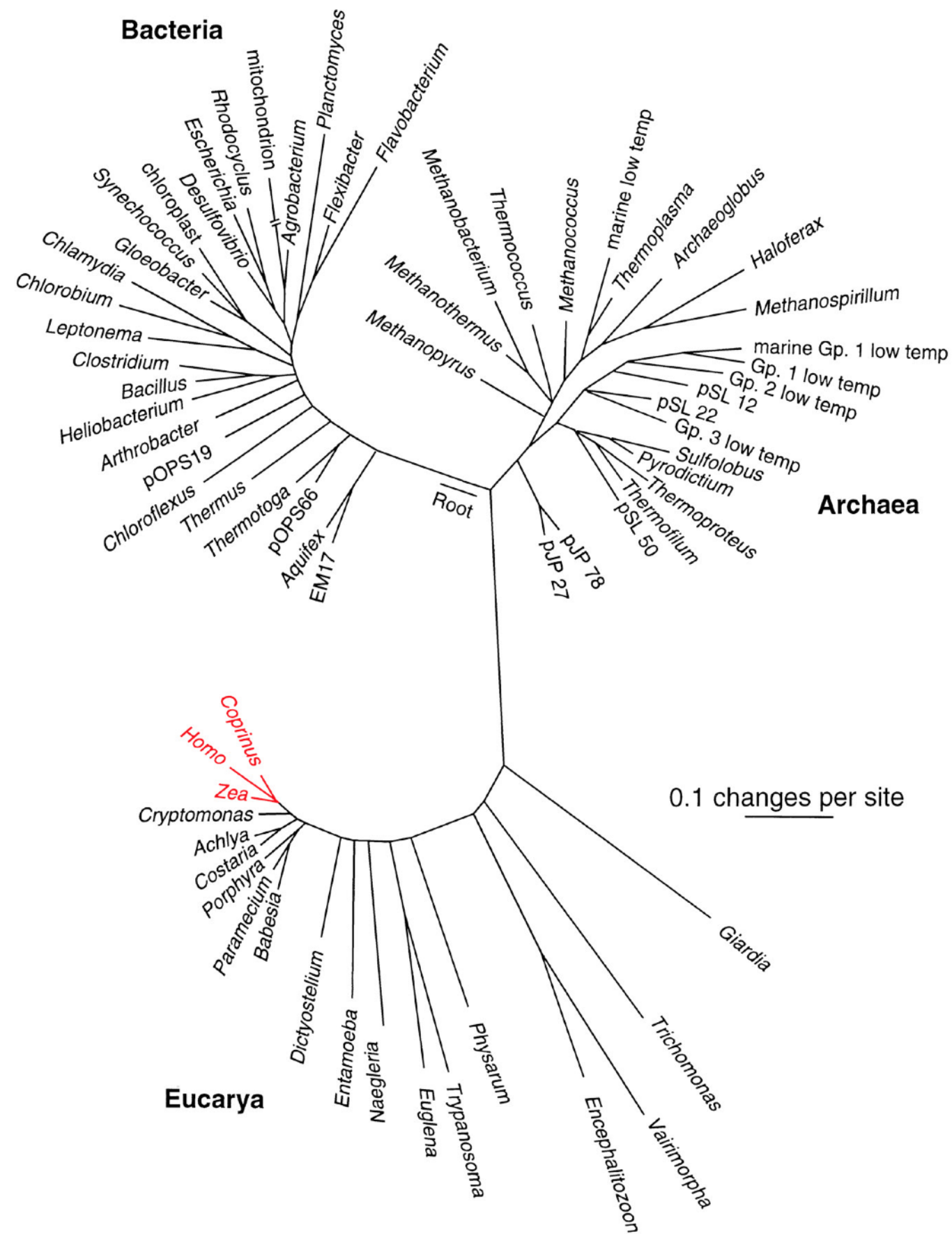


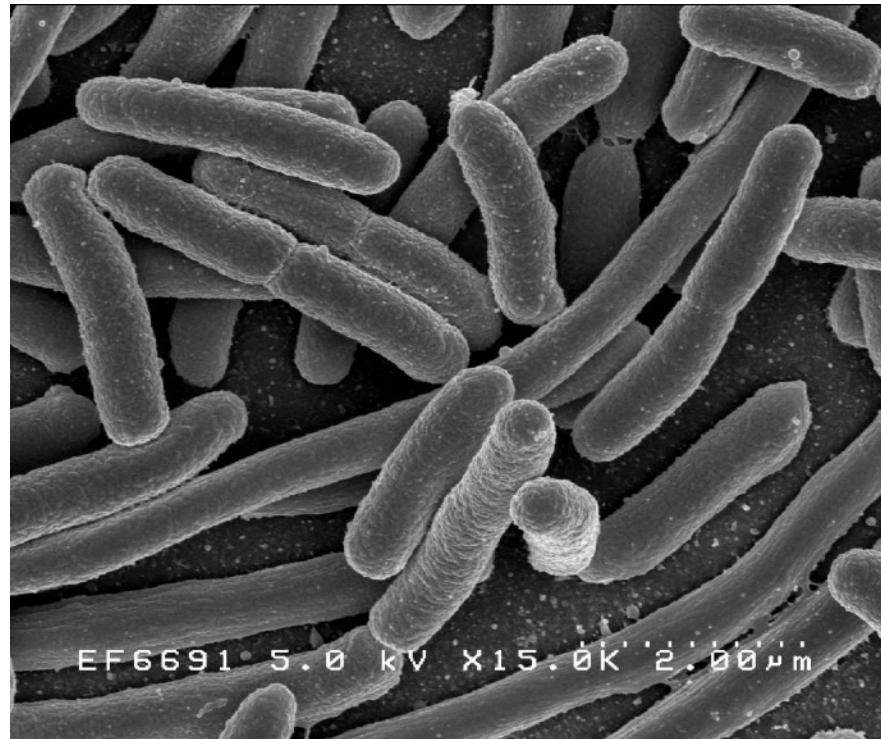
OTUs and ASVs

Diversity of Life



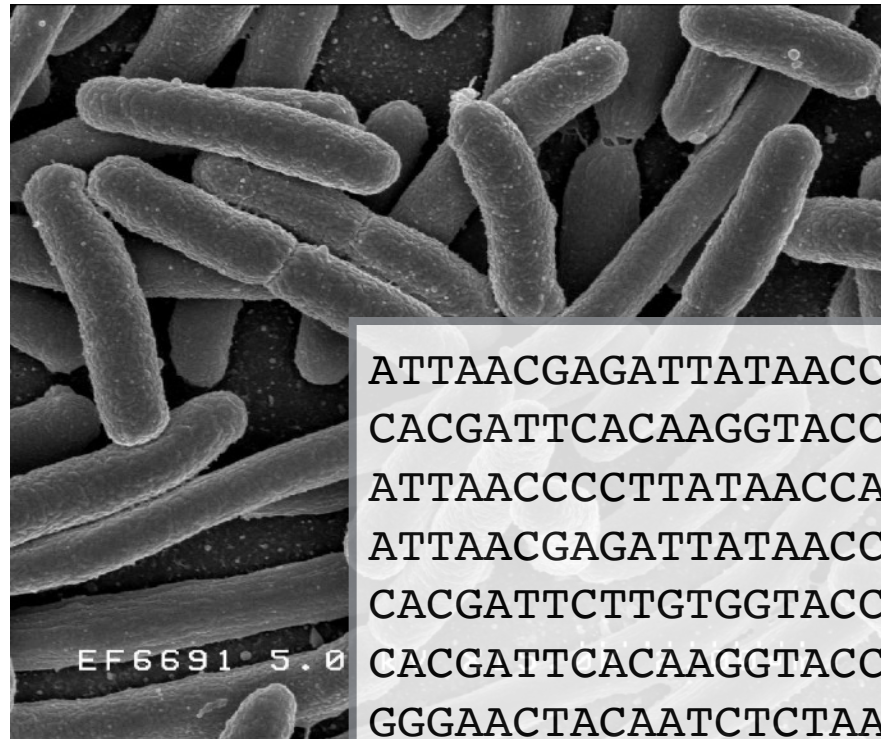
A Microbial Census

Metabarcoding or Marker-gene Sequencing



A Microbial Census

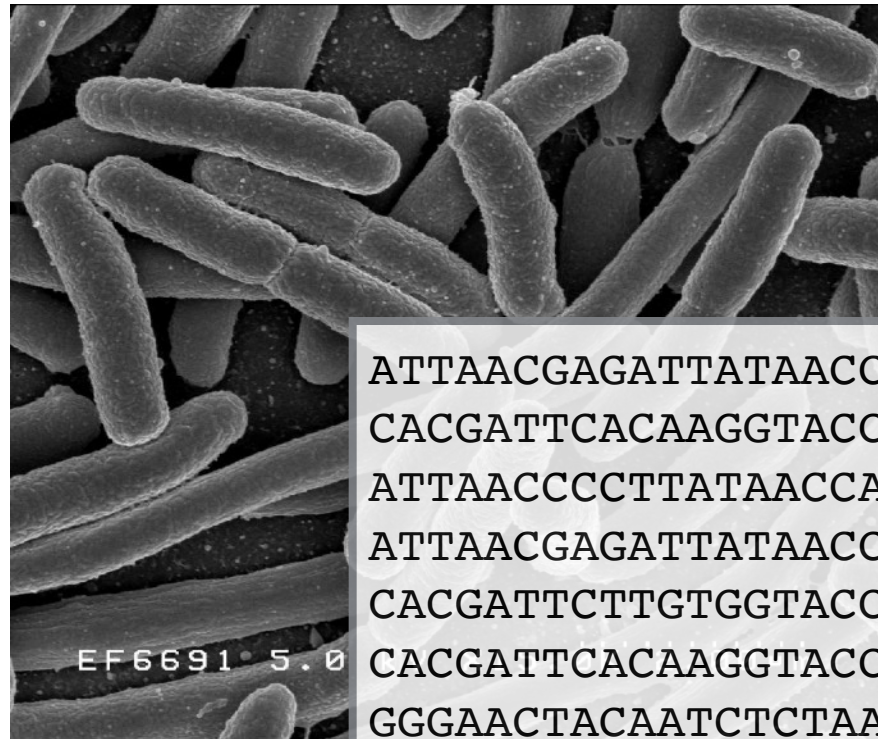
Metabarcoding or Marker-gene Sequencing



```
ATTAACGAGATTATAACCAGAGTACGAATACCGAAC  
CACGATTCACAAGGTACCACAAGGTAACATAGCTCC  
ATTAACCCCTTATAACCAGAGTACGAATACCGAAC  
ATTAACGAGATTATAACCAGAGAGAGAATACCGAAC  
CACGATTCTTGTGGTACCACAAGGTAACATAGCTCC  
CACGATTCACAAGGTACCACAAGGTAACATAGCTCC  
GGGA ACTACAATCTCTAAGGTGAAGTCTCAGTCTAT  
ATTAACGAGATTATAACCAGAGTACGAATACCGAAC  
CACGATTCACAAGGTACCACAAGGTAACATAGCTCC  
ATTAACGAGATTATAACCAGAGTACGAATACCGAAC
```

A Microbial Census

Metabarcoding or Marker-gene Sequencing

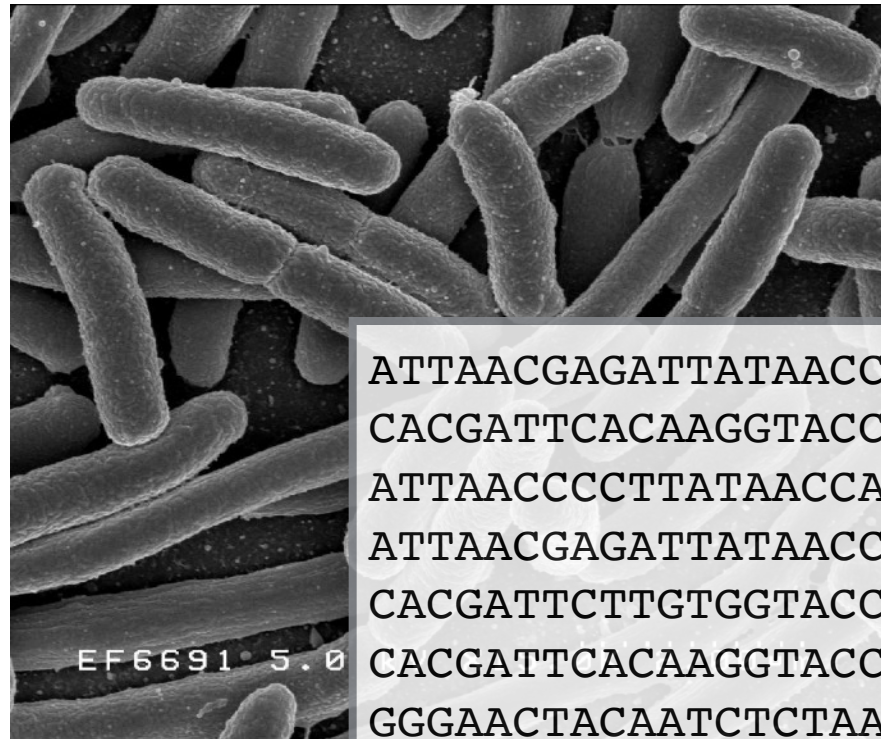


ATTAACGAGATTATAACCAGAGTACGAATACCGAAC
CACGATTCACAAGGTACCACAAGGTAACATAGCTCC
ATTAACCCCTTATAACCAGAGTACGAATACCGAACA
ATTAACGAGATTATAACCAGAGAGAGAATACCGAAC
CACGATTCTTGTGGTACCACAAGGTAACATAGCTCC
CACGATTCACAAGGTACCACAAGGTAACATAGCTCC
GGGA ACTACAATCTCTAAGGTGAAGTCTCAGTCTAT
ATTAACGAGATTATAACCAGAGTACGAATACCGAAC
CACGATTCACAAGGTACCACAAGGTAACATAGCTCC
ATTAACGAGATTATAACCAGAGTACGAATACCGAAC

<i>Lactobacillus crispatus</i>	1300	5	0	882	596
<i>Ureaplasma urealytica</i>	15	0	220	0	0
<i>Gardnerella vaginalis</i>	22	0	1	0	412
<i>Prevotella intermedia</i>	0	0	8	12	0
...

A Microbial Census

Metabarcoding or Marker-gene Sequencing



ATTAACGAGATTATAACCAGAGTACGAATACCGAAC
CACGATTCACAAGGTACCACAAGGTAACATAGCTCC
ATTAACCCCTTATAACCAGAGTACGAATACCGAAC
ATTAACGAGATTATAACCAGAGAGAGAATACCGAAC
CACGATTCTTGTGGTACCACAAGGTAACATAGCTCC
CACGATTCACAAGGTACCACAAGGTAACATAGCTCC
GGGAAC TACAATCTCTAAGGTGAAGTCTCAGTCTAT
ATTAACGAGATTATAACCAGAGTACGAATACCGAAC
CACGATTCACAAGGTACCACAAGGTAACATAGCTCC
ATTAACGAGATTATAACCAGAGTACGAATACCGAAC

<i>Lactobacillus crispatus</i>	1300	5	0	882	596
<i>Ureaplasma urealytica</i>	15	0	220	0	0
<i>Gardnerella vaginalis</i>	22	0	1	0	412
<i>Prevotella intermedia</i>	0	0	8	12	0
...

Visualization

Exploration

Inference

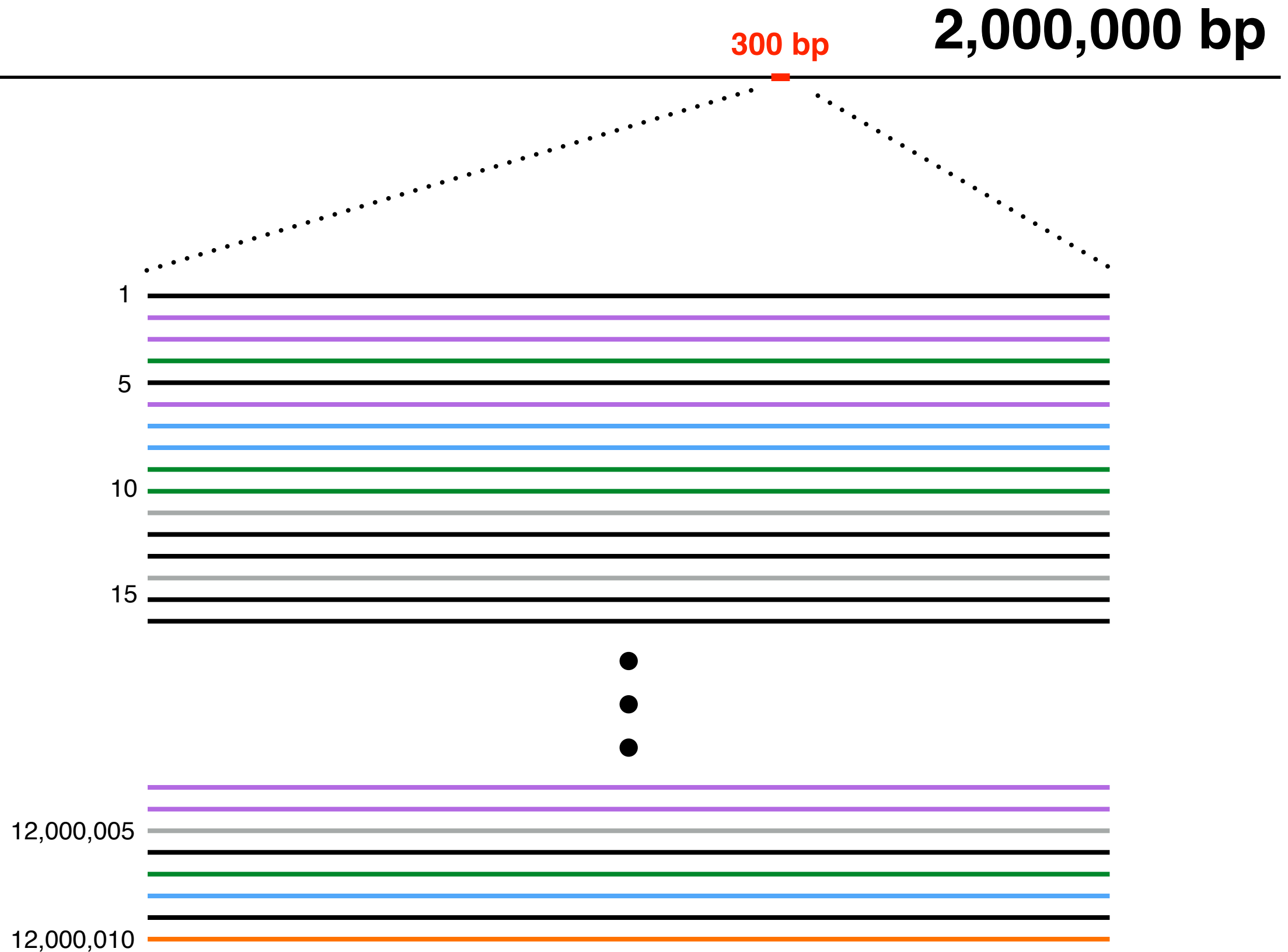
Marker-gene Sequencing

300 bp

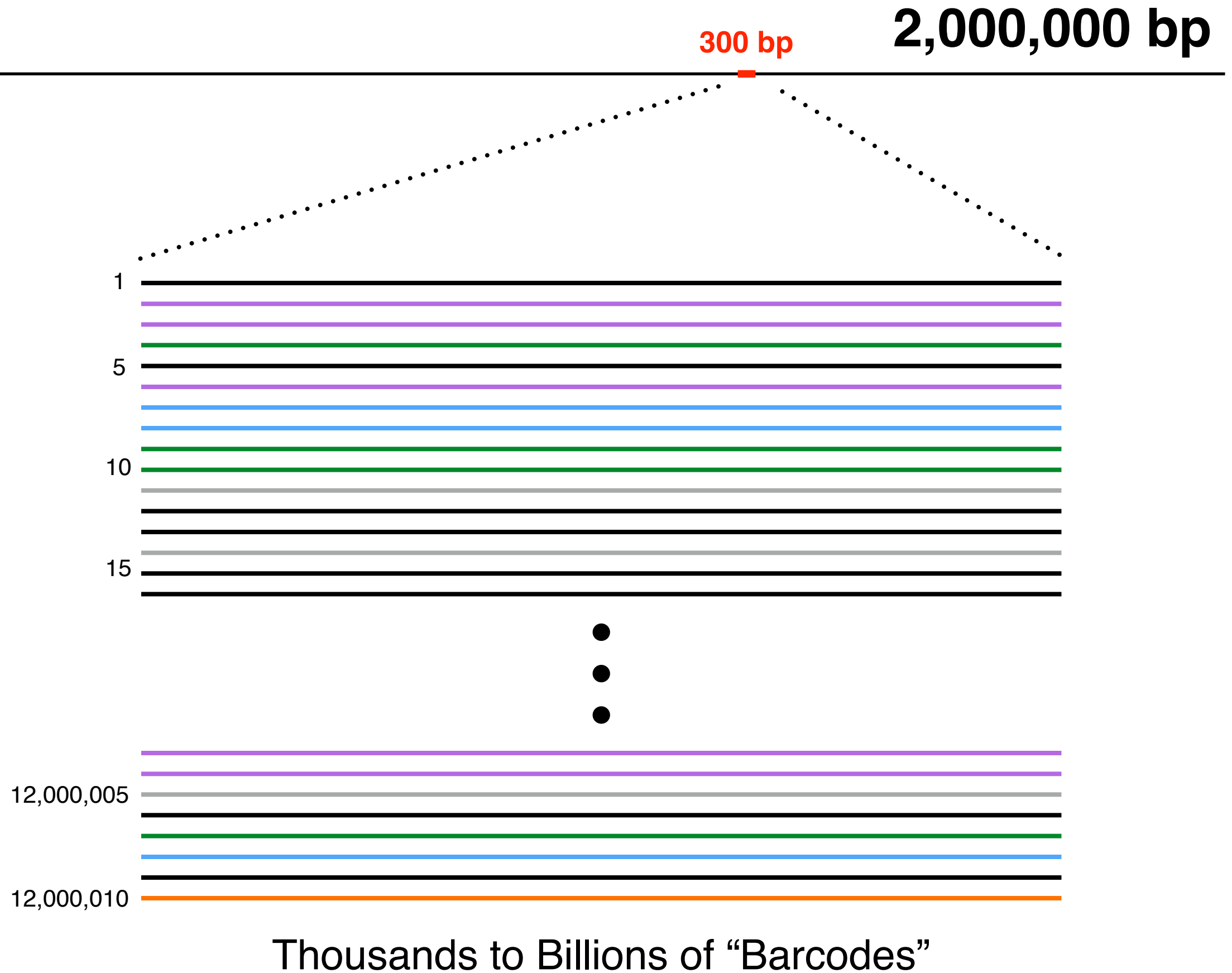
2,000,000 bp



Marker-gene Sequencing

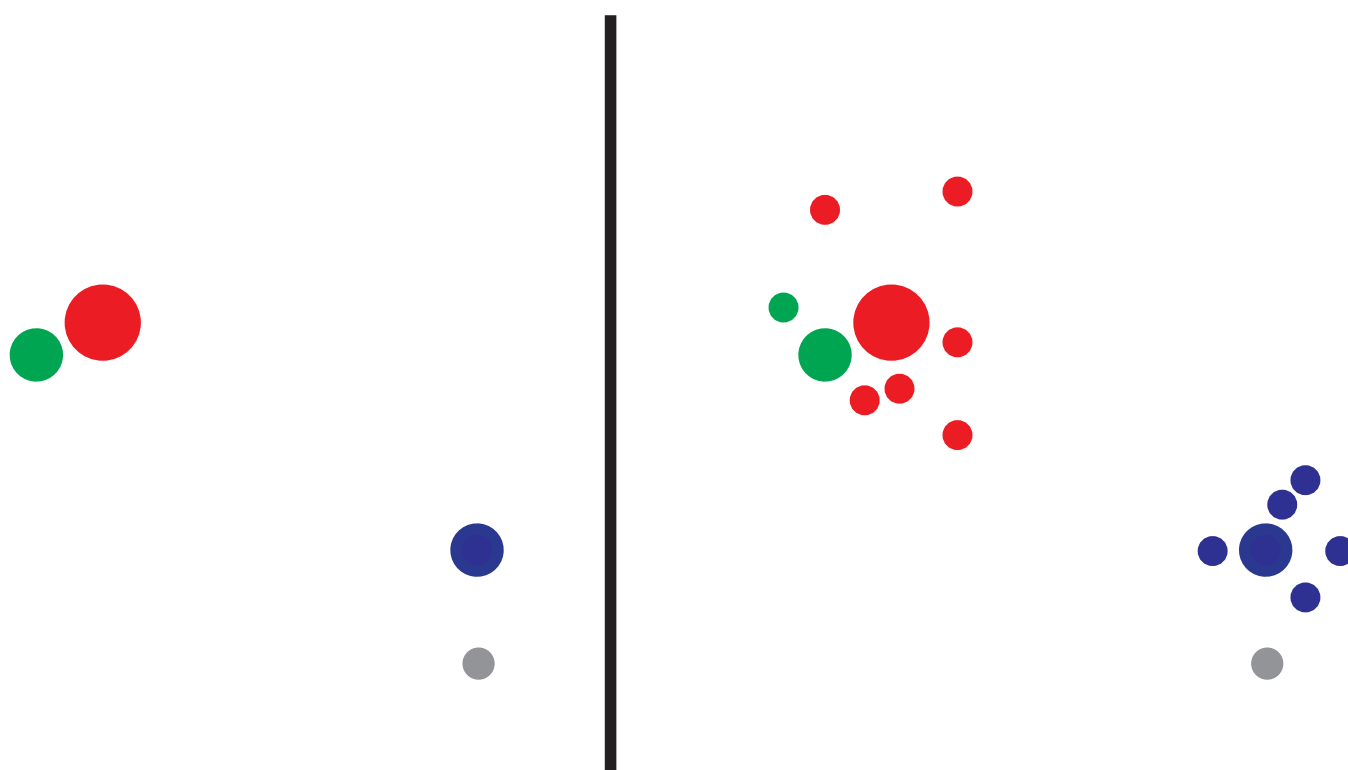


Marker-gene Sequencing



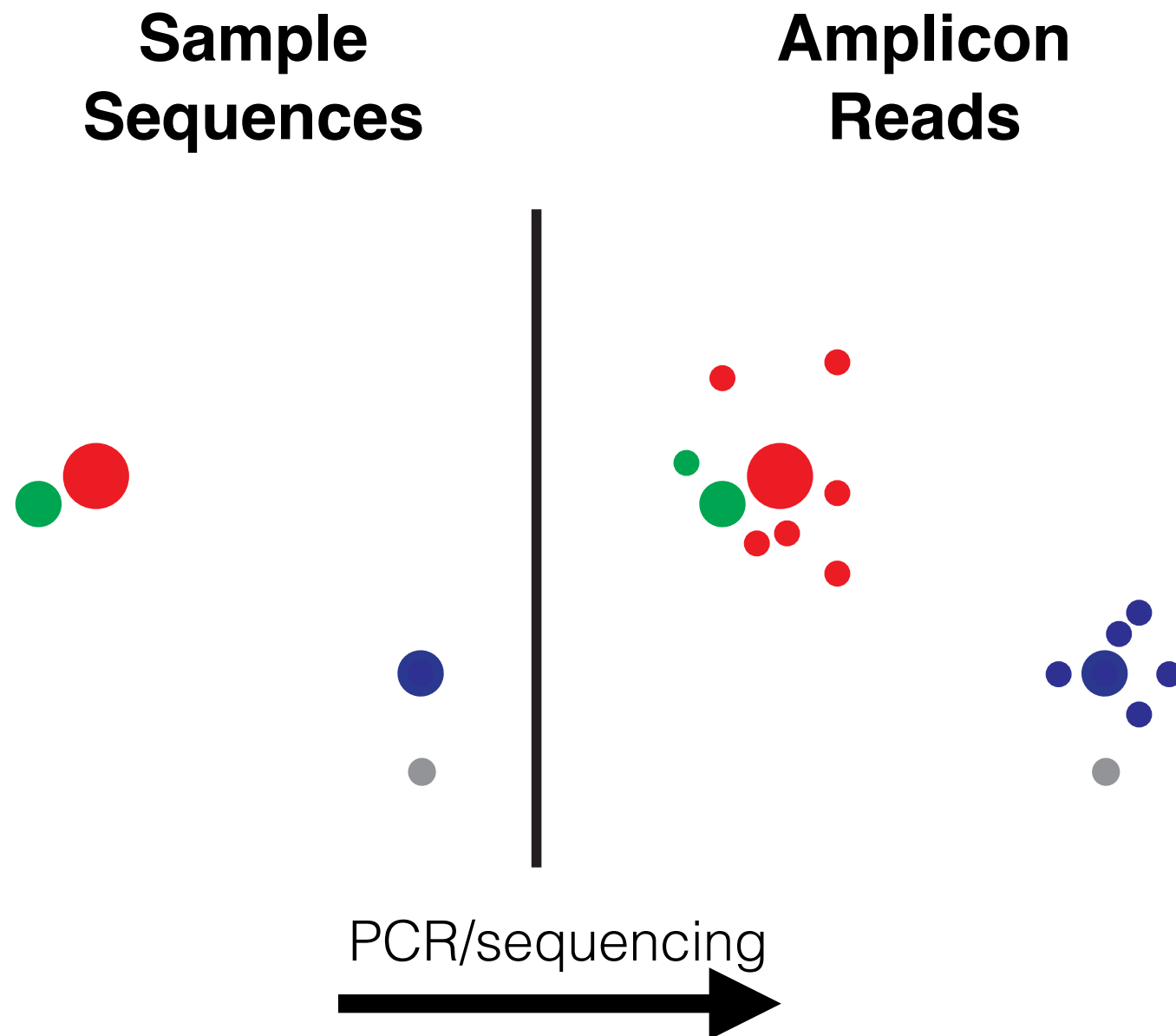
**Sample
Sequences**

**Amplicon
Reads**



PCR/sequencing





Two Challenges

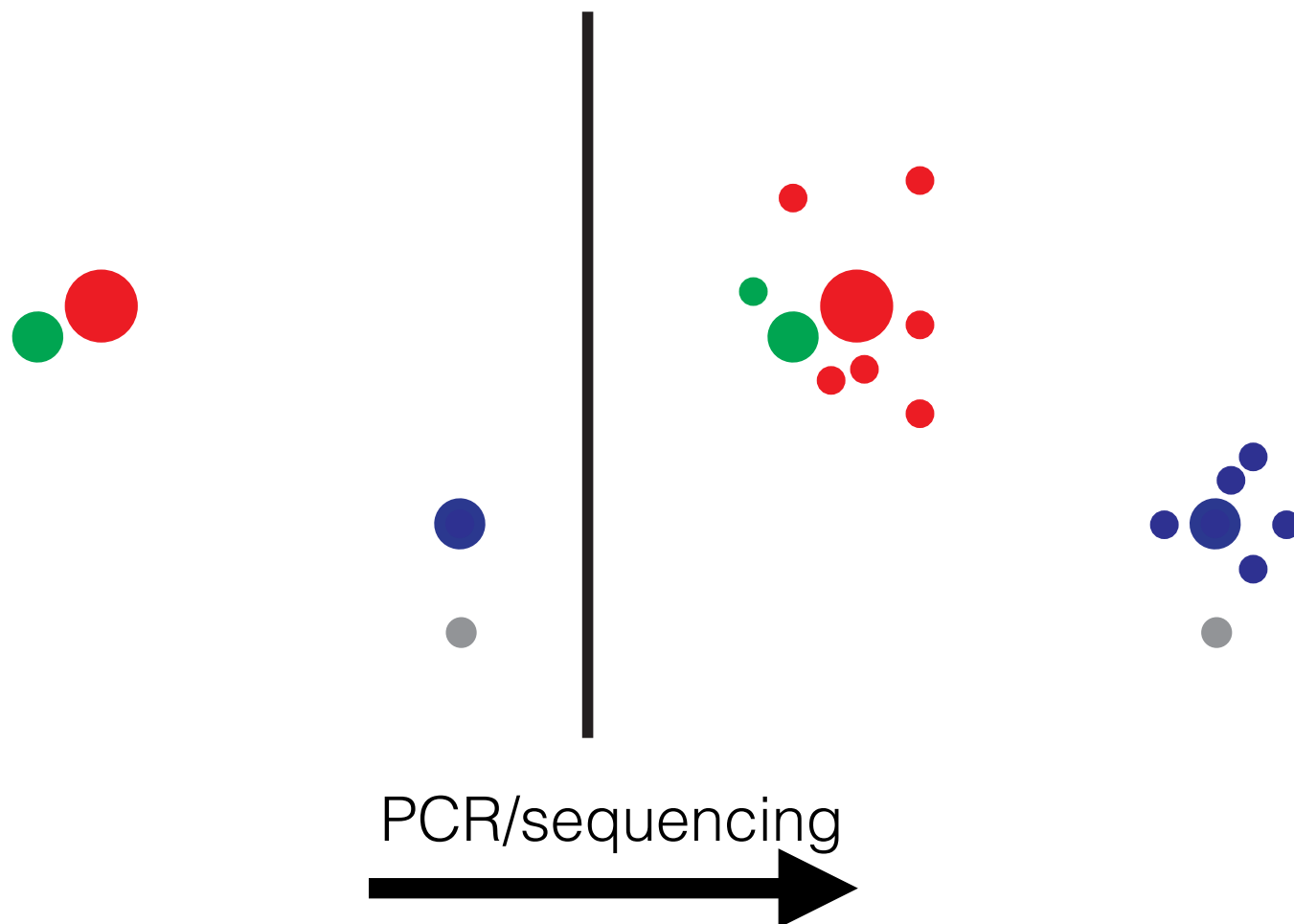
How do we deal with errors?

How do we define the units of our analysis?

Closed reference OTUs

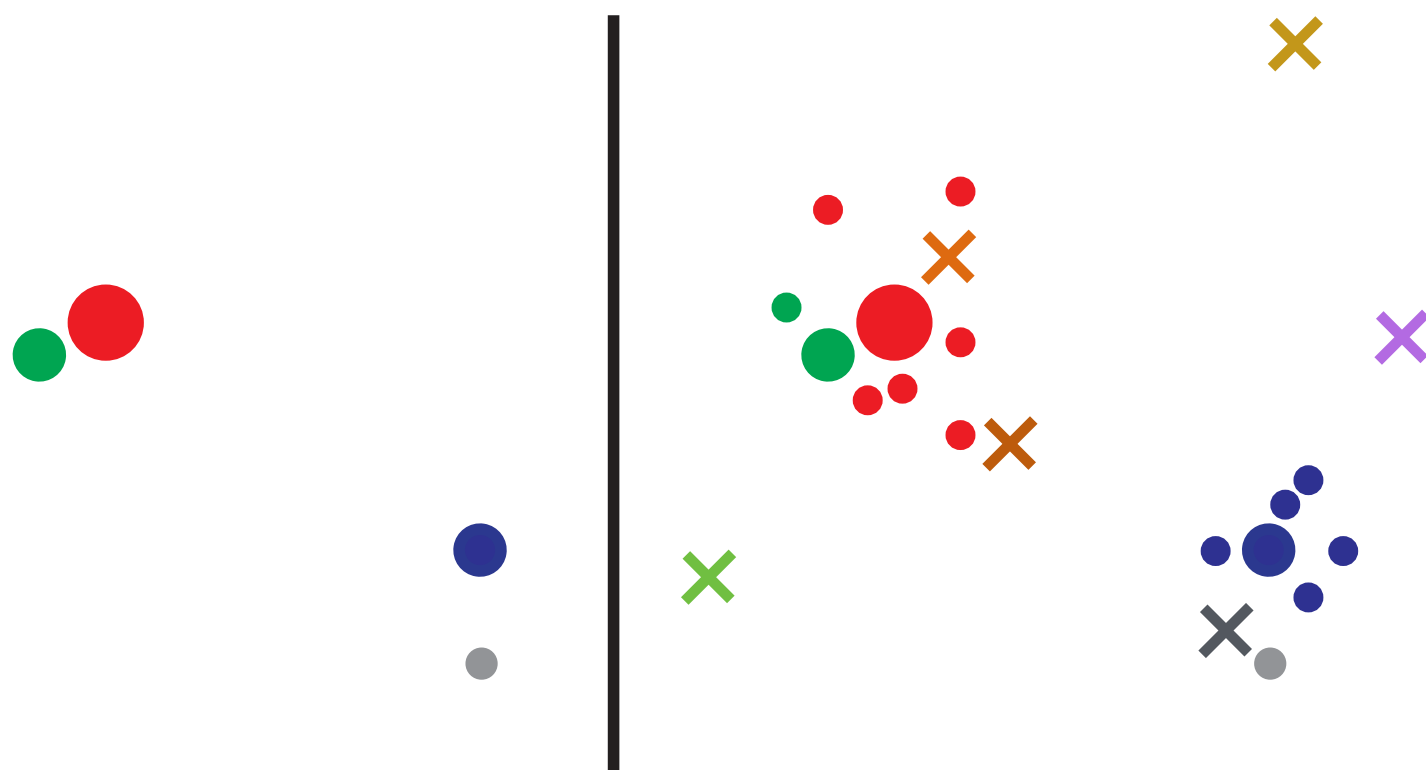
**Sample
Sequences**

**Amplicon
Reads**

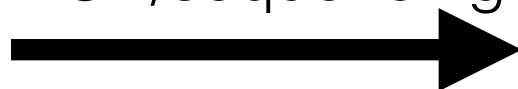


**Sample
Sequences**

**Amplicon
Reads**



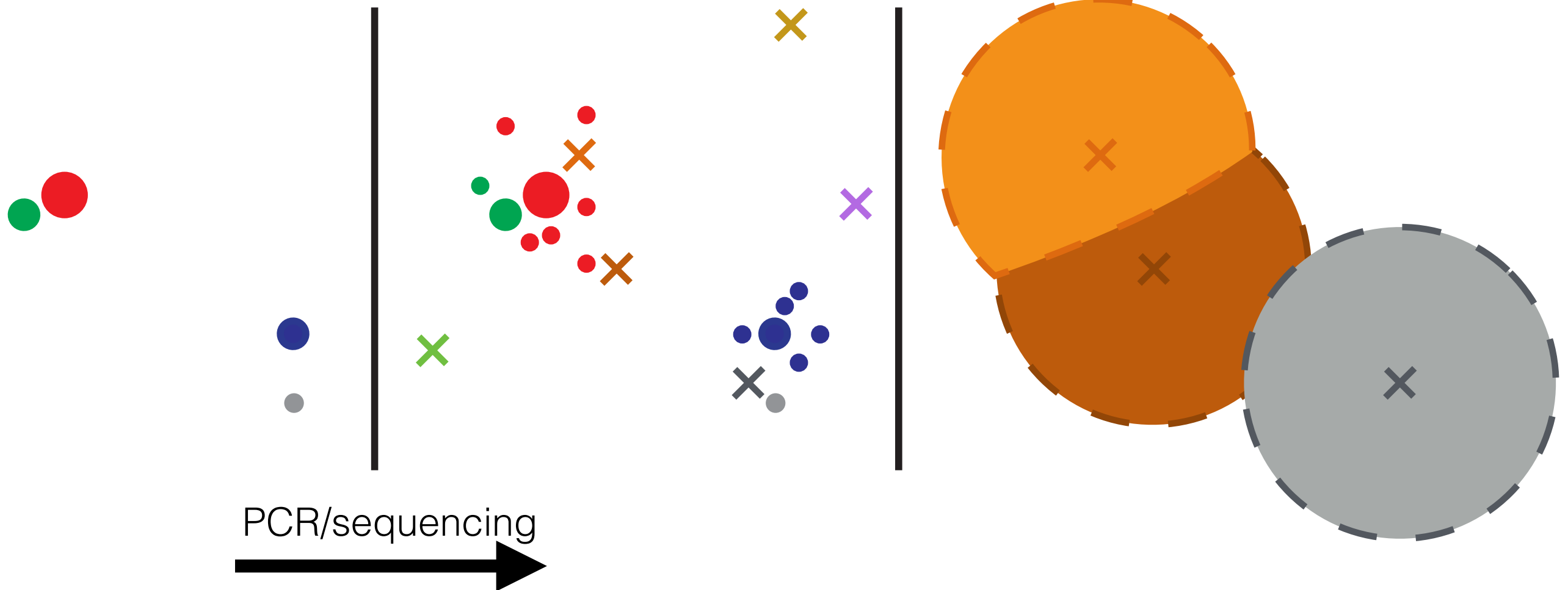
PCR/sequencing



Sample Sequences

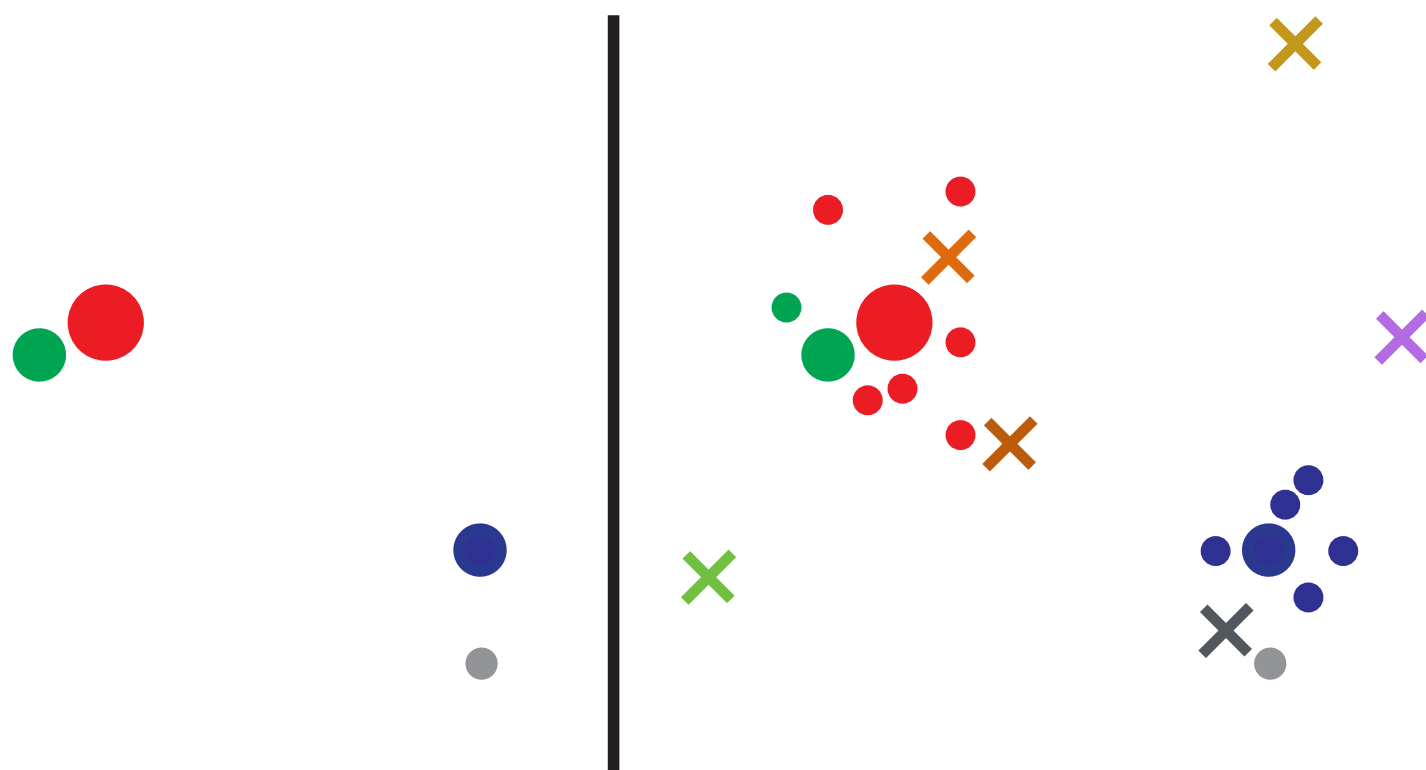
Amplicon Reads

Operational Taxonomic Units (closed reference)

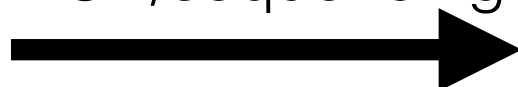


**Sample
Sequences**

**Amplicon
Reads**

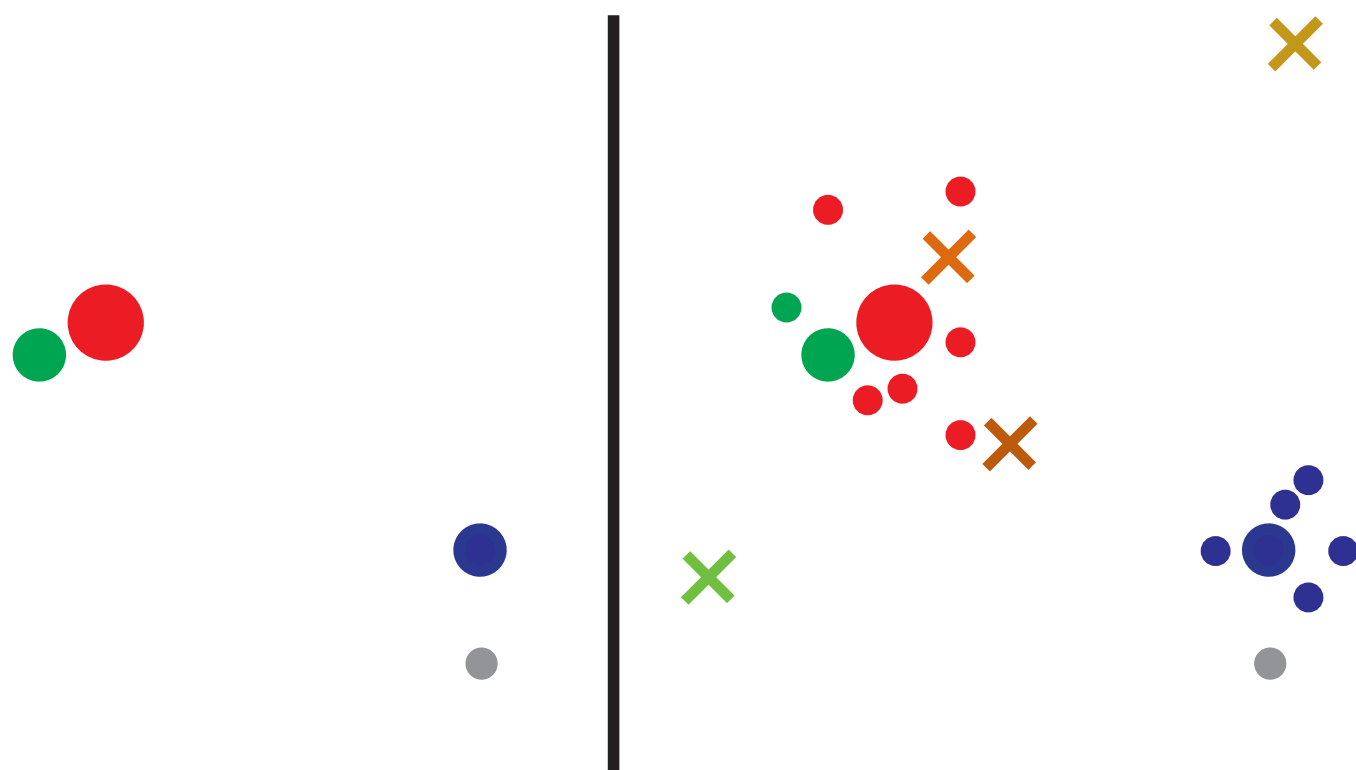


PCR/sequencing

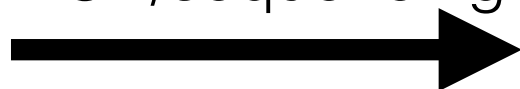


**Sample
Sequences**

**Amplicon
Reads**



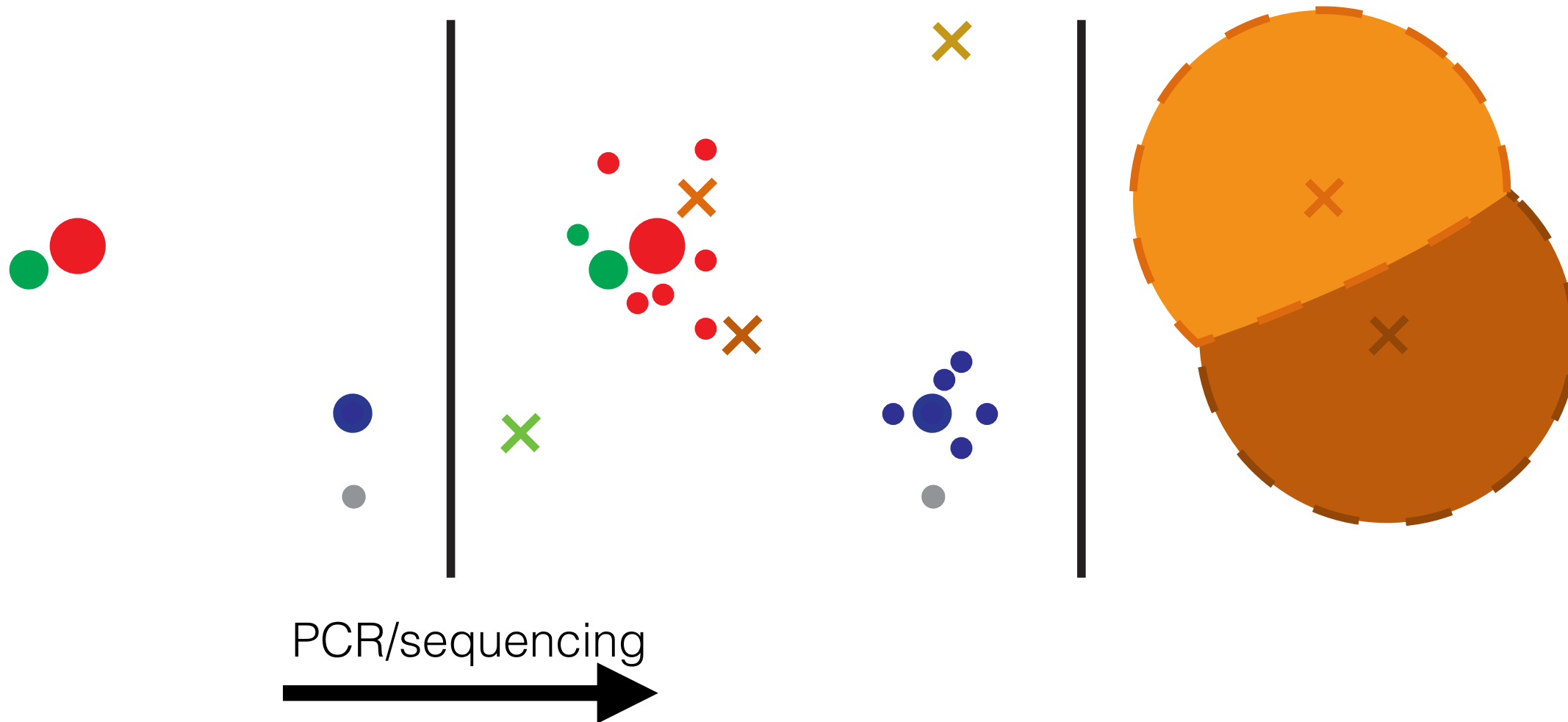
PCR/sequencing



**Sample
Sequences**

**Amplicon
Reads**

**Operational
Taxonomic Units
(closed reference)**

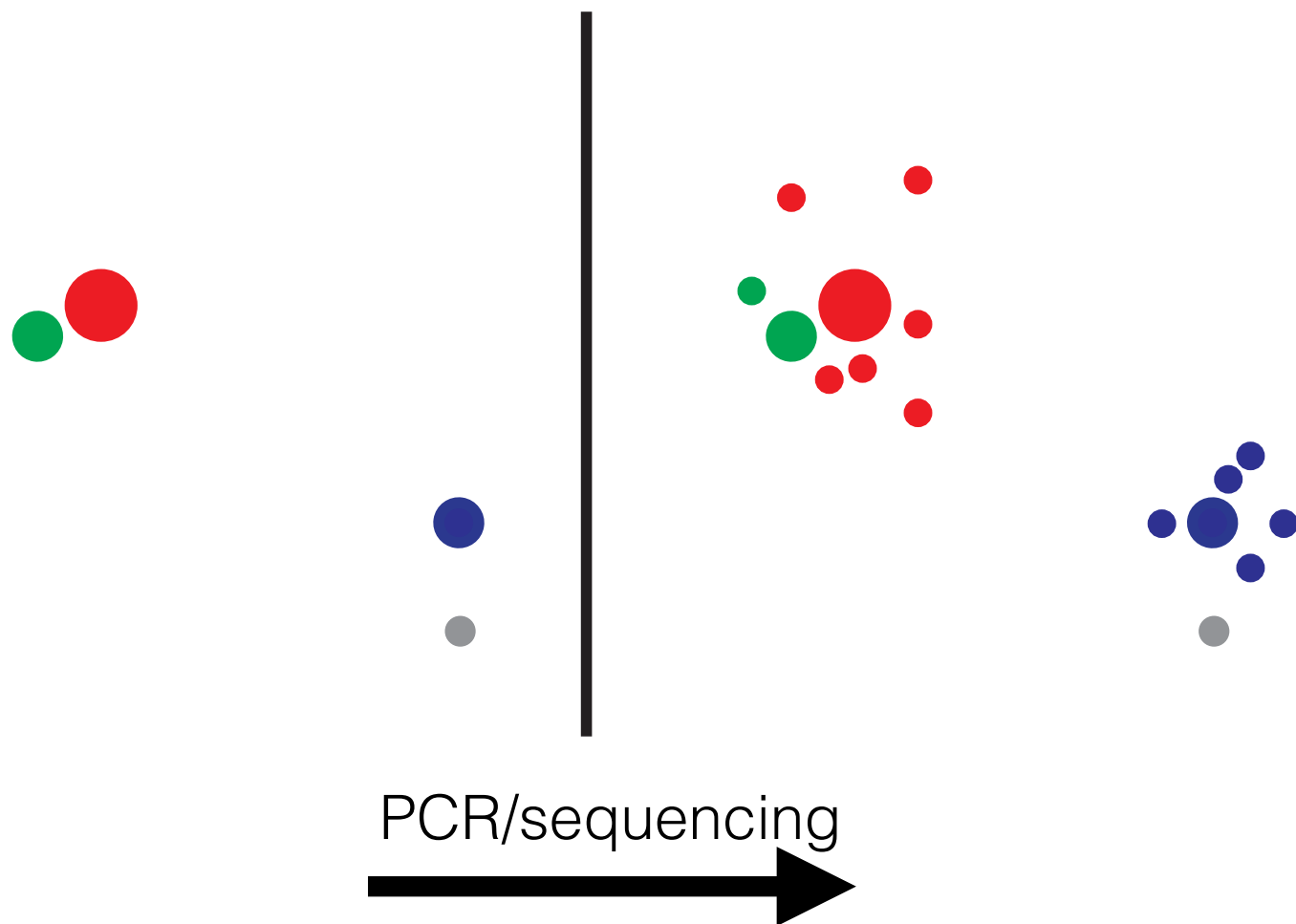


Incomplete OTU table!

de novo OTUs

**Sample
Sequences**

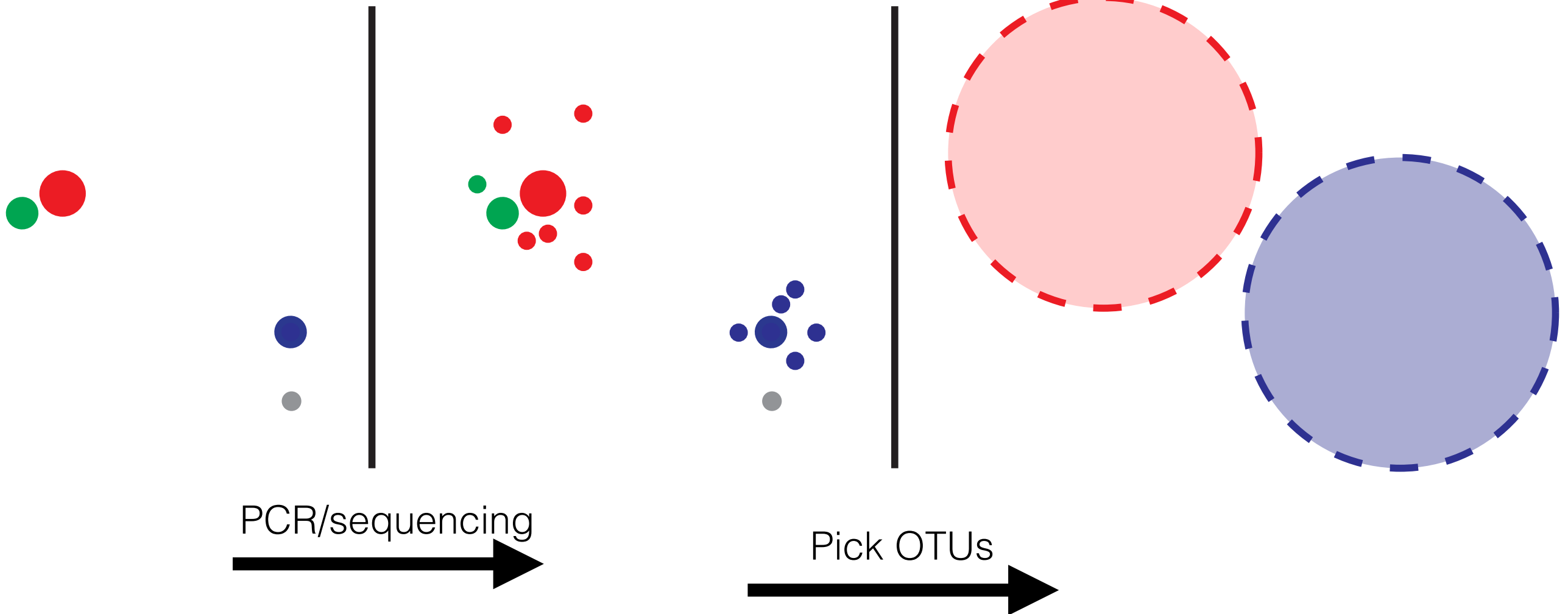
**Amplicon
Reads**



**Sample
Sequences**

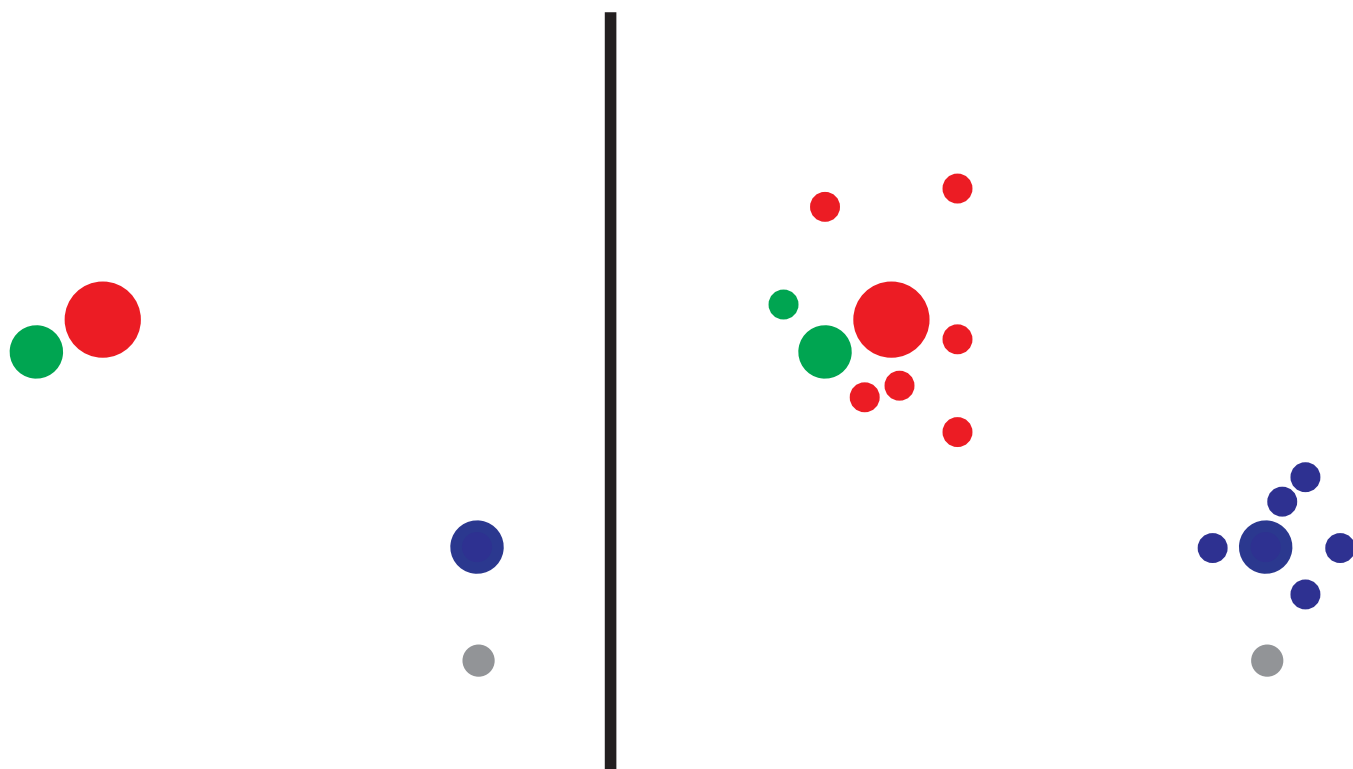
**Amplicon
Reads**

**Operational
Taxonomic Units
(de novo)**

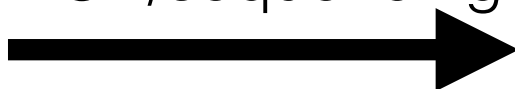


**Sample
Sequences**

**Amplicon
Reads**

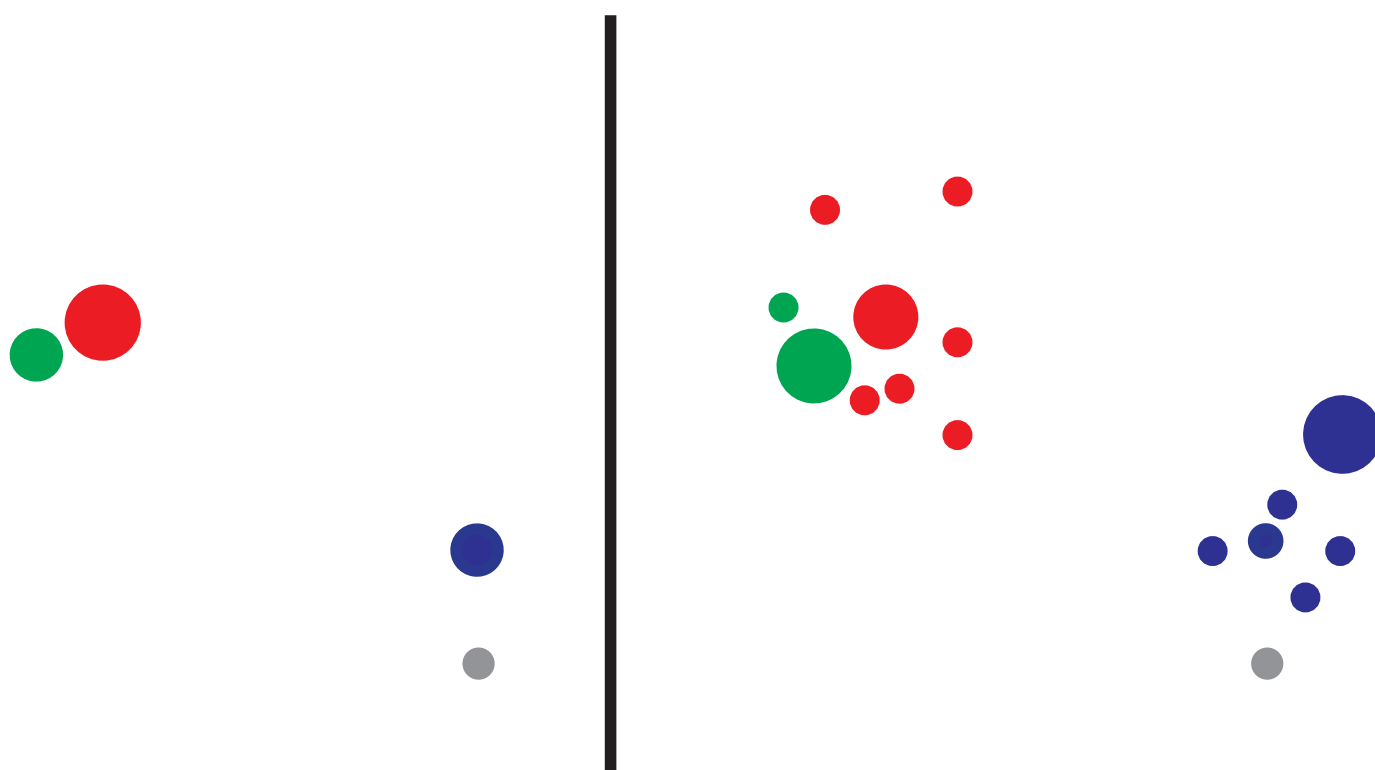


PCR/sequencing



**Sample
Sequences**

**Amplicon
Reads**



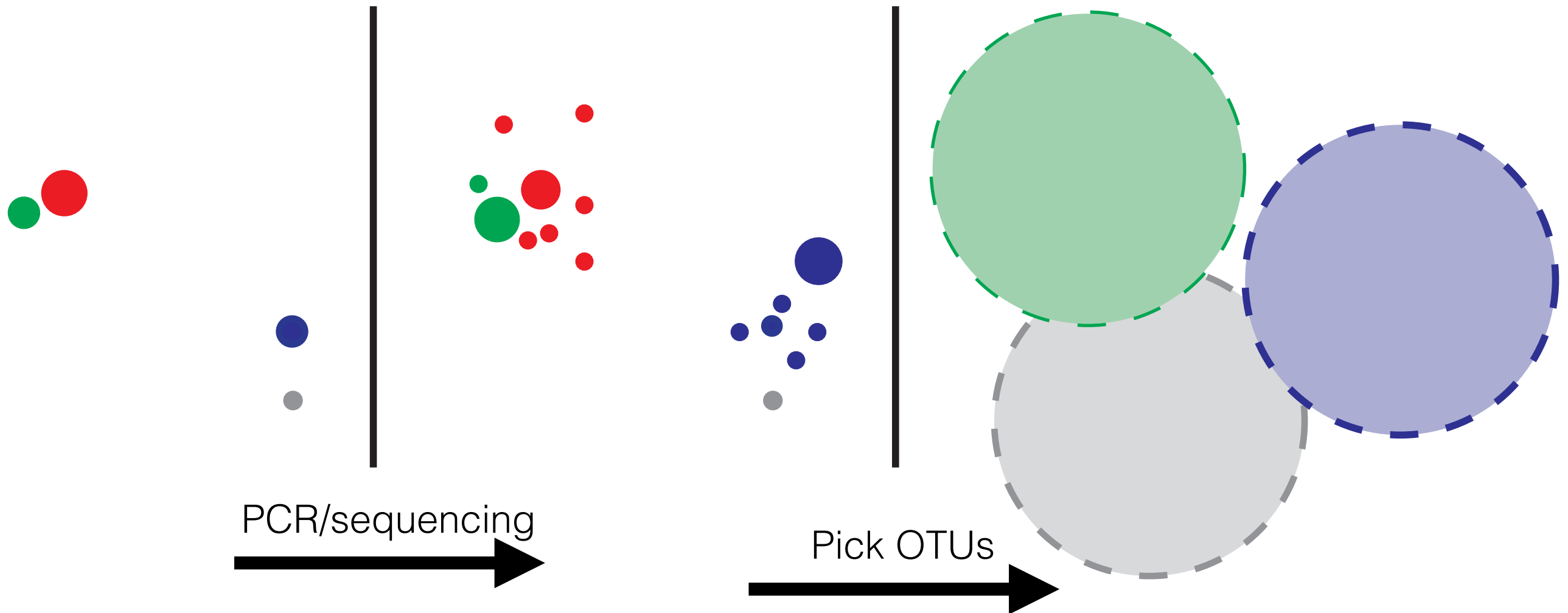
PCR/sequencing



**Sample
Sequences**

**Amplicon
Reads**

**Operational
Taxonomic Units
(de novo)**



Different OTUs!

(In)consistent labels

OTU85 is ***not a consistent label***

OTU85 is *predictive of a disease?* **Not in future data!**

OTU85 is *associated w/ X and Y?* **Can't be tested!**

OTU85 is *in this community?* **OTUs don't exist in nature!**

Exact Sequence Variants...

OPEN

The ISME Journal (2017), 1–5

www.nature.com/ismej

PERSPECTIVE

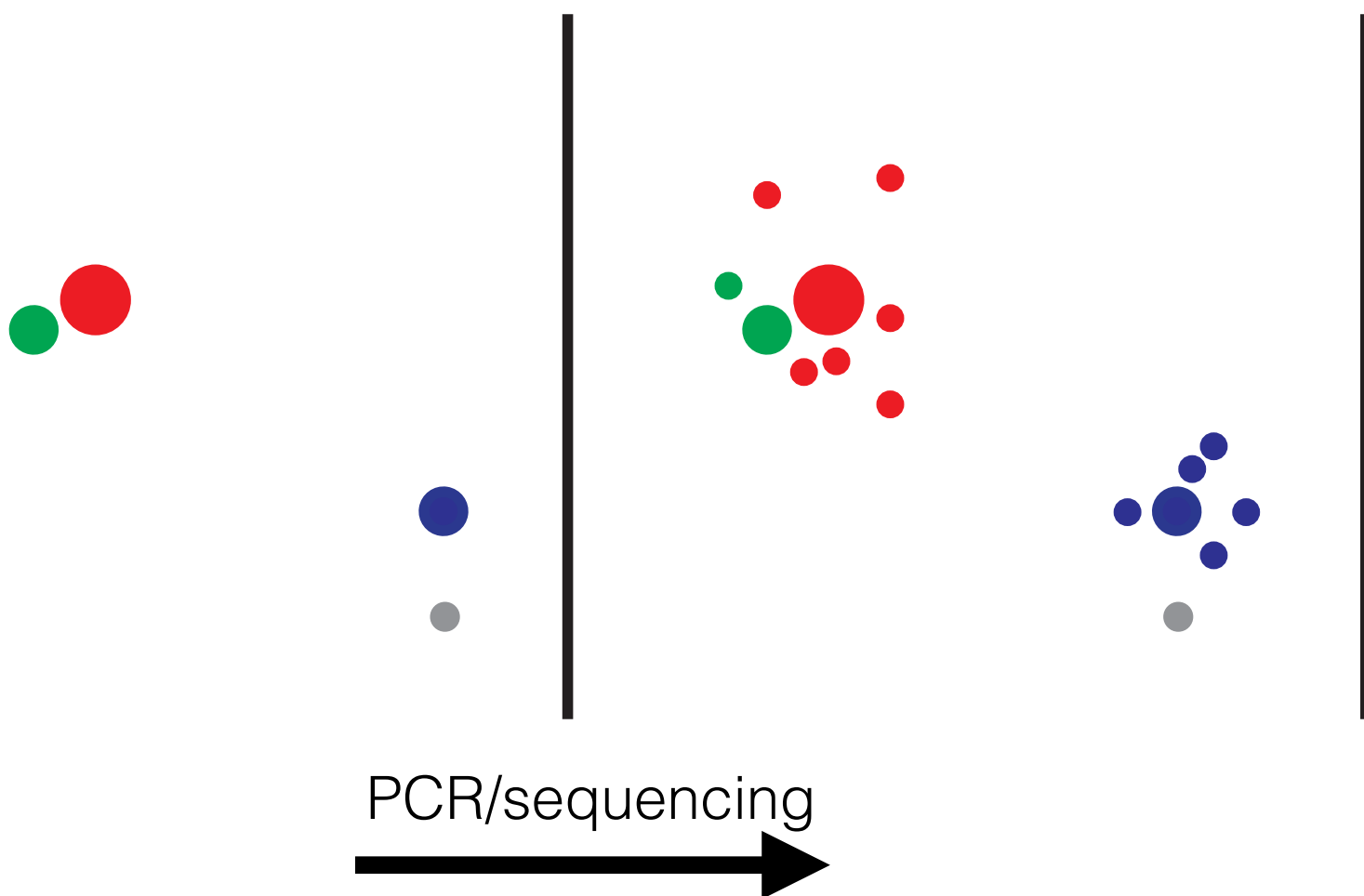
Exact sequence variants should replace operational taxonomic units in marker-gene data analysis

Benjamin J Callahan¹, Paul J McMurdie² and Susan P Holmes³

¹*Department of Population Health and Pathobiology, NC State University, Raleigh NC, USA;* ²*Whole Biome Inc, San Francisco CA, USA and* ³*Department of Statistics, Stanford University, Stanford CA, USA*

**Sample
Sequences**

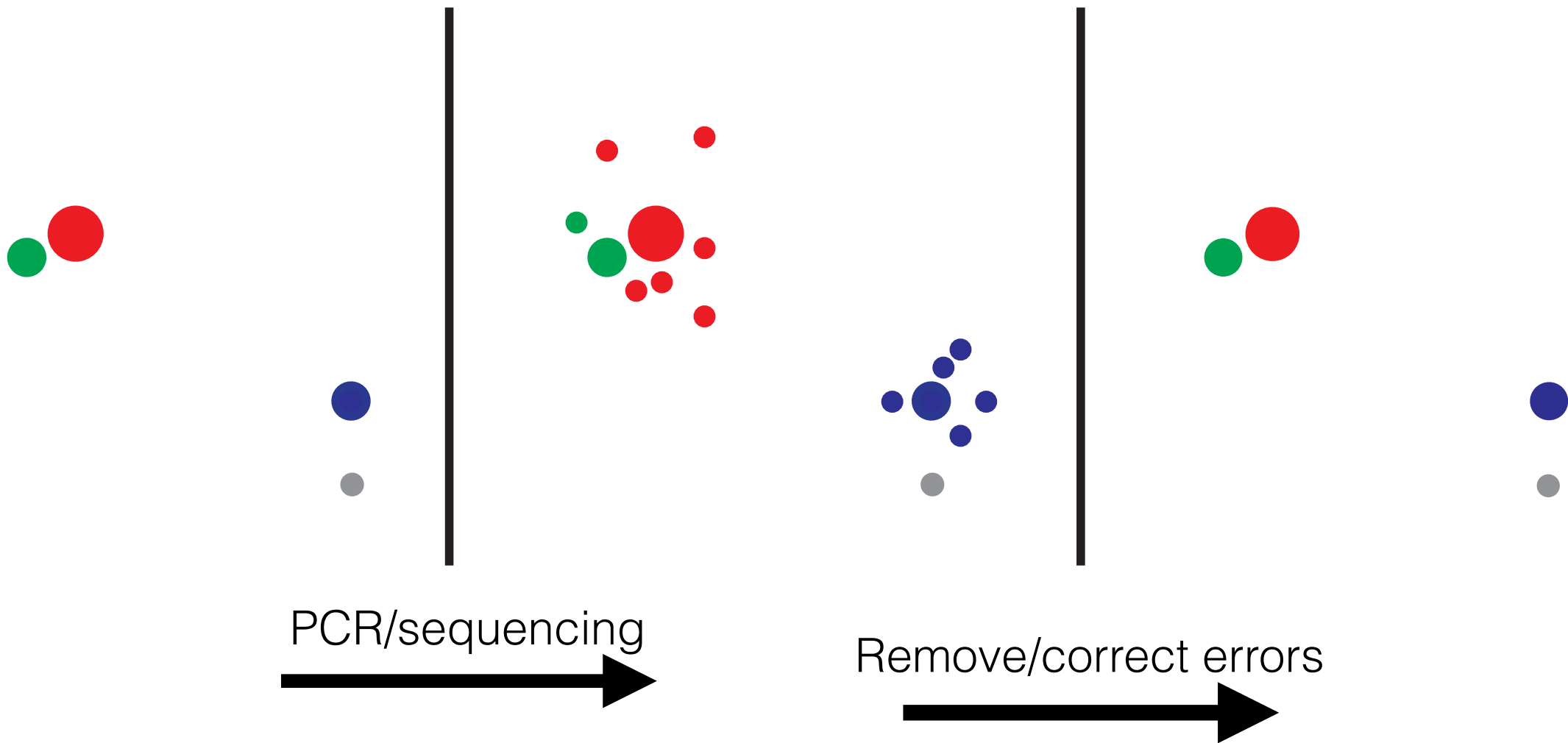
**Amplicon
Reads**



**Sample
Sequences**

**Amplicon
Reads**

**Amplicon
Sequence Variants
(ASVs)**



Exact Sequence Variants

OTU85 is ***not*** a consistent label, but...

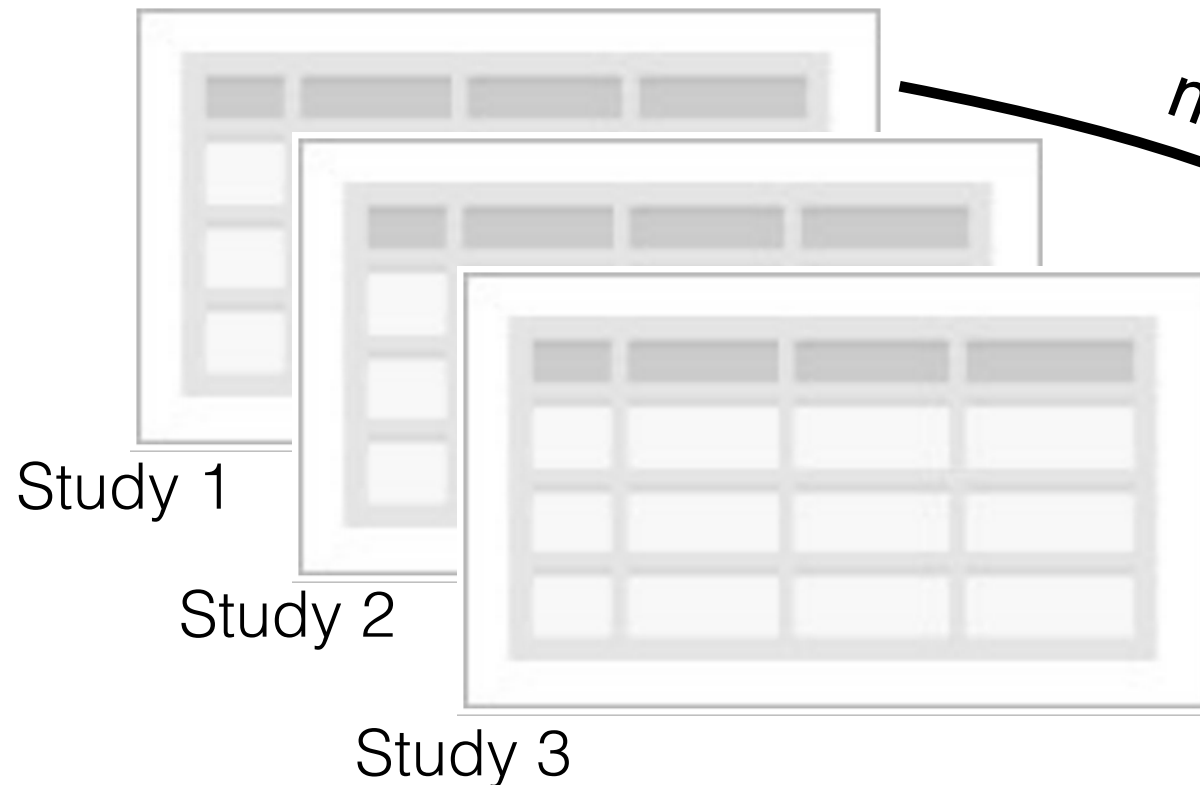
ATTAAACGAGATTATAACCCAGAGTACGAATA...

is consistent!



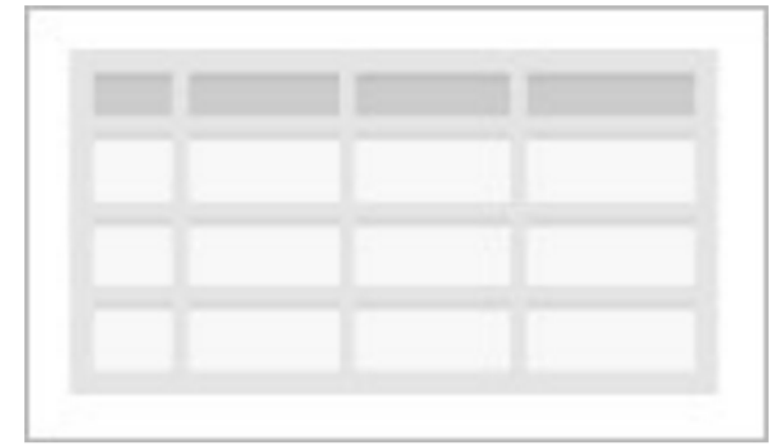
The Sequence is the Label

Sequence Tables



merge

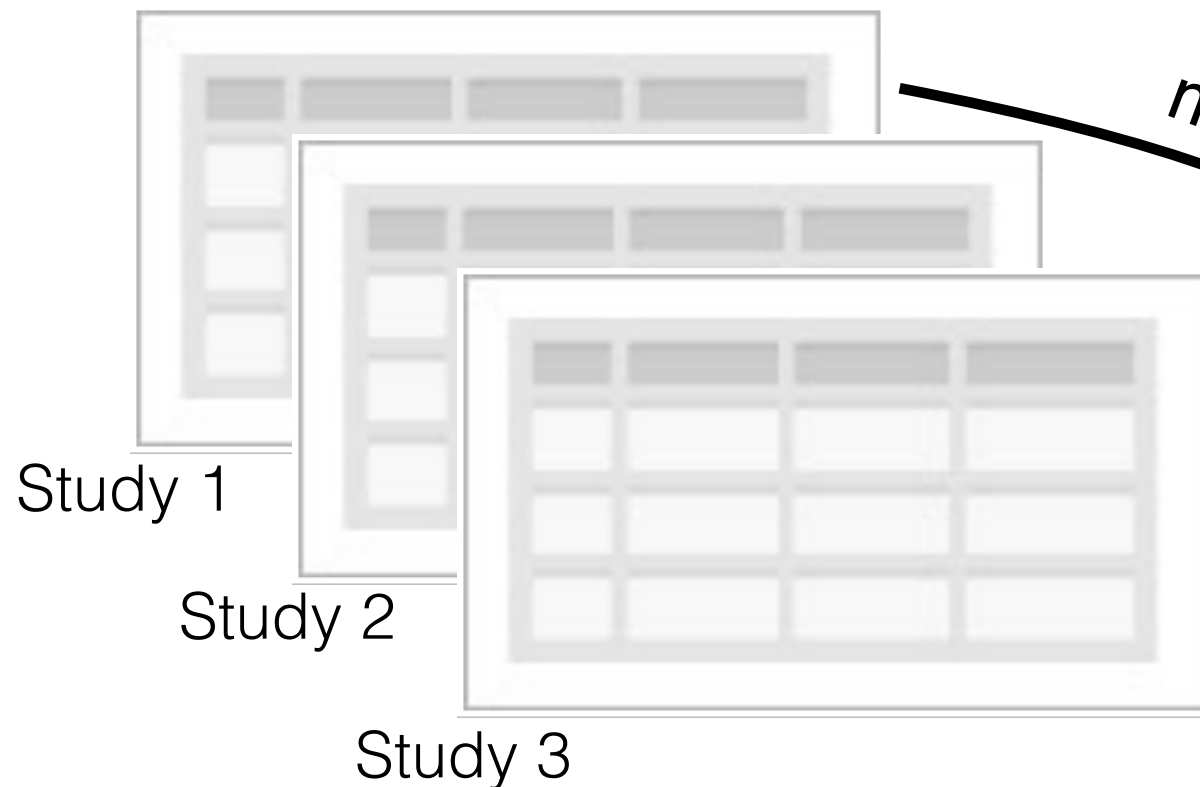
Cross-study comparison



Eliminates need for joint reprocessing of raw data.

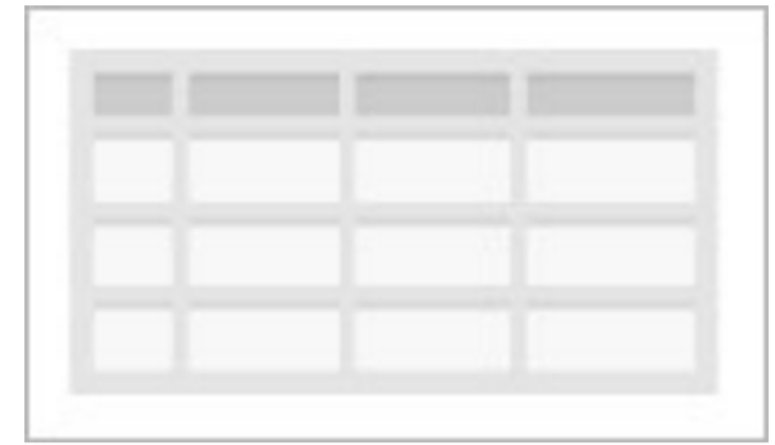
The Sequence is the Label

Sequence Tables



merge

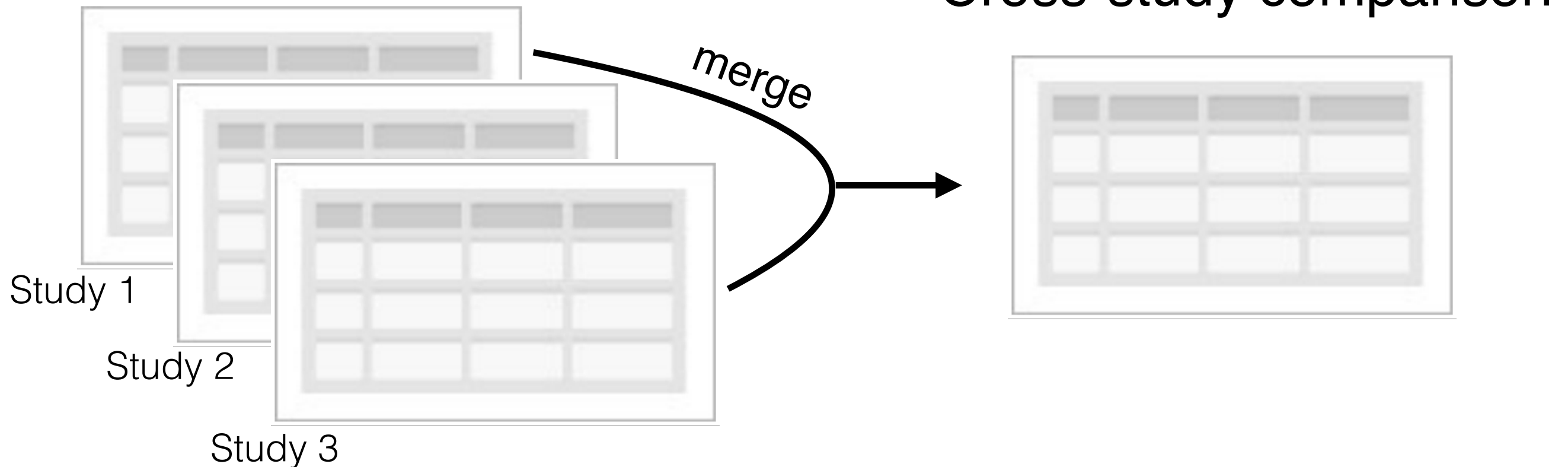
Cross-study comparison



Eliminates need for joint reprocessing of raw data.
Continuous data integration. Unlimited dataset size.

The Sequence is the Label

Sequence Tables

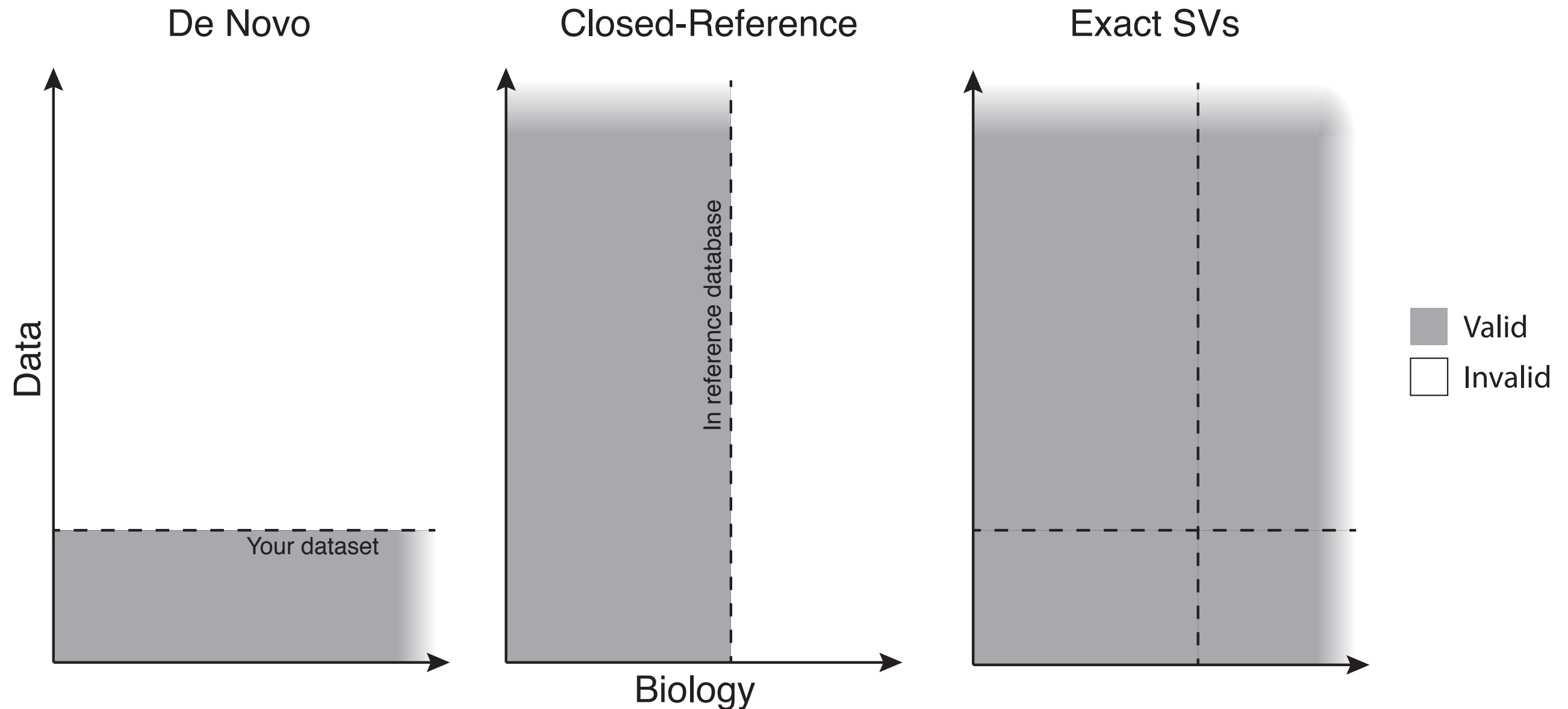


Eliminates need for joint reprocessing of raw data.

Continuous data integration. Unlimited dataset size.

You in 2 years can work directly with you today.

The Sequence is the Label



“Replacing OTUs with ASVs makes marker-gene sequencing more precise, reusable, reproducible and comprehensive.”

A rose by another name

Amplicon Sequence Variants (ASVs)

- Needham et al. 2017
- Callahan et al. 2017

Exact Sequence Variants (ESVs)

- Callahan et al. 2017 (by accident)

sub-OTUs (sOTUs)

- Amir et al. 2017

Zero radius OTUs (zOTUs)

- Edgar 2017

Haplotypes, oligotypes, ...

A rose by another name

Amplicon Sequence Variants (ASVs)

- Needham et al. 2017
- Callahan et al. 2017

Exact Sequence Variants (ESVs)

- Callahan et al. 2017 (by accident)

sub-OTUs (sOTUs)

- Amir et al. 2017

Zero radius OTUs (zOTUs)

- Edgar 2017

Haplotypes, oligotypes, ...

All the same thing! All the same (qualitative) benefits!

ASVs vs. OTUs

	ASVs	De novo	Closed-ref
Precise	✓	~	~
Tractable	✓	~	✓
Reproducible	✓	✗	✓
Comprehensive	✓	✓	✗

ASVs vs. OTUs

	ASVs	De novo	Closed-ref
Precise	✓	~	~
Tractable	✓	~	✓
Reproducible	✓	✗	✓
Comprehensive	✓	✓	✗

ASVs vs. OTUs

	ASVs	De novo	Closed-ref
Precise	✓	~	~
Tractable	✓	~	✓
Reproducible	✓	✗	✓
Comprehensive	✓	✓	✗

Are OTUs *wrong*?

ASVs vs. OTUs

	ASVs	De novo	Closed-ref
Precise	✓	~	~
Tractable	✓	~	✓
Reproducible	✓	✗	✓
Comprehensive	✓	✓	✗

Are OTUs *wrong*? No.

ASVs vs. OTUs

	ASVs	De novo	Closed-ref
Precise	✓	~	~
Tractable	✓	~	✓
Reproducible	✓	✗	✓
Comprehensive	✓	✓	✗

Are OTUs *wrong*? No.

Are ASVs *always* the best unit of analysis?

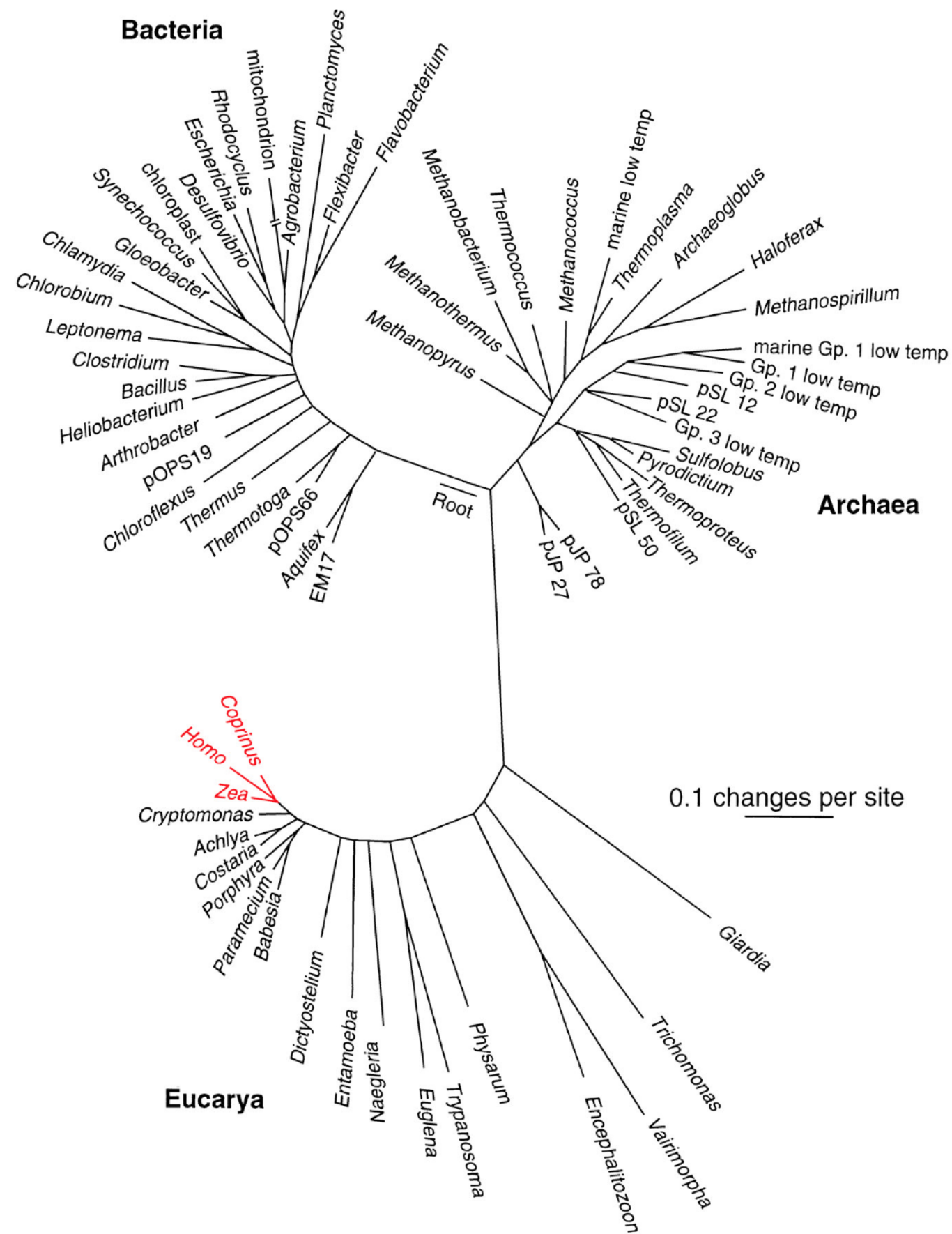
ASVs vs. OTUs

	ASVs	De novo	Closed-ref
Precise	✓	~	~
Tractable	✓	~	✓
Reproducible	✓	✗	✓
Comprehensive	✓	✓	✗

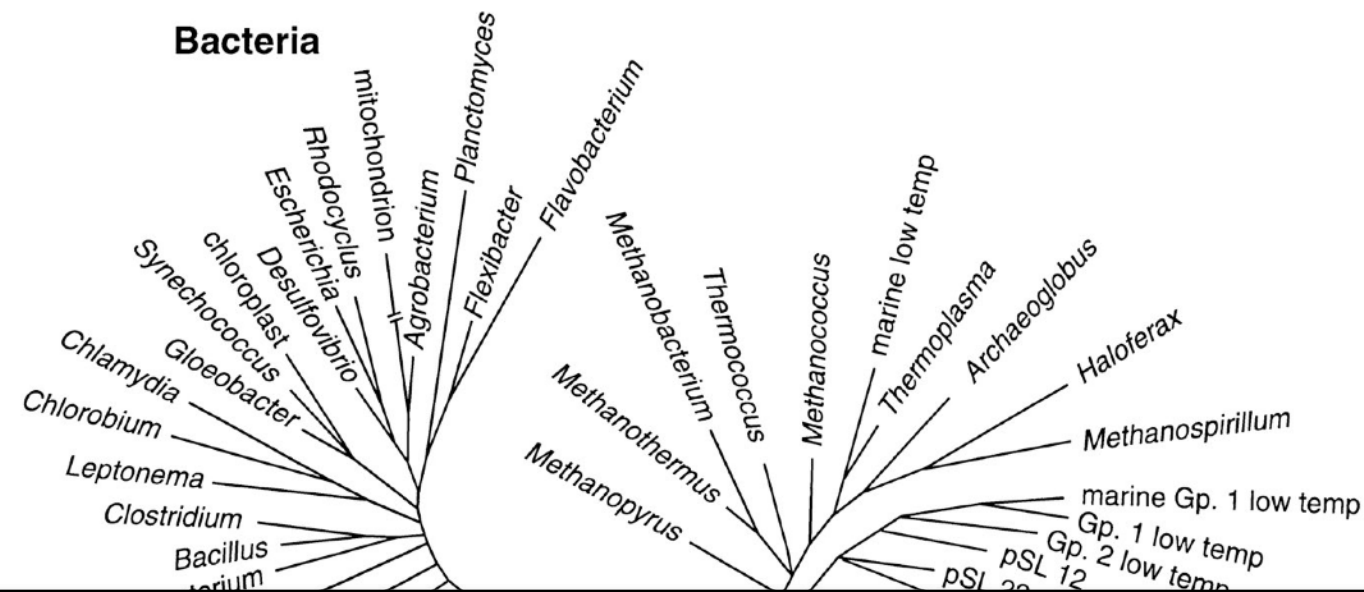
Are OTUs *wrong*? No.

Are ASVs *always* the best unit of analysis? No.

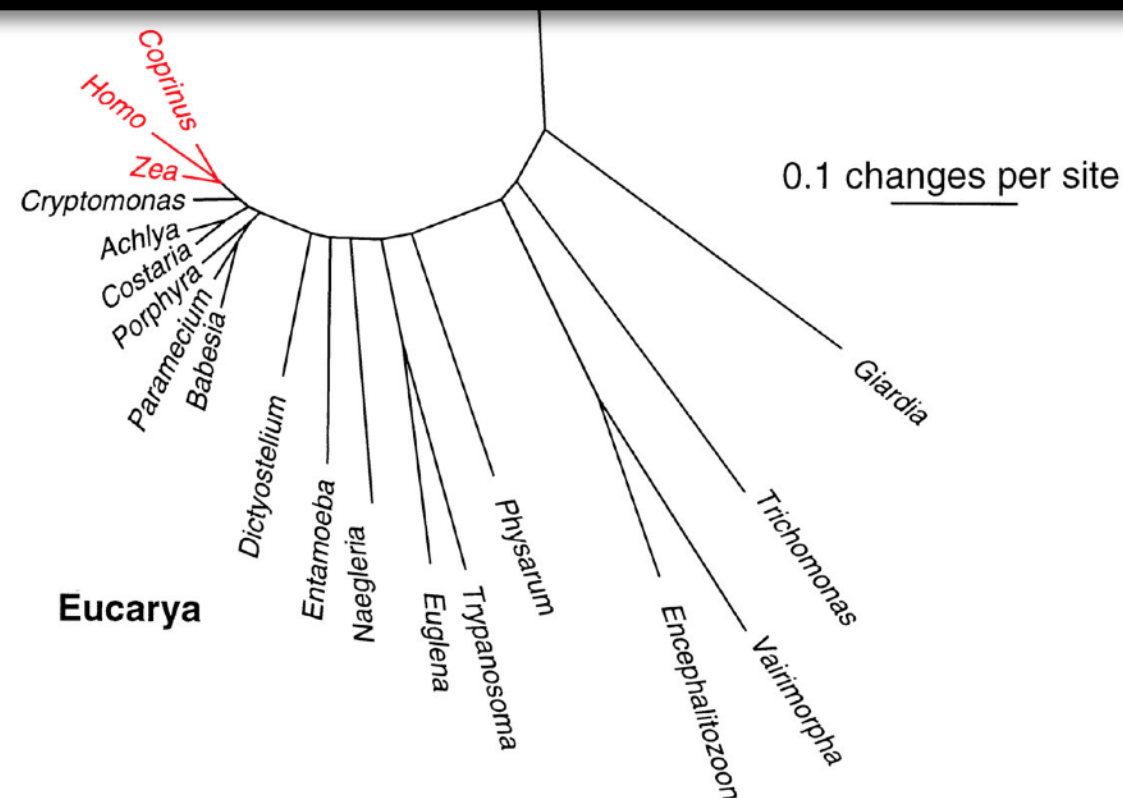
Phylogenetic Scale



Phylogenetic Scale



What “taxonomic unit” best captures the phylogenetic scale at which the function relevant to a scientific question varies?



Recommendations...

- **You should (probably) start with ASVs**
- Consider analysis at *multiple* taxonomic levels
- Reproducible: ASVs, taxonomy. Not: *de novo* OTUs.
- Comprehensive: ASVs, *de novo* OTUs. Not: taxonomy.
- Use domain knowledge on relevant phylogenetic scale
- Deposit and share your ASV table

Can agglomerate up, but can't divide down.

Acknowledgements

Exact Sequence Variants



Susan Holmes



Joey McMurdie



Michael Rosen

...for ITS



Naga Betrapally

Contamination



Nicole Davis

Bias



Michael McLaren

