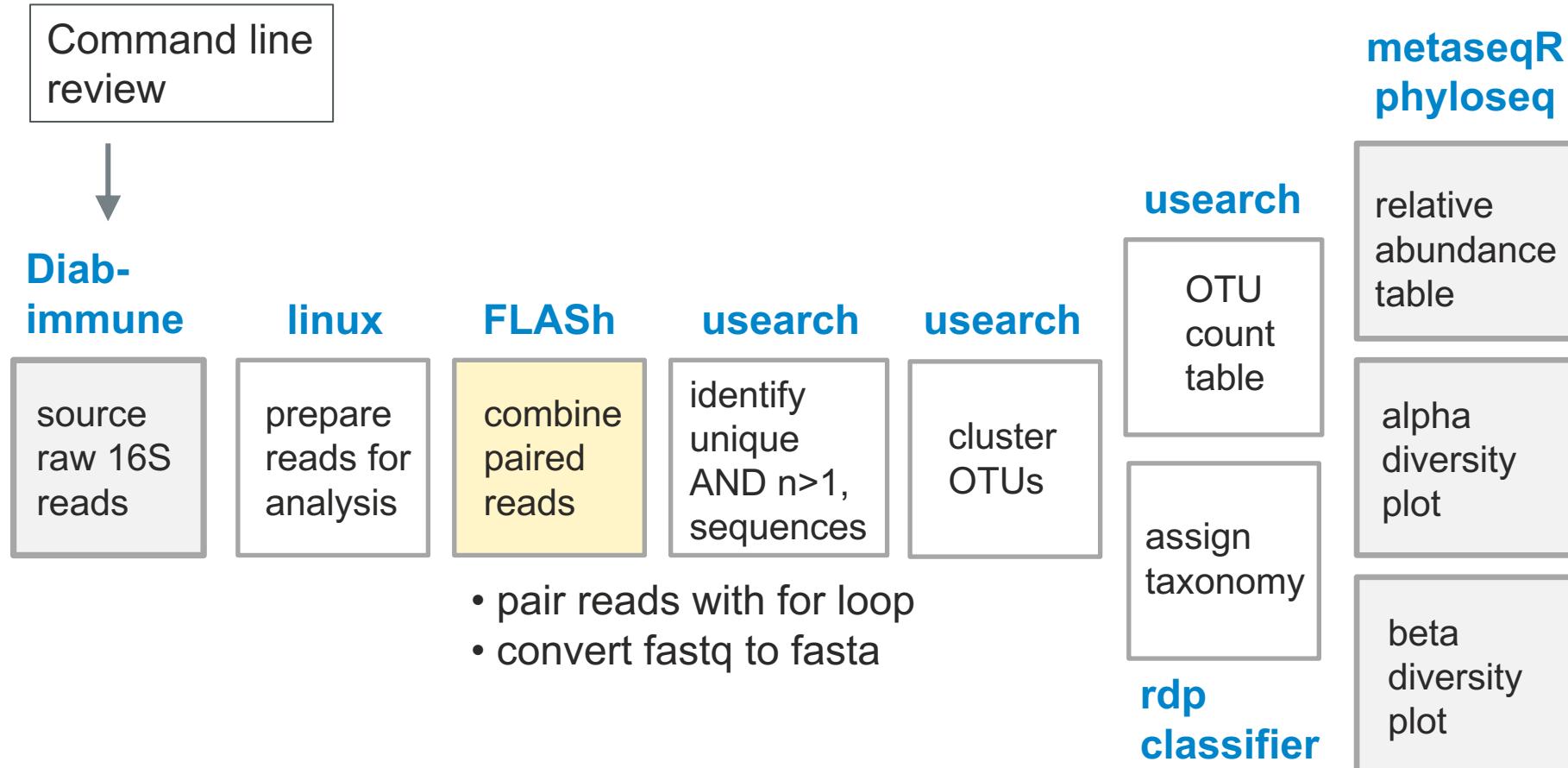
The exercise: an overview



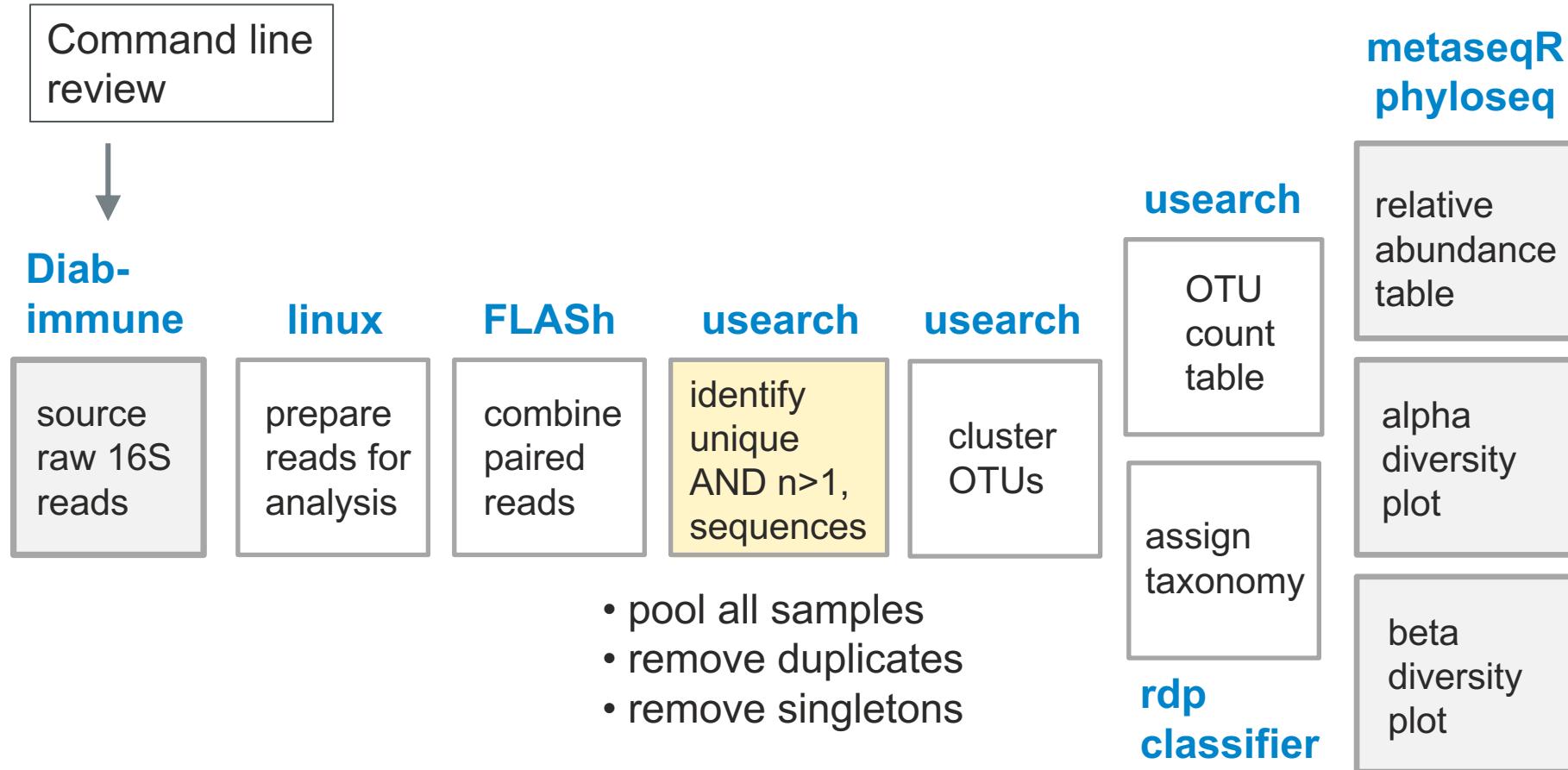
The exercise: an overview



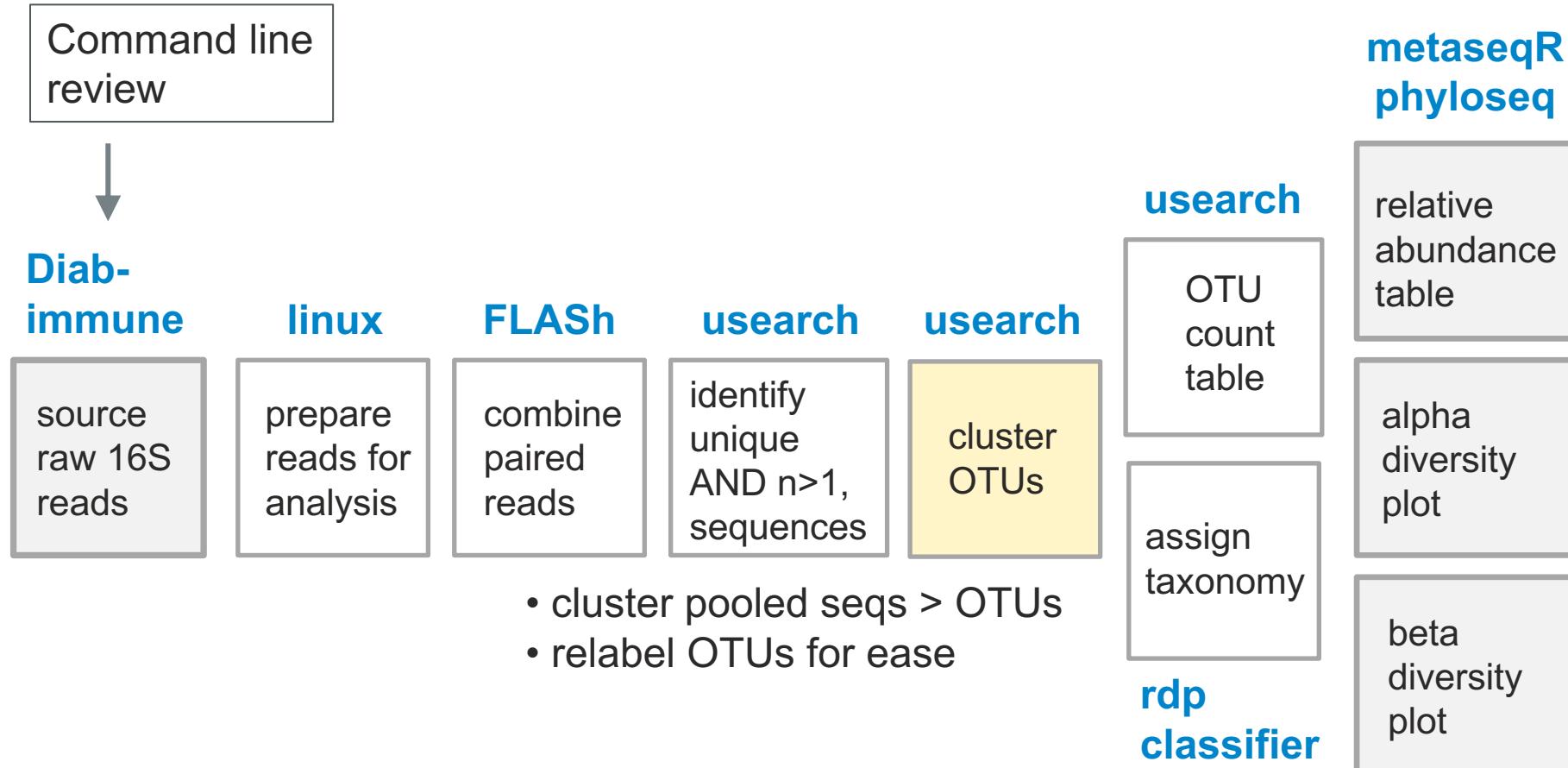
The exercise: an overview



The exercise: an overview



The exercise: an overview



The exercise: an overview

Command line
review



**Diab-
immune**

source
raw 16S
reads

linux

prepare
reads for
analysis

FLASH

combine
paired
reads

usearch

identify
unique
AND $n > 1$,
sequences

usearch

cluster
OTUs

- assign taxonomy for each OTU

usearch

OTU
count
table

metaseqR
phyloseq

relative
abundance
table

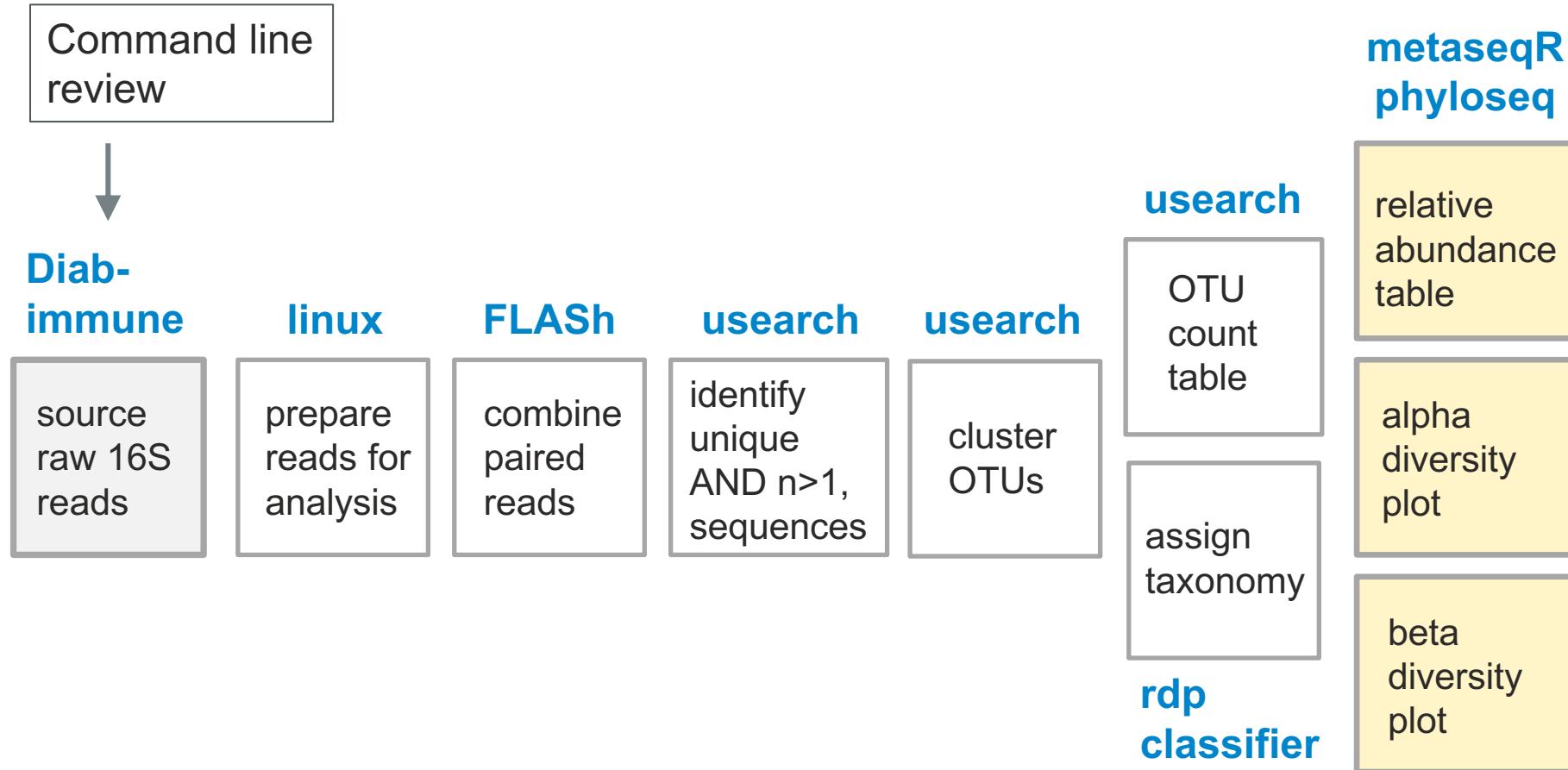
alpha
diversity
plot

assign
taxonomy

rdp
classifier

beta
diversity
plot

The exercise: an overview



Revisiting our goals

Conceptual goals for students

- why do we sequence the 16S gene?
- how do we decide which 16S sequences belong to which organism?
- what can relative abundance and diversity plots tell us?
- develop hypothesis and test it

Data science goals for students

- file manipulation from the command line
- writing and executing "for loops"
- running commands in terminal
- data visualization in R
- writing R notebooks



Revisiting our goals

Teaching implementation goals

- allows for individual student curiosity & open-endedness
- real-world medical relevance
- easily sourced raw data
- do not have to wait for code to run
- avoid painful coding “typos”
- avoid version/compatibility issues

