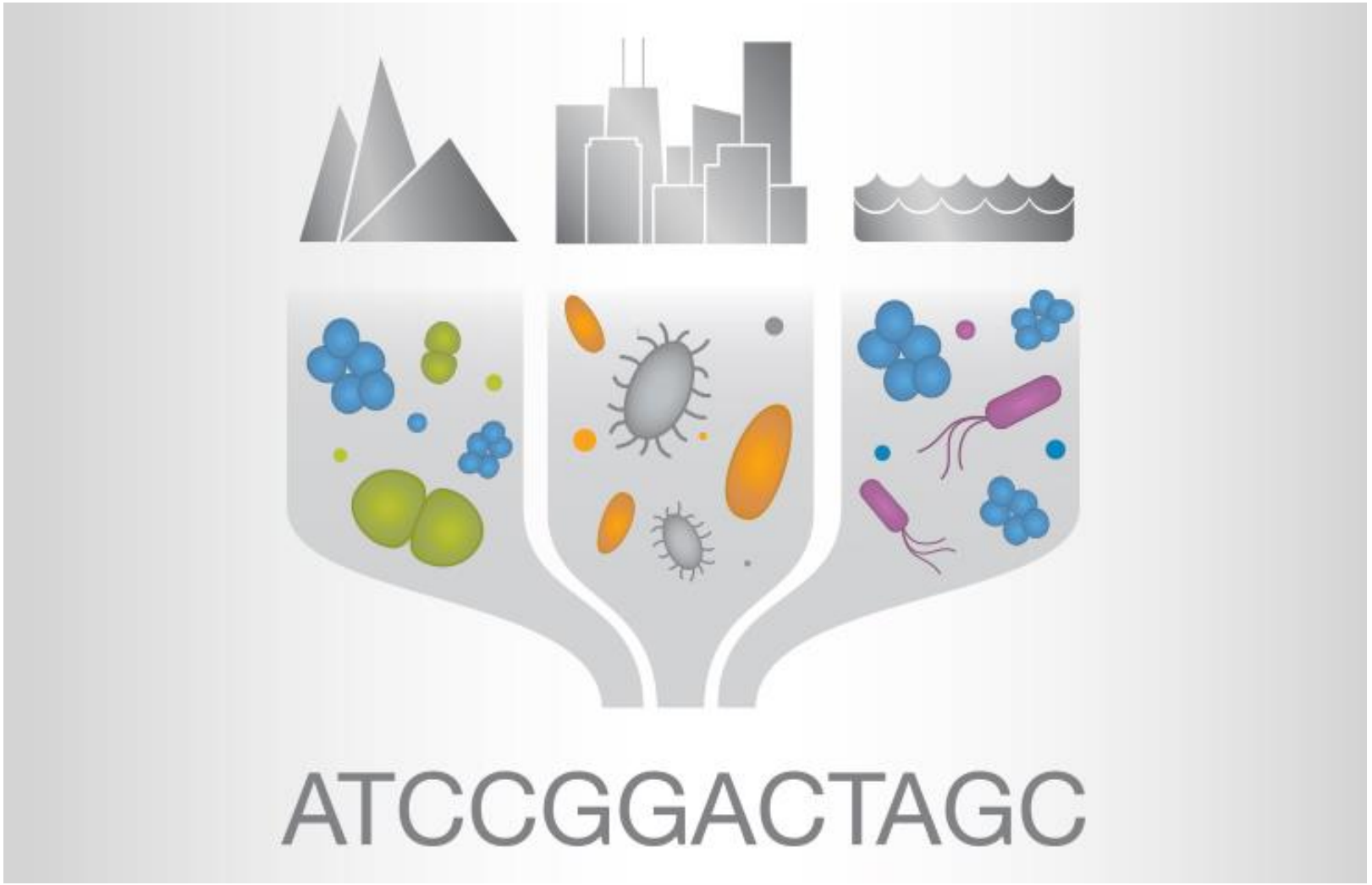
Metagenome sequencing

Andrés Cumsille



Metagenomics

Cultivation-independent genome-level characterization of communities.

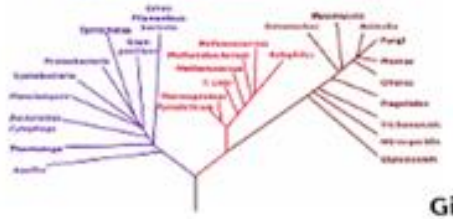
Commonly used to investigate complex microbial communities sampled directly from the environment, without isolating any organism (Ghosh *et al*., 2018).

Metagenomics timeline

(Escobar-Zepeda *et al.*, 2015)

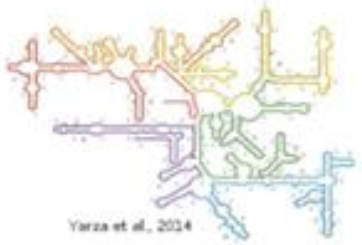
Robert Koch isolates microorganisms using solid cultures

1888



Carl Woese propose rRNA as marker for taxonomy

1977



Giovannoni et al., perform the first microbial community study by 16S rRNA libraries

1990

1676

Leeuwenhoek reports his observations about oral microbiota

1931

Winogradsky microbial ecology experiments

1980

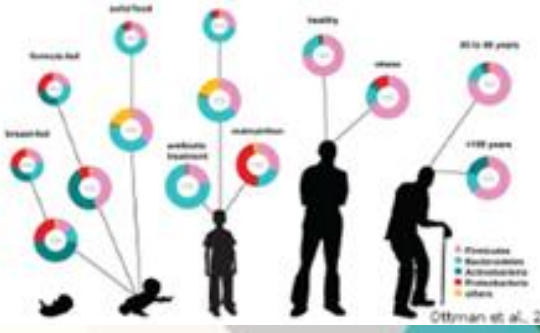
Kary Mullis develops PCR

1998

Handelsman et al., propose the term 'metagenomics'

2005

First NGS machine released by Roche



2006

GA sequencer from Solexa is released

2008

Human Microbiome Project publication

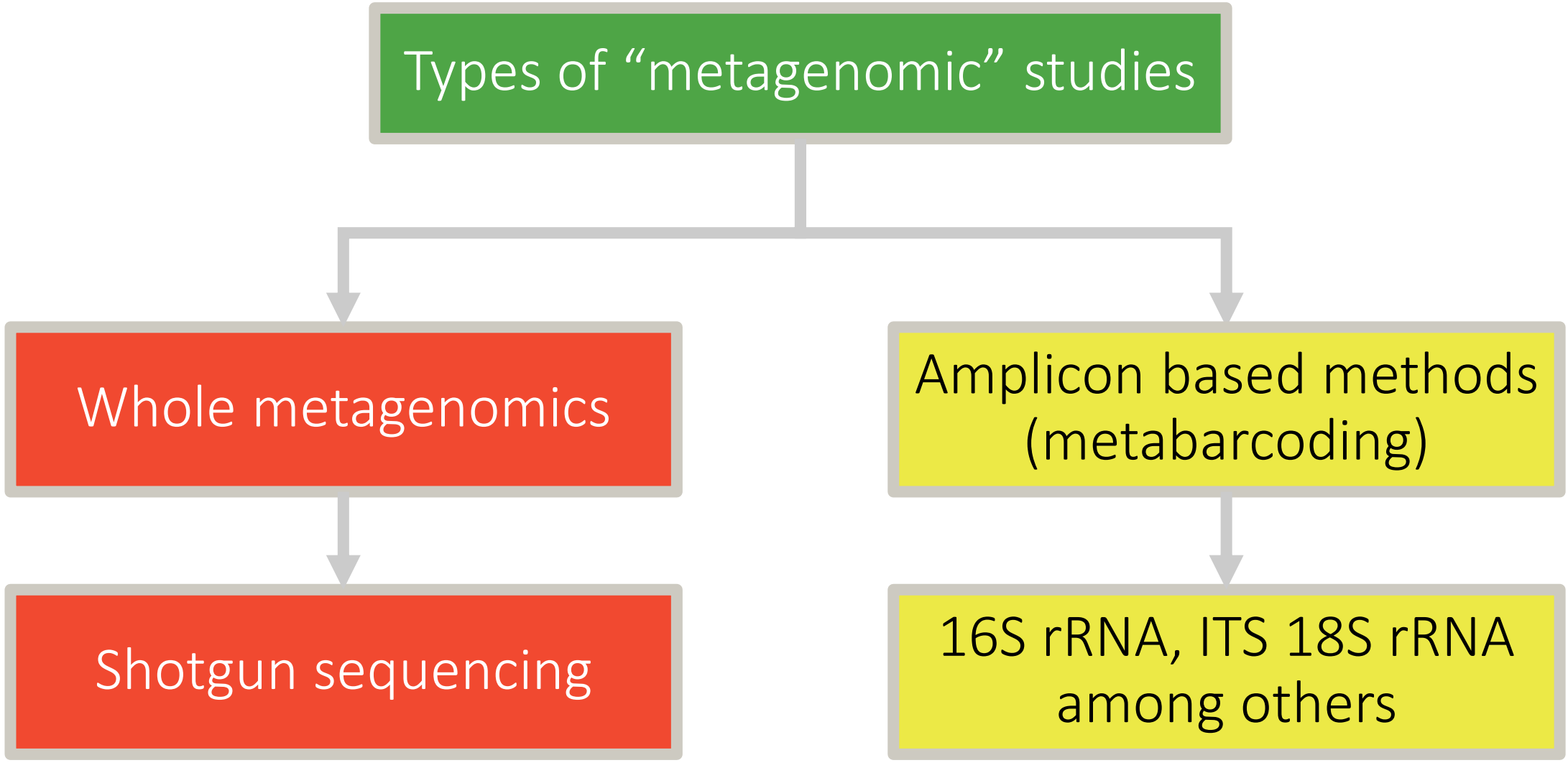
2011

PacBio RS sequencer is released

2015

Ocean Samplig Day

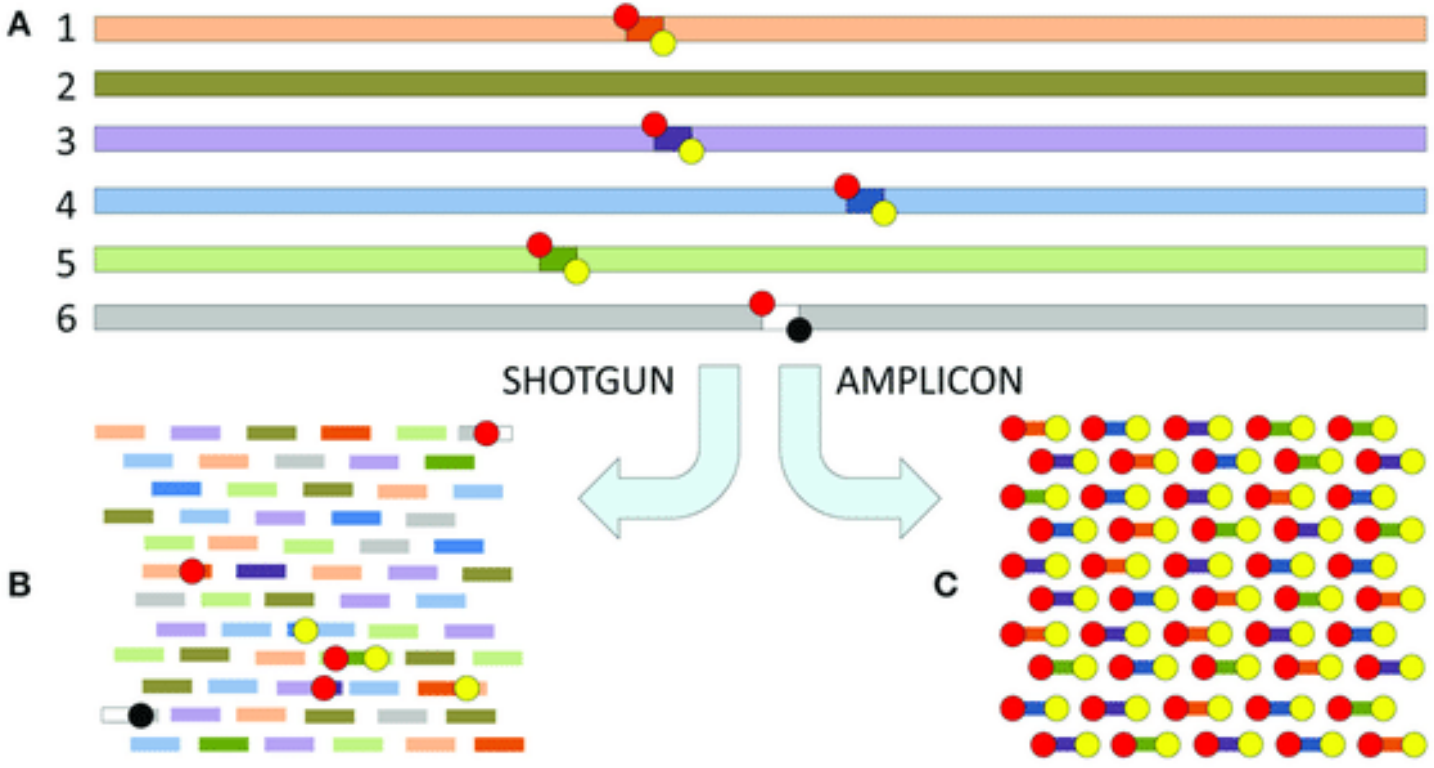
Modified from https://www.hhmi.org/biointeractive/


Types of “metagenomic” studies
Whole metagenomics
Shotgun sequencing
Amplicon based methods (metabarcoding)
16S rRNA, ITS 18S rRNA among others
What are they doing?
Who is out there?

Types of “metagenomic” studies

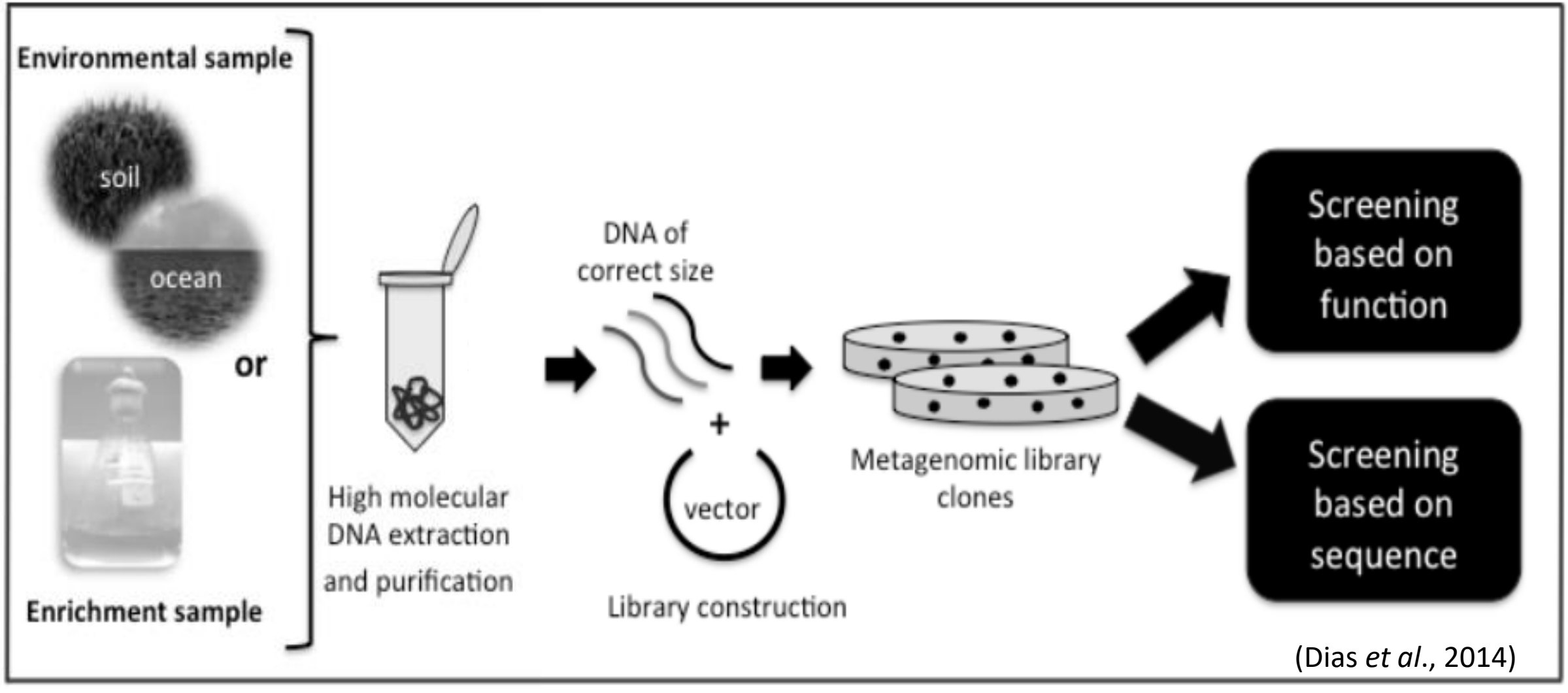
6 different genomes

Conserved motifs



(Sekse *et al.*, 2017)

Early metagenomics



(Dias *et al.*, 2014)

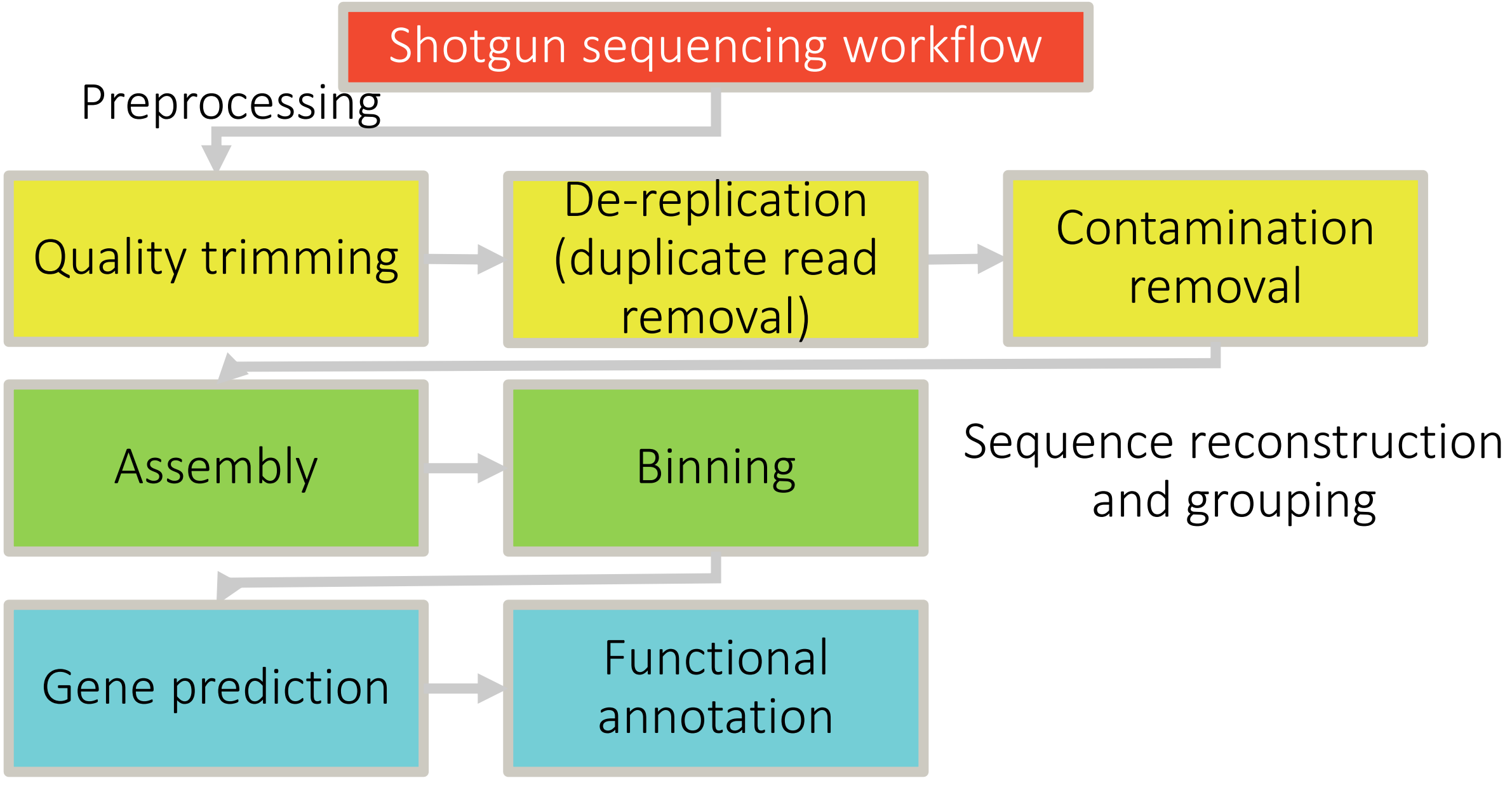
Shotgun sequencing → Reconstruct large fragments or even complete genomes from organism in a community

(Escobar-Zepeda *et al.*, 2015)

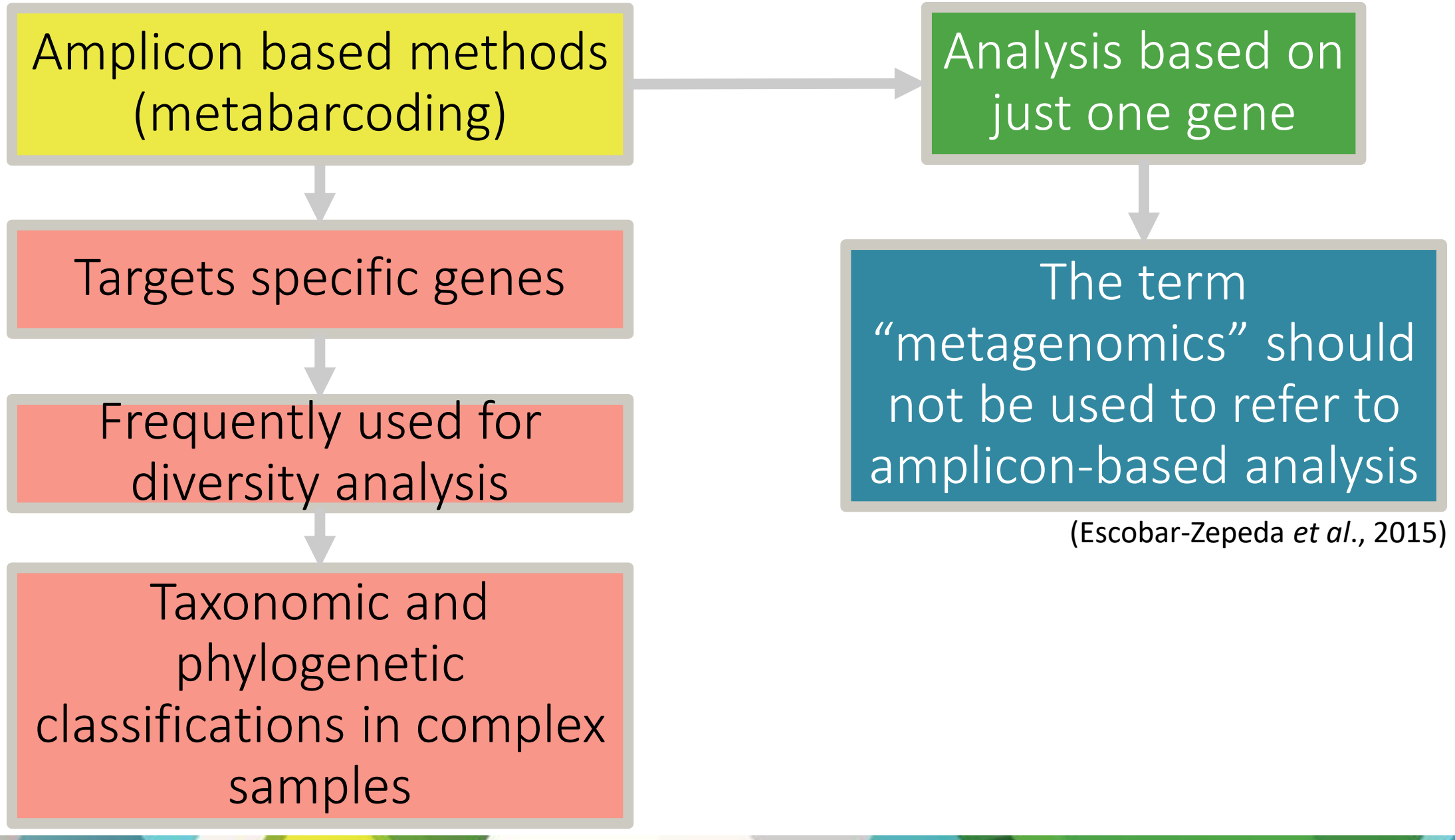
Sequence based screens → Describe the microbial diversity and genomes of a particular sample

Functional screens → Identify some functional gene products without necessarily determine the species from which is originated

(Ghosh *et al.*, 2018)

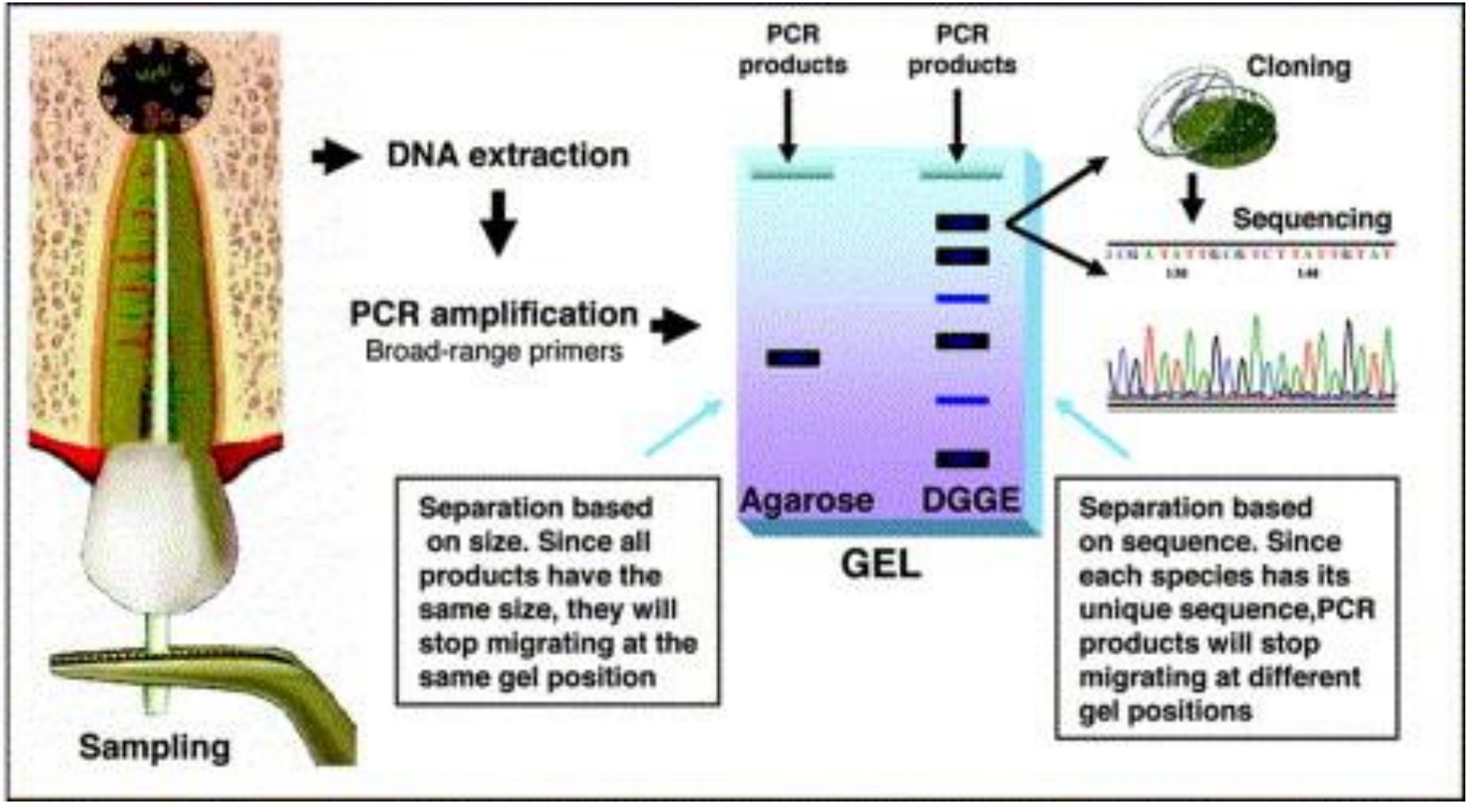


(Ghosh *et al.*, 2018)



(Escobar-Zepeda *et al*., 2015)

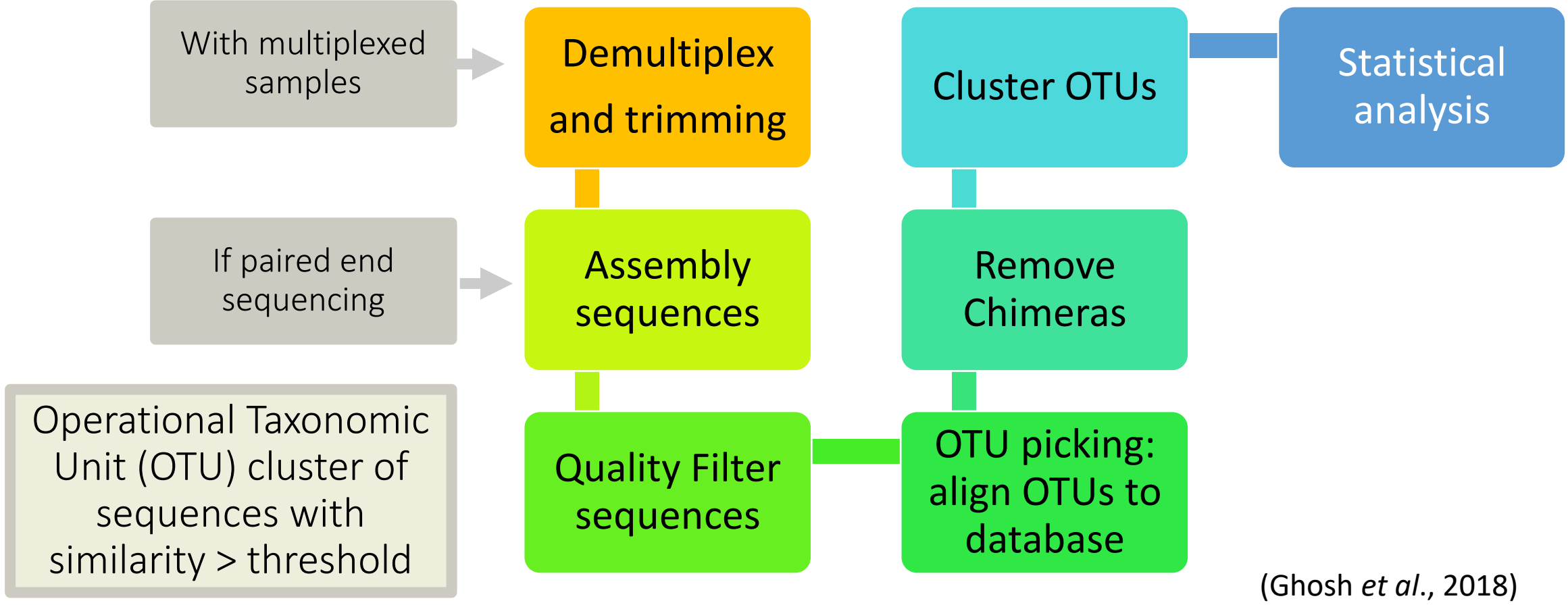
Denaturing Gradient Gel Electrophoresis (DGGE)



(Siqueira *et al.*, 2005)

Metabarcoding workflow

Qiime
Mothur
DADA2

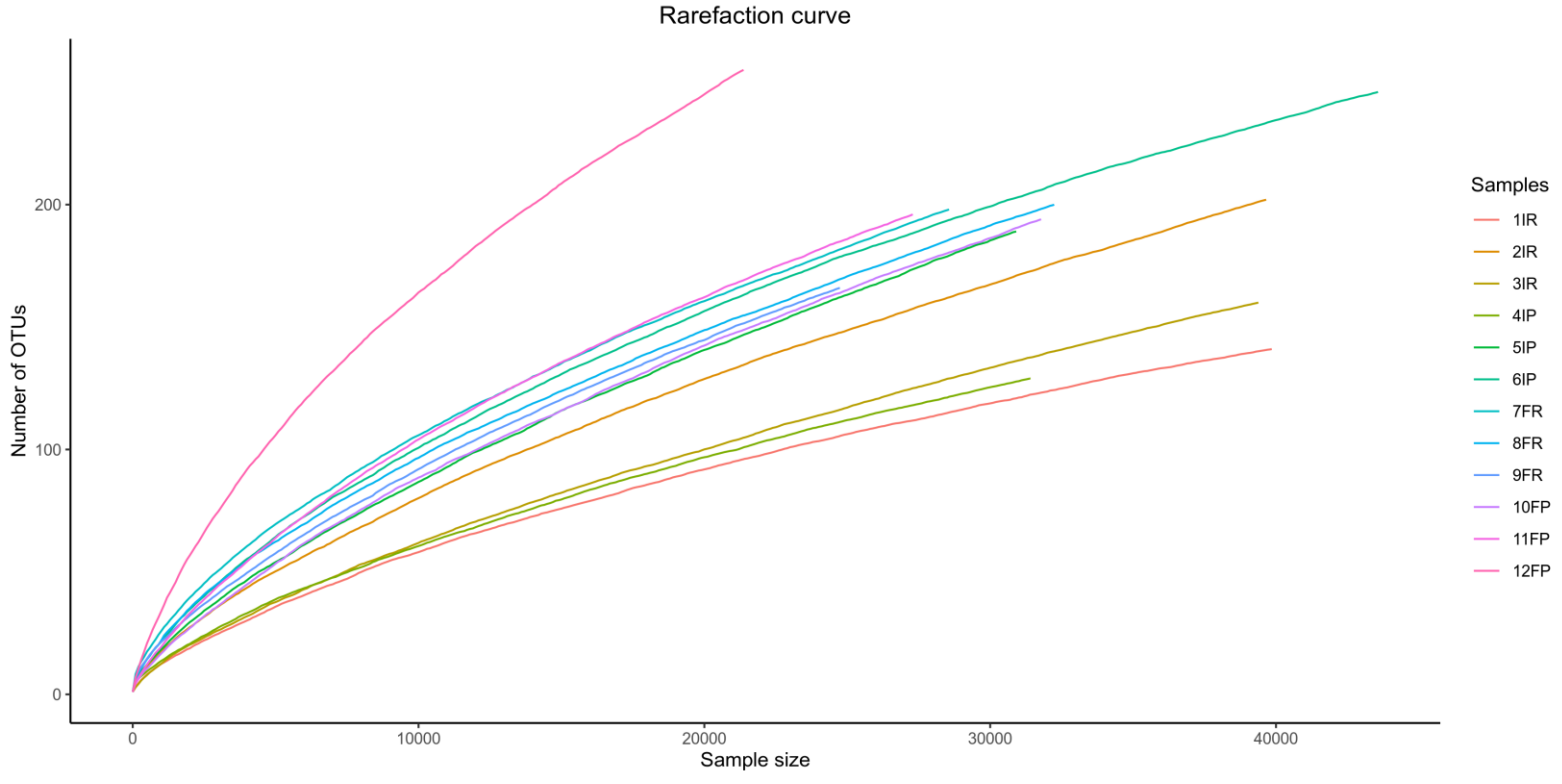


(Ghosh *et al.*, 2018)

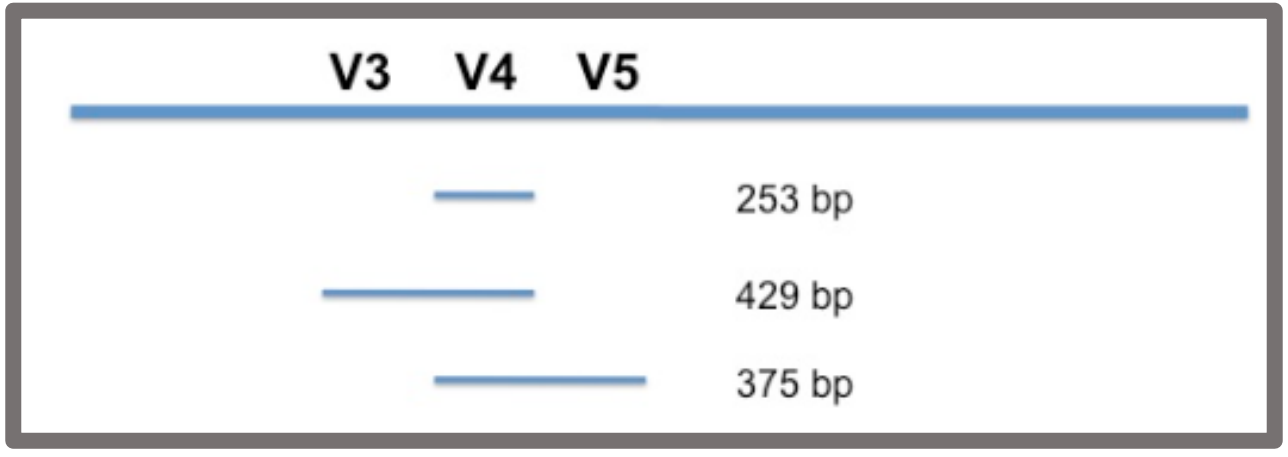
Metabarcoding considerations

- Number of reads
- Sequence length
- Cost

Rarefaction curve



Error rates



Amount of overlap for 2x250 bp reads:

V4: 247 bp
V34: 71 bp
V45: 125 bp

Study sequencing a mock community

REGION	LENGTH	ERROR RATE
V3-V4	429	0.41
V4	253	0.04
V4-V5	375	0.57

Schloss oral presentation, 2017