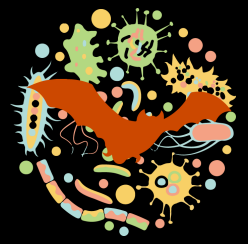
“My Field For Dummies”: Microbiomes and Bats

Melissa R. Ingala, Ph.D.
22 February 2023

@bat_biomes

This webinar will take place in three parts.

Welcome to the Microbiome



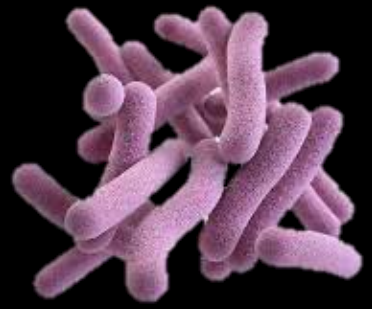
A “How To” Guide for Microbiome Research in Bats

Applying Microbiomes to Ecology & Evolution

Part I: Welcome to the Microbiome

Organisms studied by microbiologists

Bacteria & Archaea



Fungi

Protists

Viruses, virions, prions

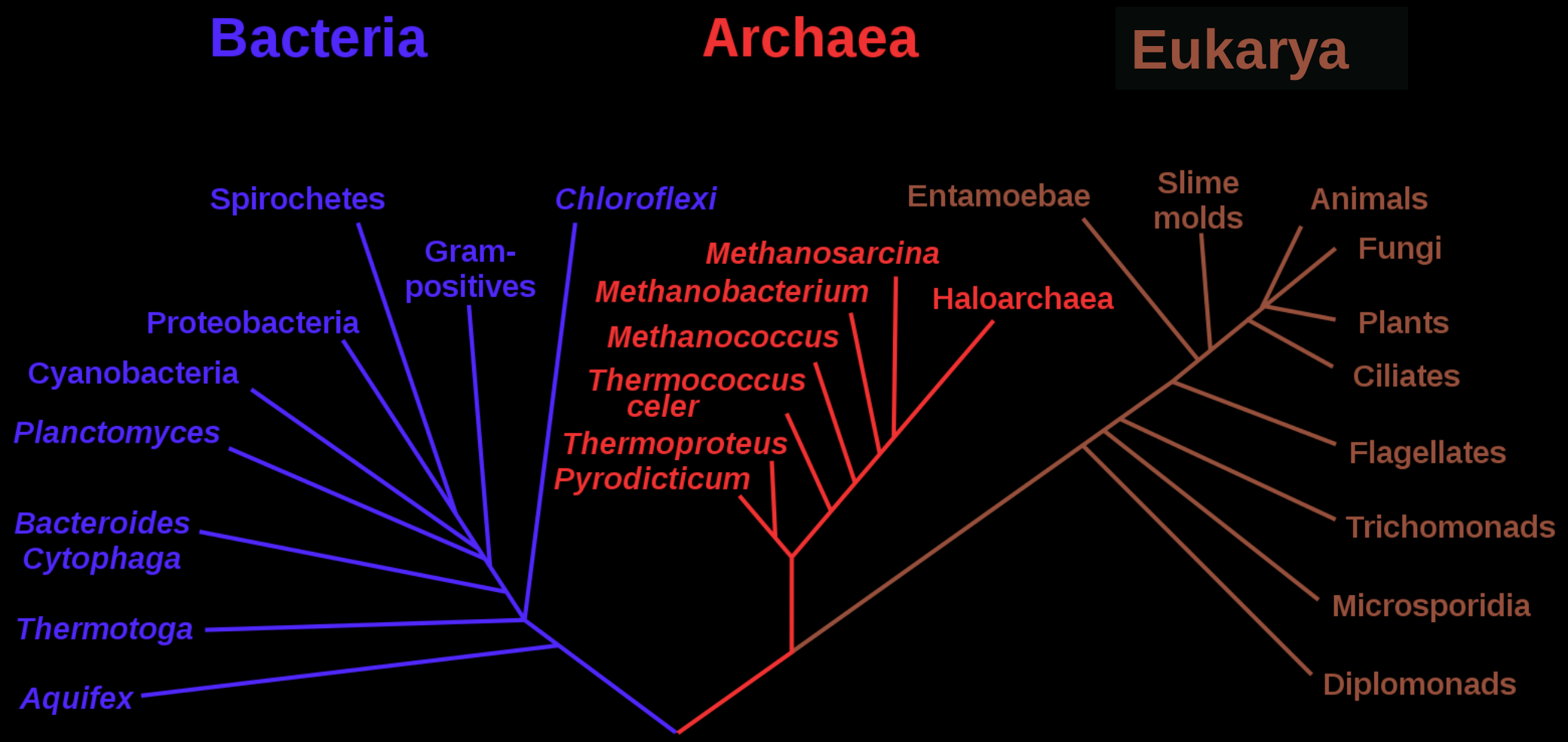
Cellular

DNA

Acellular

DNA/RNA

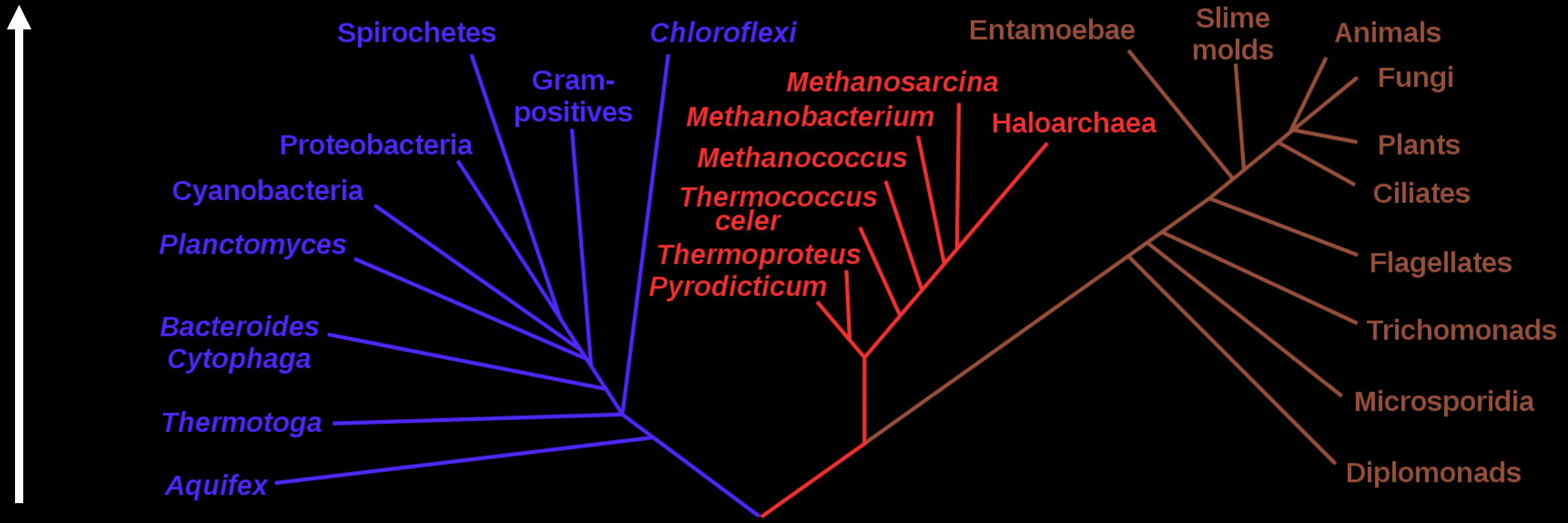
The Universal Tree of Life



Woese et al. (1990)

The Universal Tree of Life

Present species



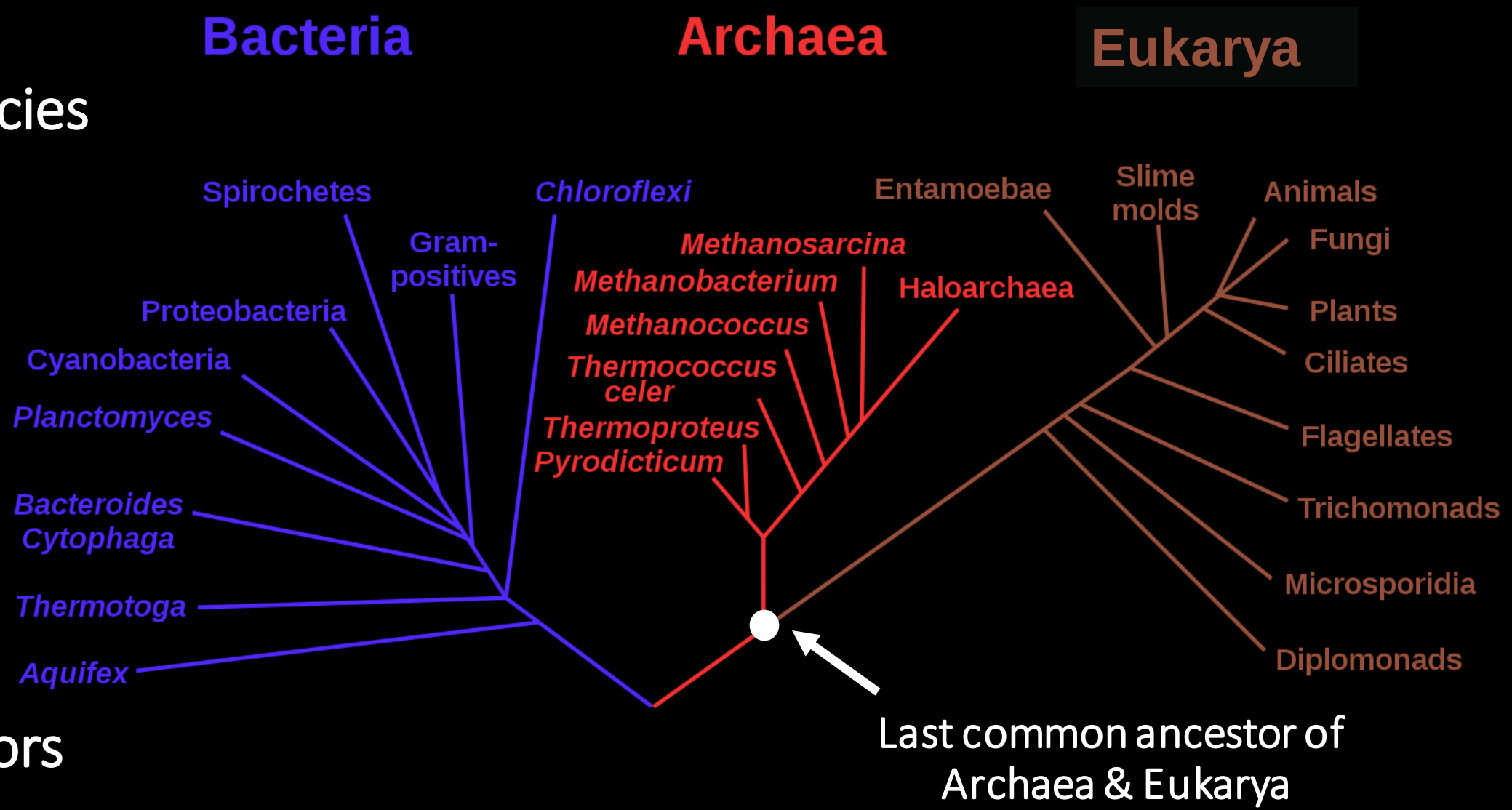
Past ancestors

Woese et al. (1990)

The Universal Tree of Life

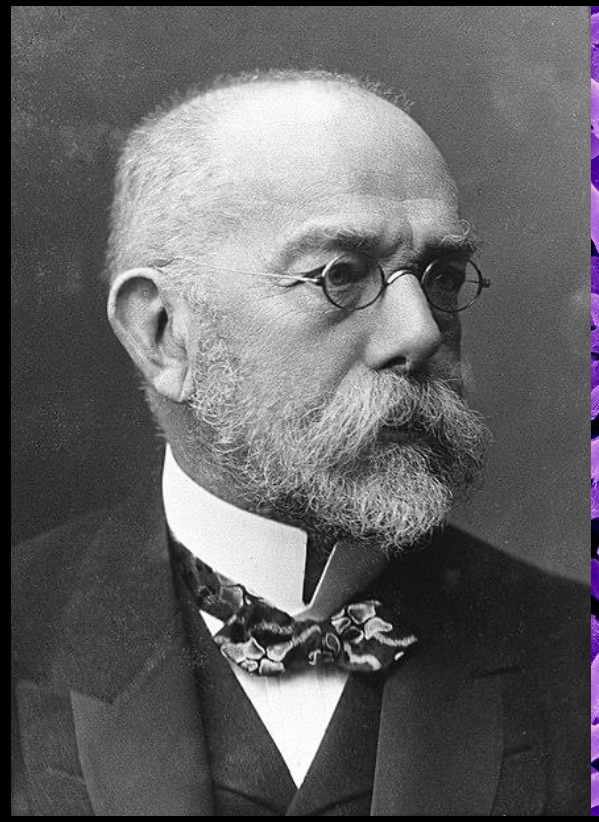
Present species

Past ancestors

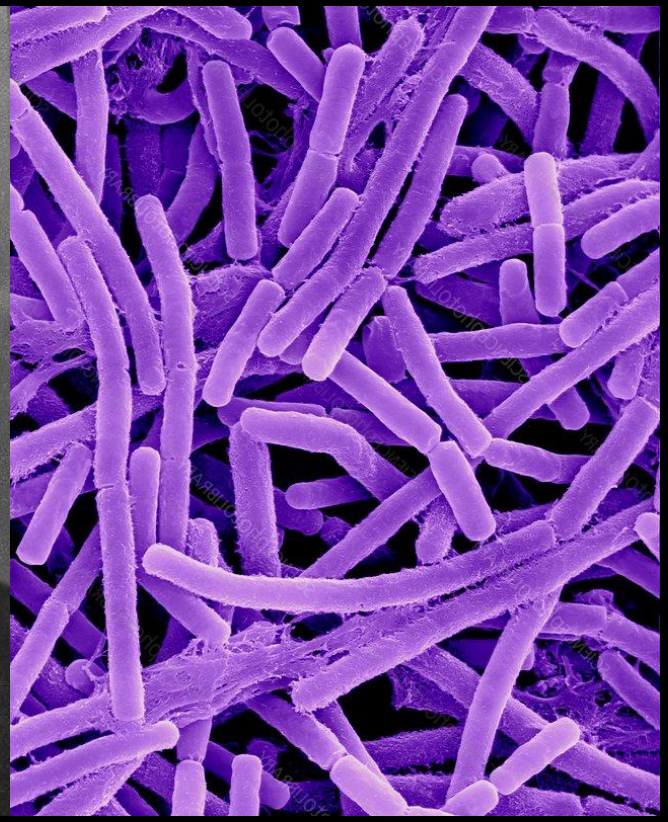


Your bacteria and you

For several centuries, our relationship to bacteria was defined primarily by their capacity to cause disease.



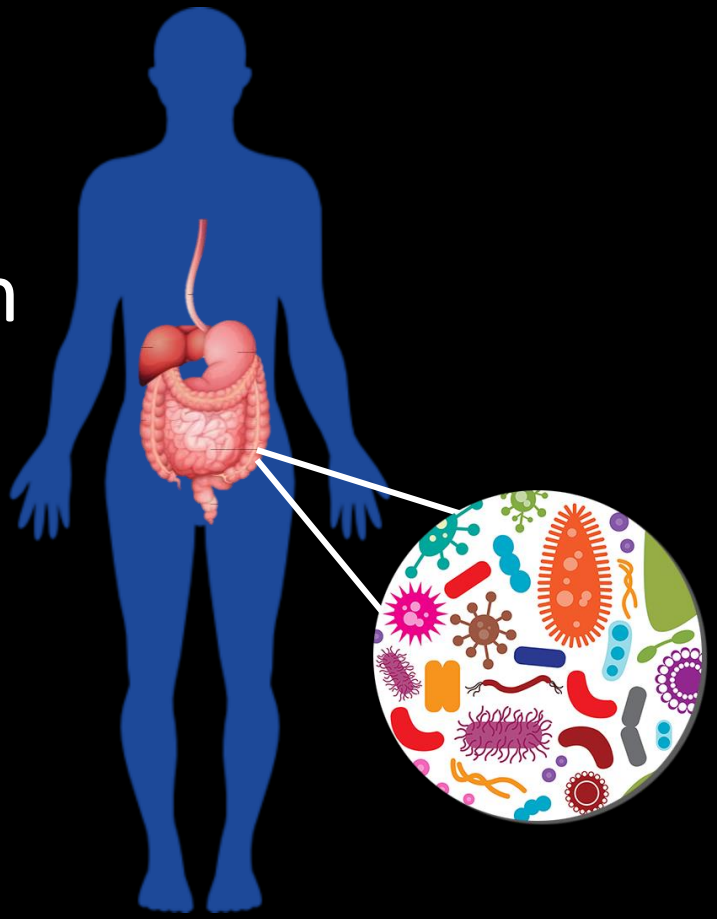
Robert Koch (1843-1910)



Bacillus anthracis, *Causative agent of anthrax disease*

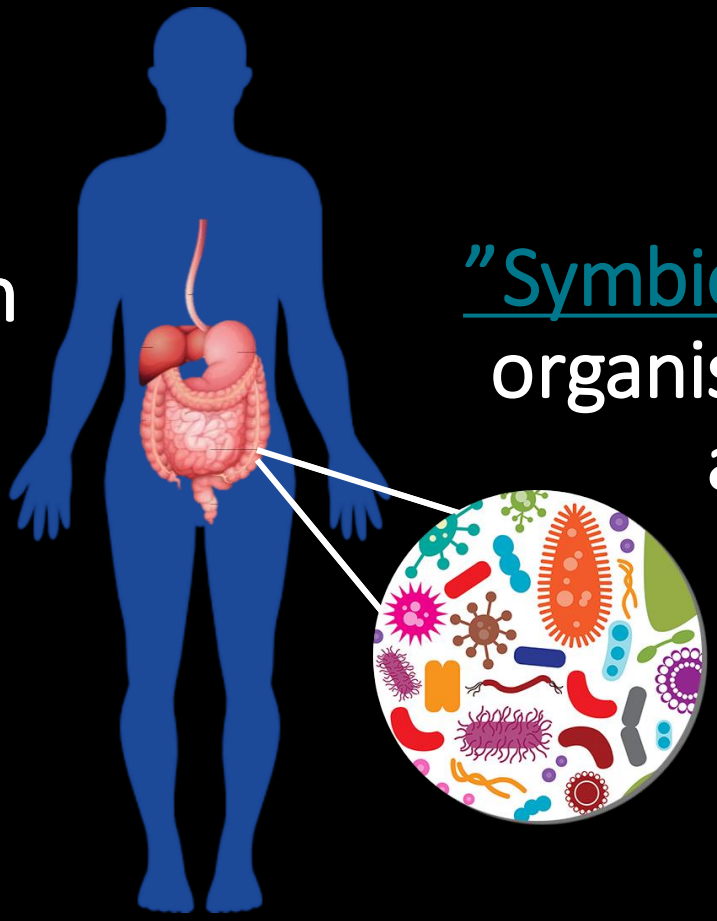
So what is a “microbiome” anyway? Why would we study it?

We now understand that human associated bacterial communities (the microbiome) are inseparable from who we are.



So what is a “microbiome” anyway? Why would we study it?

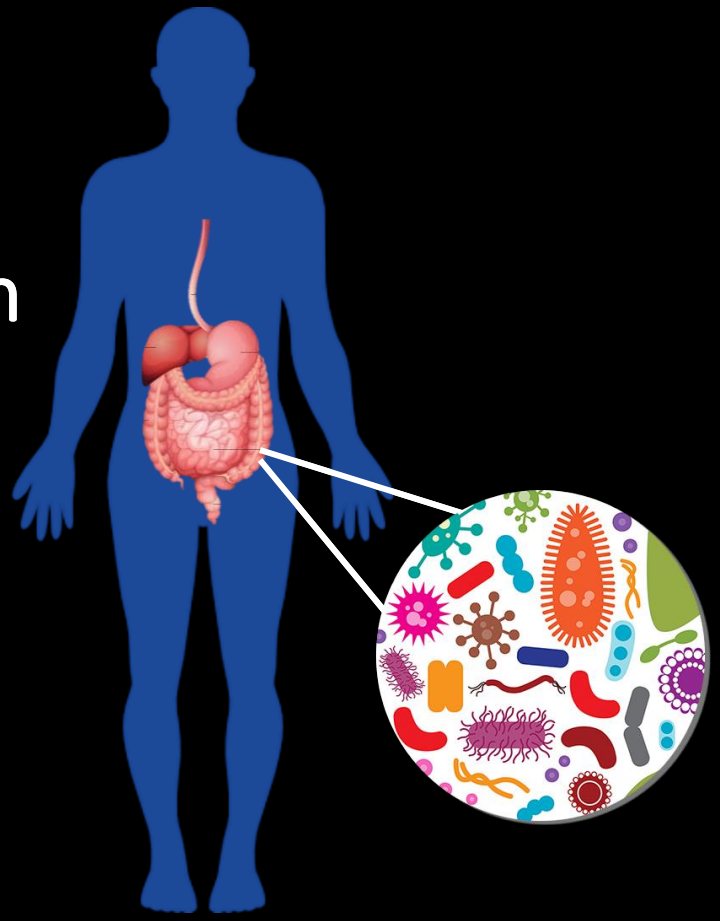
We now understand that human associated bacterial communities (the microbiome) are inseparable from who we are.



”Symbiosis”: two or more organisms living in close association

So what is a “microbiome” anyway? Why would we study it?

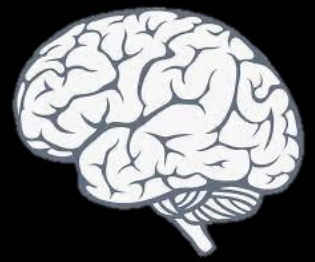
We now understand that human associated bacterial communities (the microbiome) are inseparable from who we are.



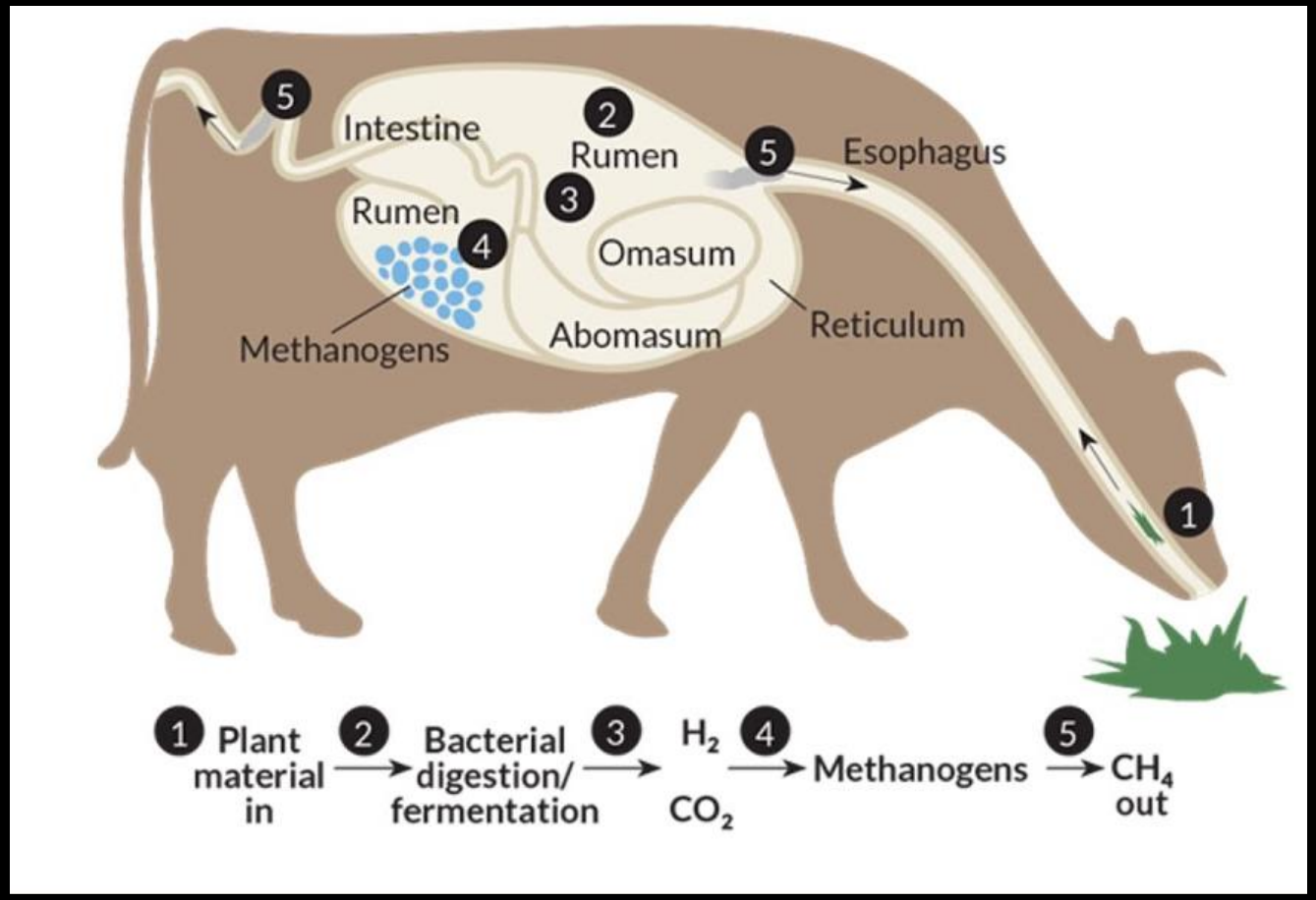
Food metabolism

Immune function

Behavior



Example: the cattle rumen microbiome



Cows rely on the bacteria & archaea in their stomachs to be able to break down the tough cell walls of grass.

What about microbiomes of wildlife?

Part II: A “How To” Guide for Microbiome Research in Bats

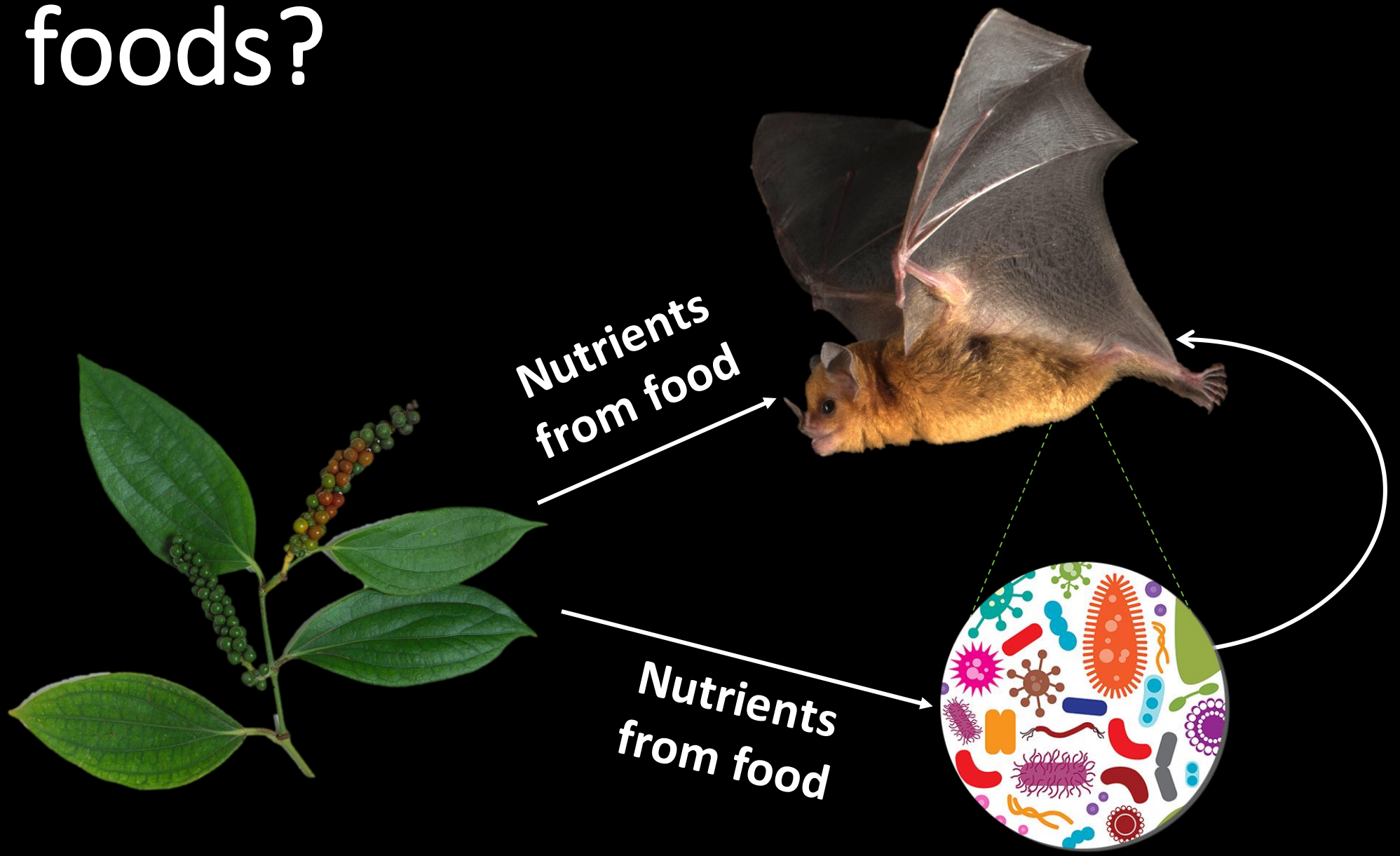
Bats are an ecologically diverse group

How do bats specialize on these foods?

Nutrients from food

Bat's own digestive processes

How do bats specialize on these foods?



Methods for studying host-associated gut microbes

1. Specimen collection

2. Bacterial Genetic Library Preparation & Sequencing

3. Data Analysis

1. Specimen Collection

Feces

Pros:

- Non-invasive
- Quick and easy
- Repeat sampling possible

Intestinal contents

Pros:

- Can do microscopy to examine where bacteria are located
- Can also assess parasites

1. Specimen Collection

Feces

Cons:

- More prone to environmental contamination

Intestinal contents

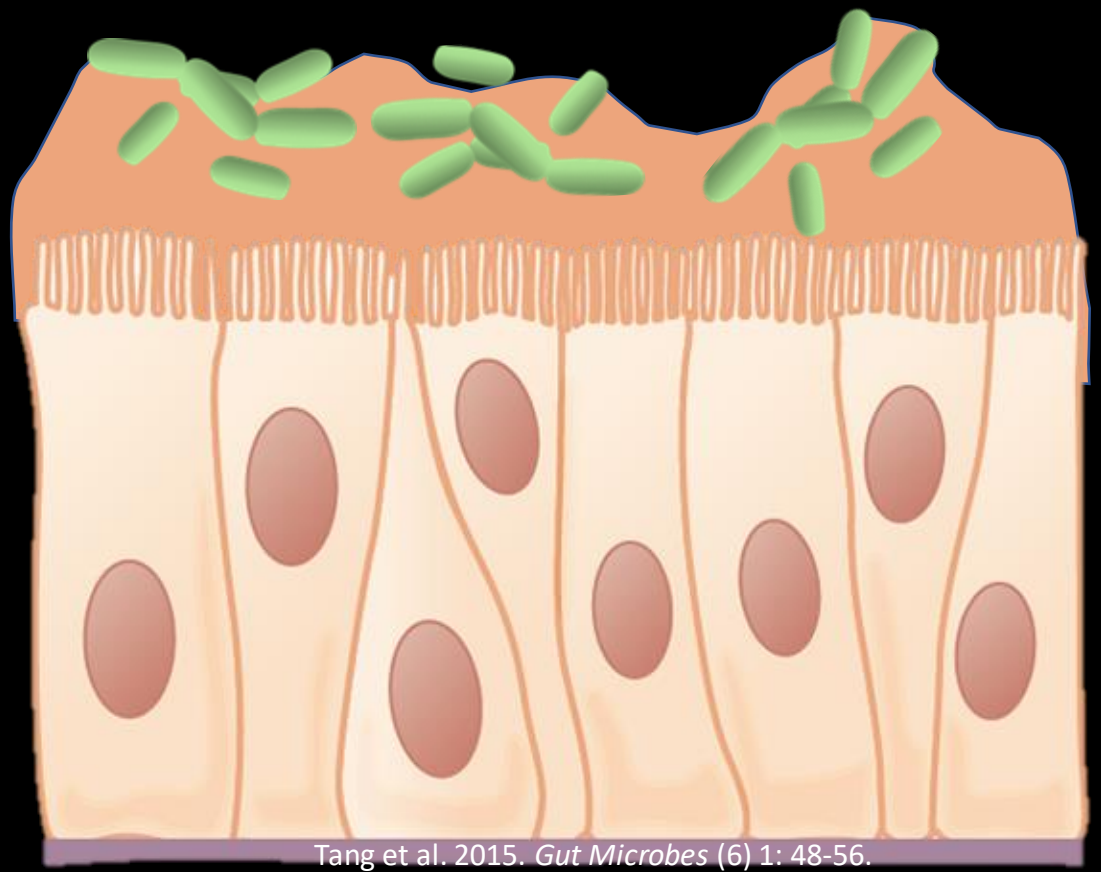
Cons:

- Lethal
- Cannot sample same individual over time

Are these communities the same?

<u>Luminal Bacteria:</u>

1. Bacteria from food items
2. Site of enzymatic digestion
3. Excreted as feces

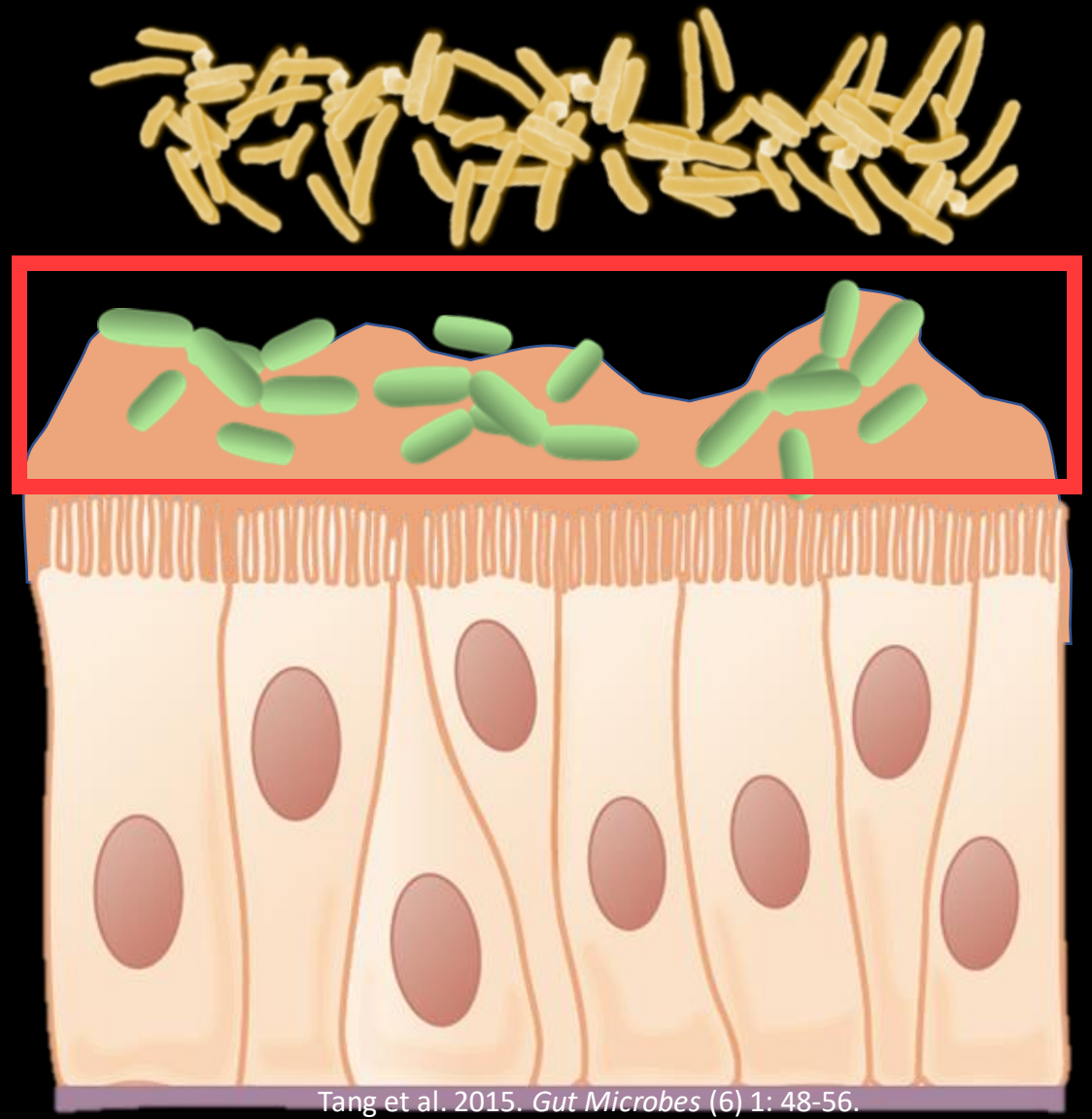


Tang et al. 2015. *Gut Microbes* (6) 1: 48-56.

Are these communities the same?

Intestinal Bacteria:

1. Interface with immune system via epithelial cells
2. Adherent to the mucosa



Tang et al. 2015. *Gut Microbes* (6) 1: 48-56.

2. Sample Preservation Choice

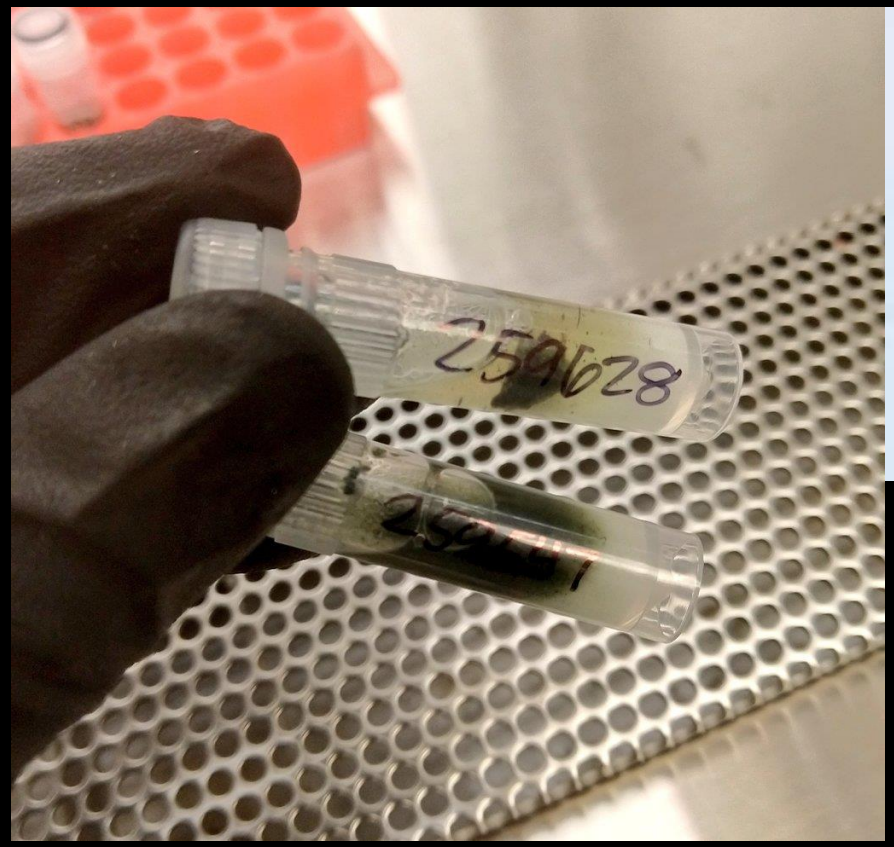
Liquid Nitrogen (No other preservative)

DNA/RNA Shield Whatman FTA cards

95% Ethanol

Gold Standard → **Bog Standard**

2. Isolation of DNA

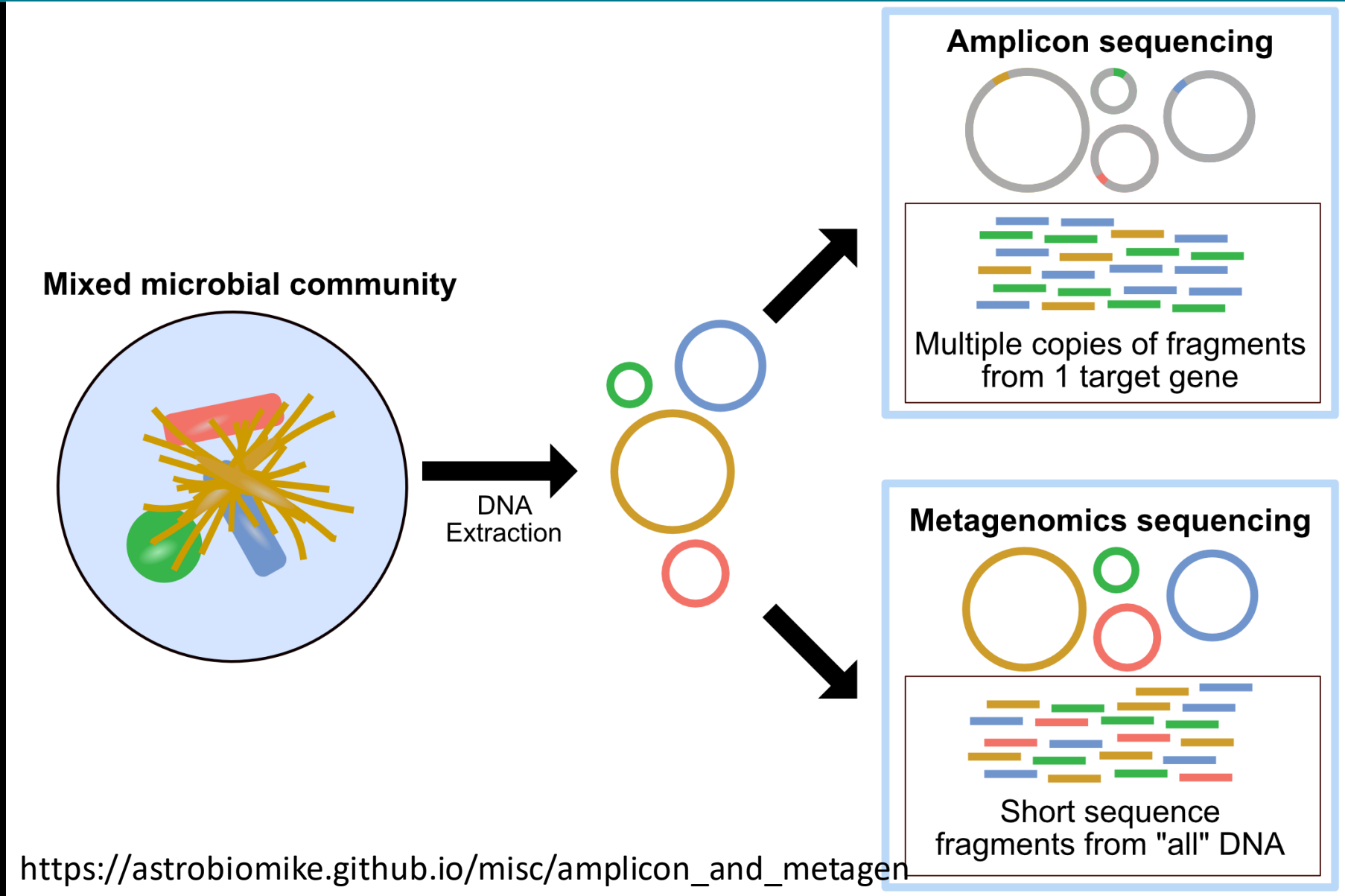


QIAGEN DNeasy Powersoil Pro Kit
$454 USD
Mfr: Germany

ZymoBIOMICS DNA Miniprep Kit
$301 USD
Mfr: USA

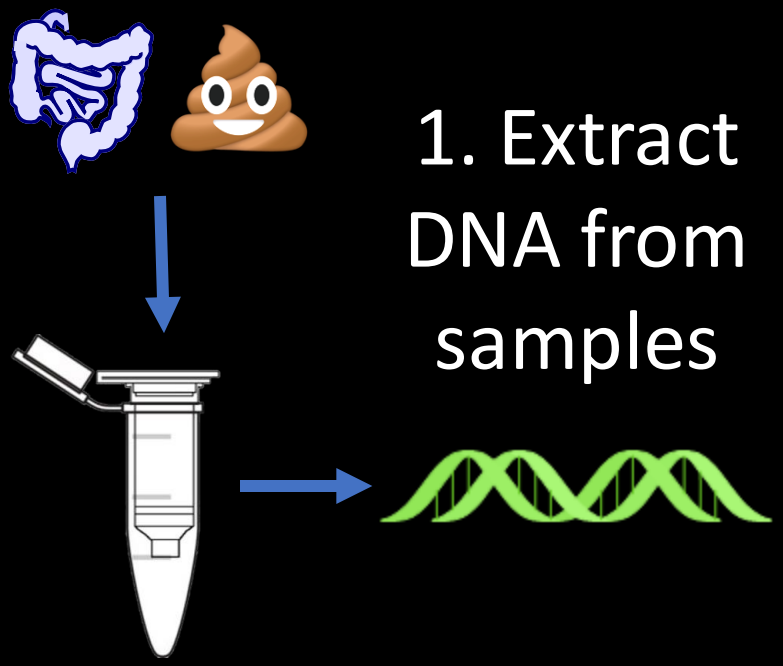
*** Any kit should have both chemical/mechanical lysis step AND humic acid precipitation step ***

2. Choice of Sequencing Technology



https://astrobiomike.github.io/misc/amplicon_and_metagen

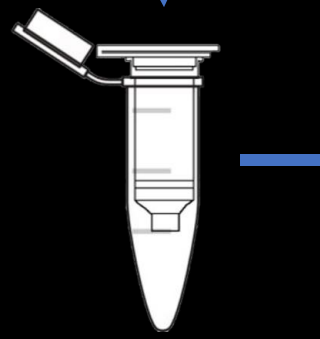
A. Amplicon Sequencing (16S)



1. Extract DNA from samples

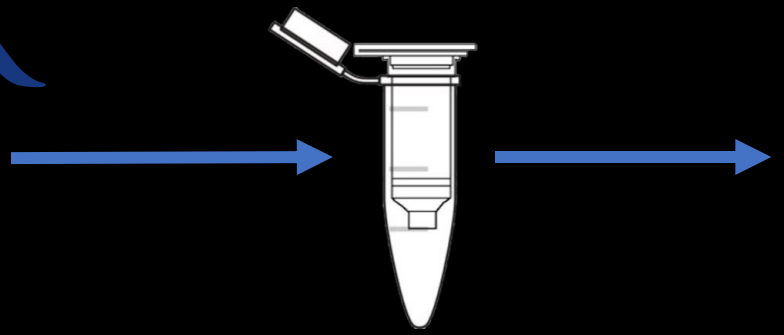
A. Amplicon Sequencing (16S)

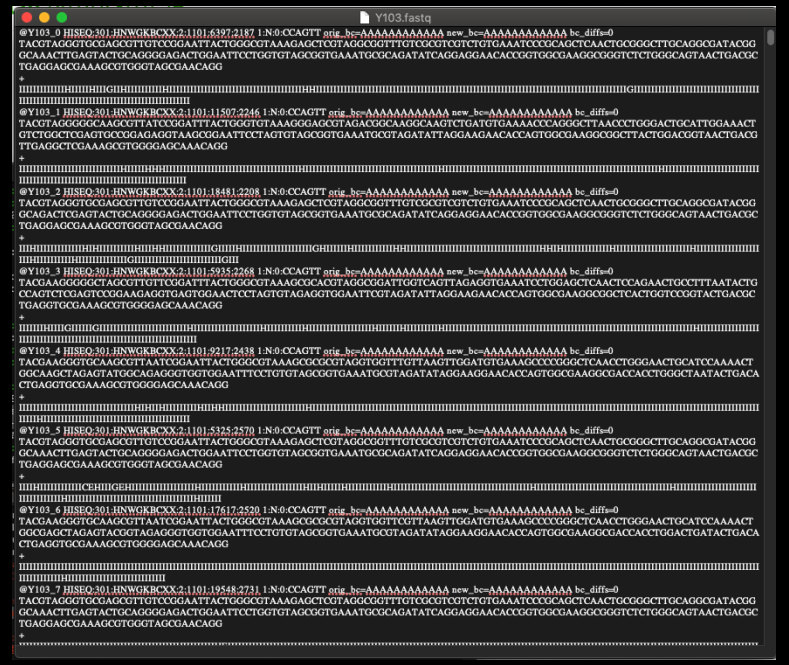
1. Extract DNA from samples



2. Prepare 16S amplicon libraries

A. Amplicon Sequencing (16S)



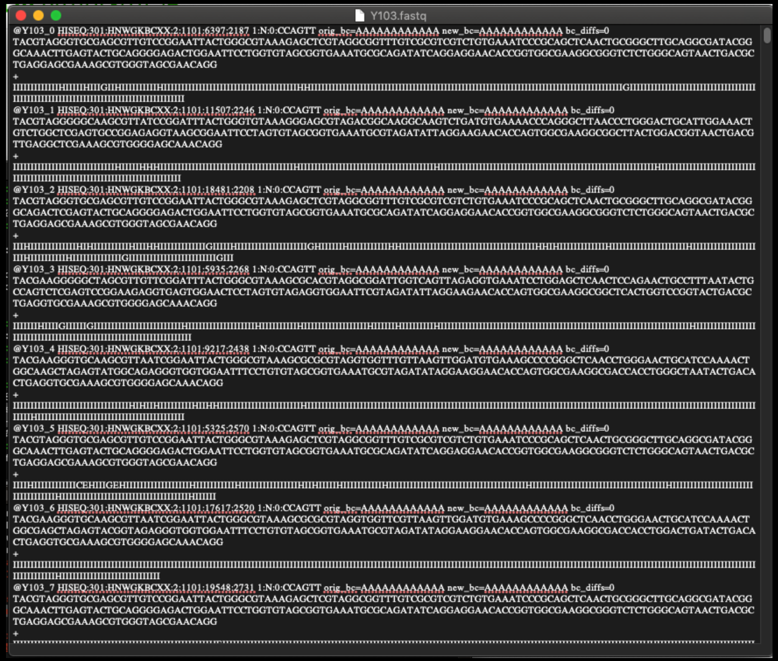


2. Prepare amplicon libraries

3. Normalize & Pool Libraries

4. Sequence

A. Amplicon Sequencing (16S)



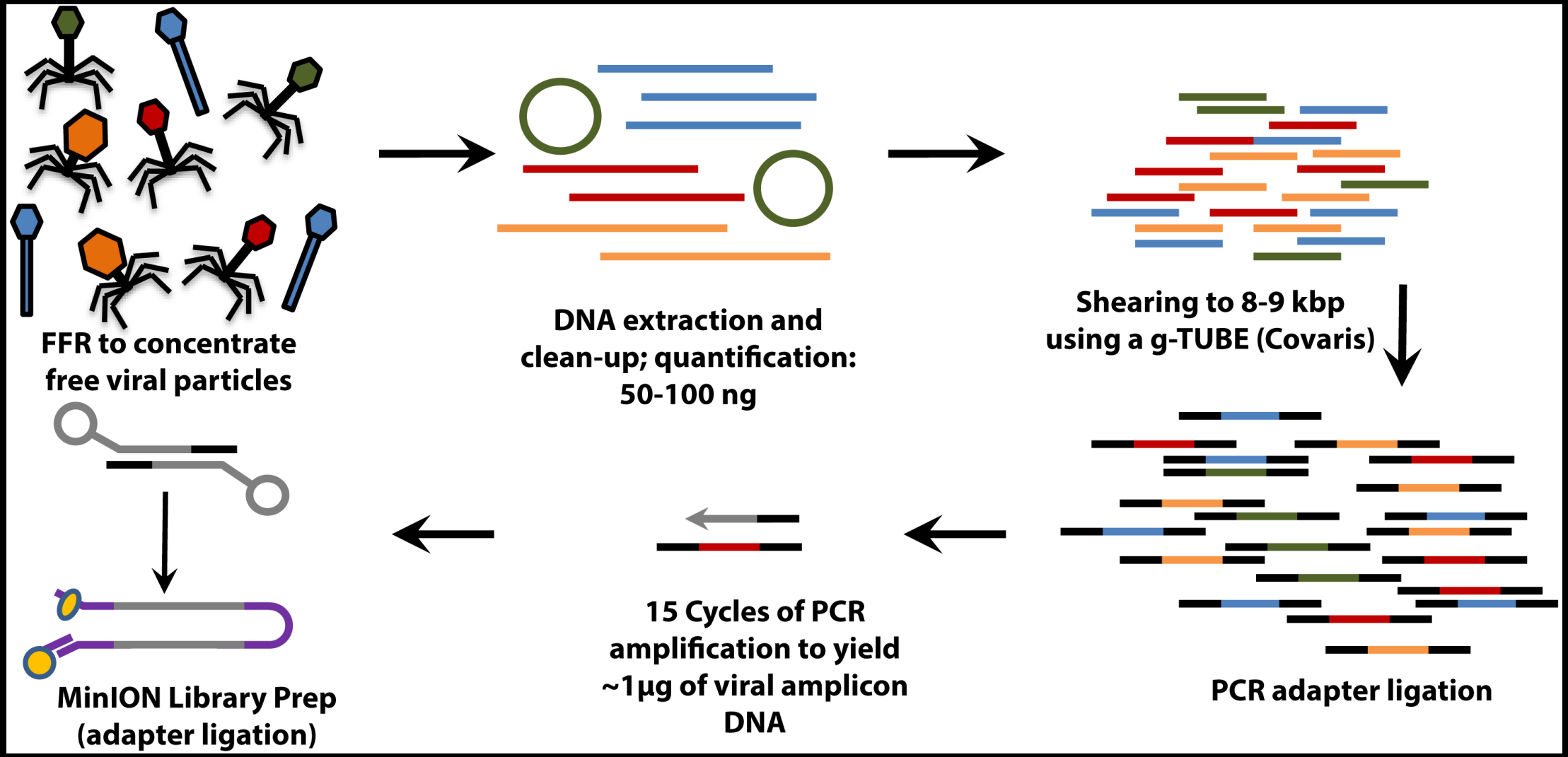
4. Demultiplex, filter, analyze

A. Amplicon Data Analysis Pipelines

Two Most Common Options:

1. QIIME2 (the one I use)
2. Mothur (I have not used this but it is fairly common)

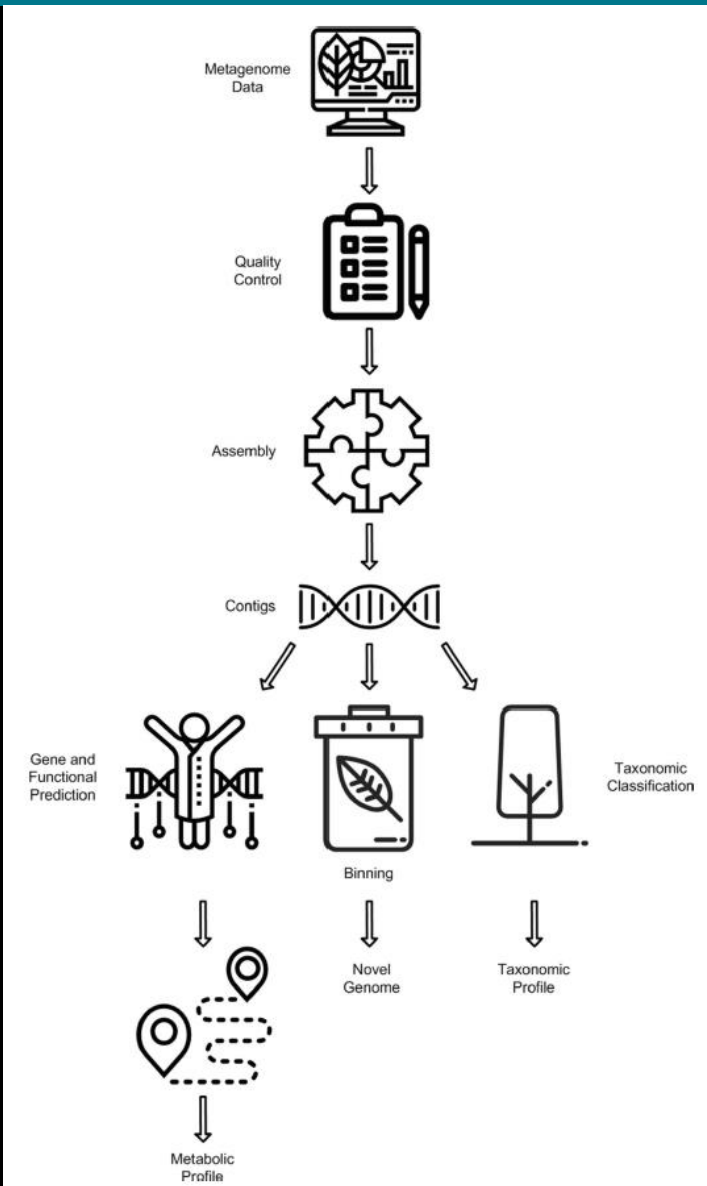
B. Shotgun Sequencing (WGS)



https://www.researchgate.net/publication/333146426_Long-read_viral_metagenomics_captures_abundant_and_microdiverse_viral_populations_and_their_niche-defining_genomic_islands

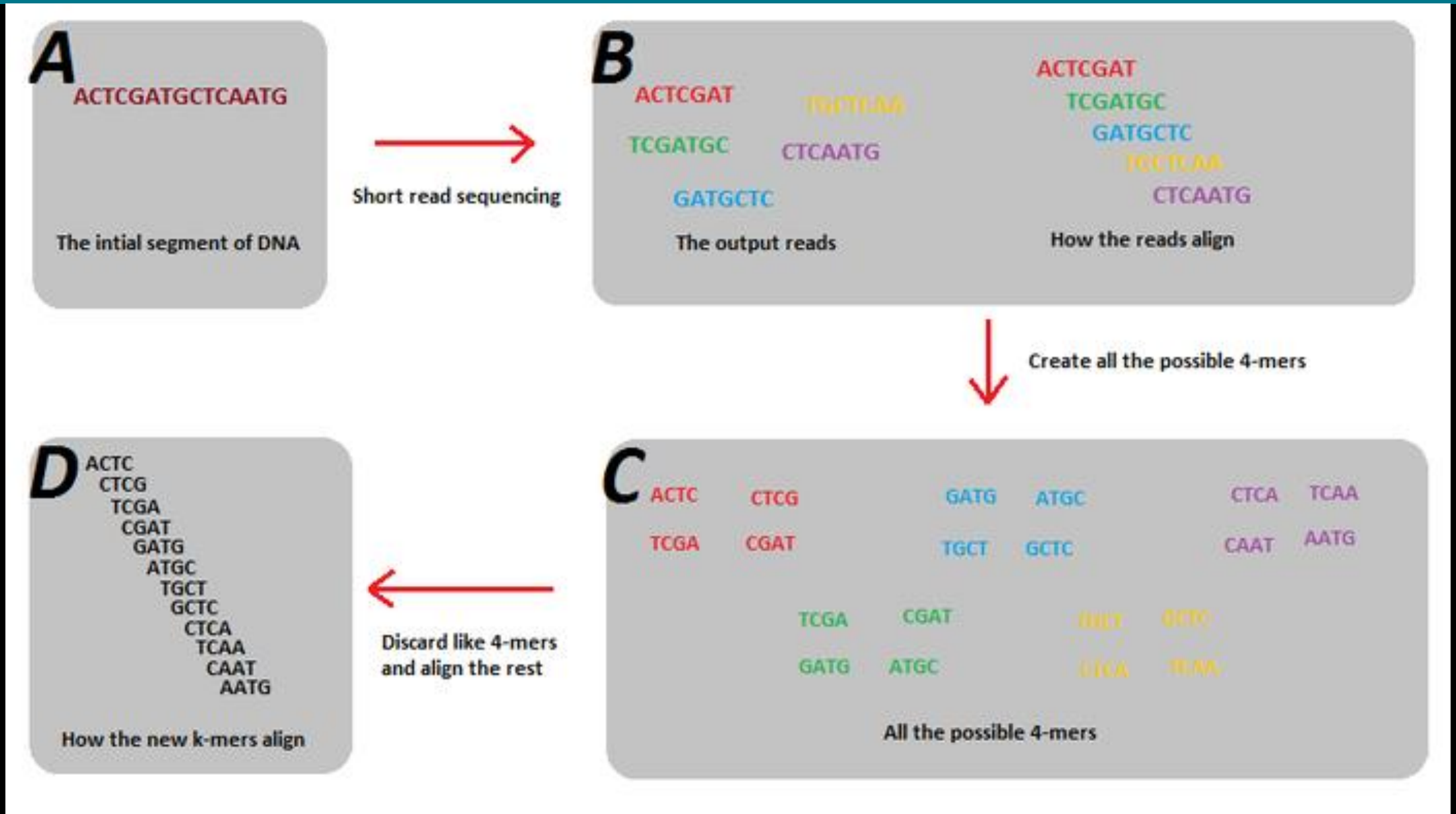
B. Shotgun Analysis Pipelines

- Examples of some available metagenomics assembly/annotation pipelines:
  - MetaSPADES (Nurk et al. 2017)
  - Metawrap (Uritskiy et al. 2018)
  - DRAGEN (Illumina, Inc.)
  - OMARU (Kishikawa et al. 2022)
  - SqueezeMeta (Tamames & Sanchez 2018)

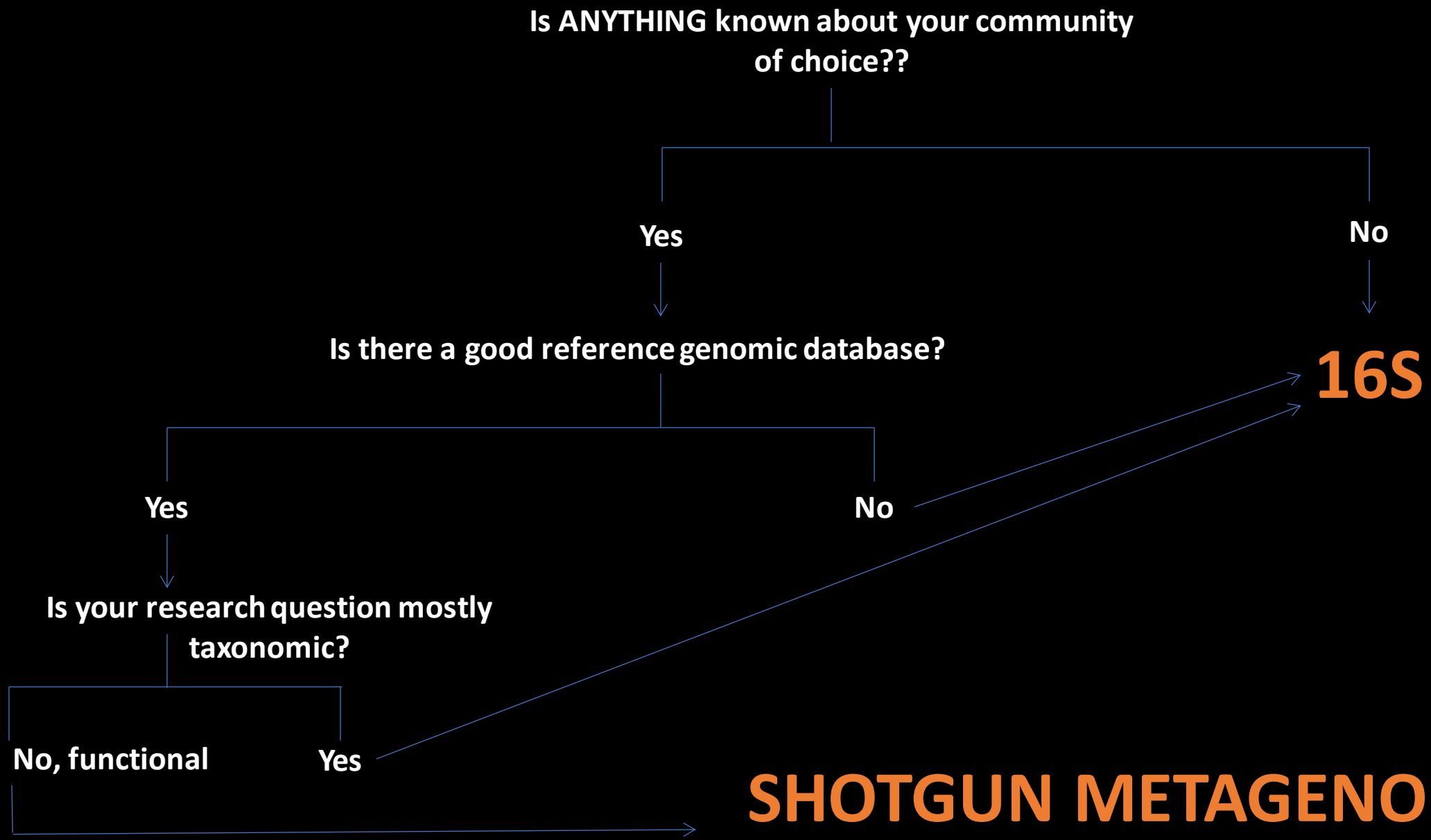


Wajid, B., Anwar, F., Wajid, I., Nisar, H., Meraj, S., Zafar, A., ... & Suchodolski, J. S. (2022). Music of metagenomics—a review of its applications, analysis pipeline, and associated tools. *Functional & integrative genomics*, 1-24.

B. Shotgun Analysis Pipelines



Which Approach should I Use??

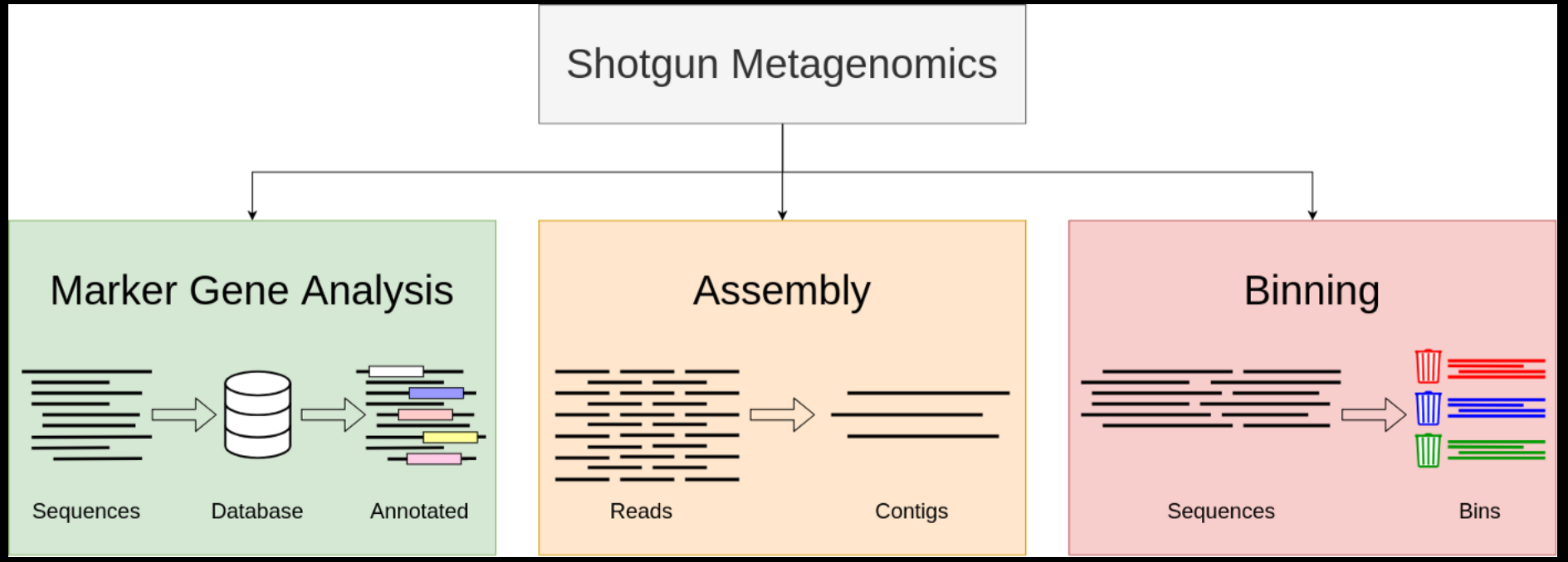


Considerations for shotgun metagenomics

1. **Sequencing coverage / depth**
    1. no easy way to estimate read depth required
    2. If host involved, need to consider accounting for all the reads that will be lost to sequencing host DNA (this can be a lot!)
    3. If you need to perform de novo assembly, best strategy for determining the coverage you need is to just go ahead and sequence a sample on one paired end Illumina HiSeq lane
2. **Read length**
    1. *de novo* assembly– start with a read length of 1 x 150 or 2 x 150 bp
    2. Only interested in measuring abundances– 1 x 90 bp or 1 x 100 bp read length should be sufficient

B. Shotgun Analysis Steps

A. Quality Control
B. Assembly
C. Binning



B. Shotgun Read Quality Control

1. Trim Primers
2. Trim Adaptors
3. Trim low-quality reads (NGS read quality tends to taper at the ends)
4. Remove known contaminants (e.g., host DNA if you study microbiomes)

Table 2 *Quality control:* The following provides a list of tools employed in quality control

1	MultiQC (Ewels et al. 2016)	2	HTSeq (Anders et al. 2015)	3	RobiNA (Lohse et al. 2012)
4	MapSplice (Wang et al. 2010)	5	Hercules (Firtina et al. 2018)	6	QuickNGS (Crispatzu et al. 2017)
7	SAMStat (Lassmann et al. 2011)	8	NextClip (Leggett et al. 2014)	9	Eoulsan (Jourdren et al. 2012)
10	wapRNA (Zhao et al. 2011)	11	Coockiecutter (Starostina et al. 2015)	12	BBT (Chu et al. 2014)
13	RseqFlow (Wang et al. 2011)	14	aRNApipe (Alonso et al. 2017)	15	NGS-Bits (Schroeder et al. 2017)
16	NGS-pipe (Singer et al. 2018)	17	ST Pipeline (Navarro et al. 2017)	18	AlmostSignifcant (Ward et al. 2016)
19	SOAPnuke (Chen et al. 2018)	20	miARma-Seq (Andrés-León et al. 2016)	21	CoVaCS (Chiara et al. 2018a, b)
22	NGS QC Toolkit (Patel and Jain 2012)	23	RNA-QC-Chain (Zhou et al. 2018)	24	miRPursuit (Chaves et al. 2017)
25	SolexaQA (Cox et al. 2010)	26	mubiomics (Smith et al. 2012)	27	OncoRep (Meißner et al. 2015)
28	S-MART (Zytnicki and Quesneville 2011)	29	ReQON (Cabanski et al. 2012)	30	TRAPLINE (Wolfien et al. 2016)
31	PyroTrimmer (Oh et al. 2012)	32	Rnnotator (Martin et al. 2010)	33	TRUFA (Kornobis et al. 2015)
34	BIGpre (Zhang et al. 2011)	35	QuaCRS (Kroll et al. 2014)	36	HYSYS (Schröder et al. 2016)
37	PRINSEQ (Schmieder and Edwards 2011a, b)	38	Pyrocleaner (Mariette et al. 2011)	39	PANDAseq (Masella et al. 2012)
40	Clustal Omega (Haider et al. 2014)	41	Taxoblast (Simon and Dittami 2017)	42	Nonpareil (Rodriguez-r and Konstantinidis 2014)
43	CorQ (Iyer et al. 2013)	44	FQC (Brown et al. 2017)	45	QA (Ramos et al. 2011)
46	FaQCs (Lo and Chain 2014)	47	AfterQC (Chen et al. 2017)	48	HTQC (Yang et al. 2013)
49	SUGAR (Sato et al. 2014)	50	KAT (Mapleson et al. 2017)	51	FastQ_brew (O'Halloran 2017)
52	QUASR (Watson et al. 2013)	53	NGS-QC Generator (Mendoza-Parra et al. 2016)	54	NGS-eval (May et al. 2015)
55	Mistagging (Esling et al. 2015)	56	Blobology (Kumar et al. 2013)	57	sigQC (Dhawan et al. 2017)
58	StatsDB (Ramirez-Gonzalez et al. 2013)	59	G-CNV (Manconi et al. 2015)	60	tGCP (Lindner et al. 2013)
61	Pheniqs (Galanti et al. 2017)	62	QASDRA (Fotouhi et al. 2018)	63	Fast-GBS (Torkamaneh et al. 2017)
64	ClinQC (Pandey et al. 2016)	65	SeqAssist (Peng et al. 2014)	66	PathoQC (Hong et al. 2014)
67	Zseq (Alkhateeb and Rueda 2017)	68	MuffinInfo (Alic and Blanquer 2016)	69	A-Game (Chiara et al. 2018a, b)
70	MT-Toolbox (Yourstone et al. 2014)	71	StreamingTrim (Bacci et al. 2014)	72	FASTQC (Andrews 2010)
73	ACDC (Lux et al. 2016)	74	SqClean (Xie et al. 2011)	75	PhagePhisher (Hatzopoulos et al. 2016)
76	PhylOligo (Mallet et al. 2017)	77	AFS (Liu et al. 2017)	78	MCSC (Lafond-Lapalme et al. 2016)
79	DeconSeq (Schmieder and Edwards 2011a, b)	80	Kraken (Davis et al. 2013)	81	Conpair (Bergmann et al. 2016)
82	ContEst (Cibulskis et al. 2011)	83	decontam (Davis et al. 2018)	84	SIDR (Fierst and Murdock 2017)
85	MIDAS (Nayfach et al. 2016)	86	Coockiecutter (Starostina et al. 2015)		

Wajid, B., Anwar, F., Wajid, I., Nisar, H., Meraj, S., Zafar, A., ... & Suchodolski, J. S. (2022). Music of metagenomics—a review of its applications, analysis pipeline, and associated tools. *Functional & integrative genomics*, 1-24.

B. Assembly-based Analysis

- Assembly aims to stich together sequencing reads to reconstruct the original genome from which they were derived
- Because metagenomics data often contain tens of millions of reads, classification is typically done using exact matching of short words of length k **(k-mers)** rather than alignment, which would be unacceptably slow
- Example programs: MetaSPAdes, MetaVelvet, MEGAHIT

B. Assembly-based Analysis

_pondered_wea

_ed_weak_and

While_I_pon

_and_weary

Once_upon_a

Ht_dreary_while

Ght_drea

On_a_midnight

Once_upon_a

on_a_midnight

ght_drea

while_I_nodded_nearly_napping **CHALLENGE: ambiguity**

ht_dreary_while

while_I_pon **CHALLENGE: insufficient coverage**

_pondered_wea

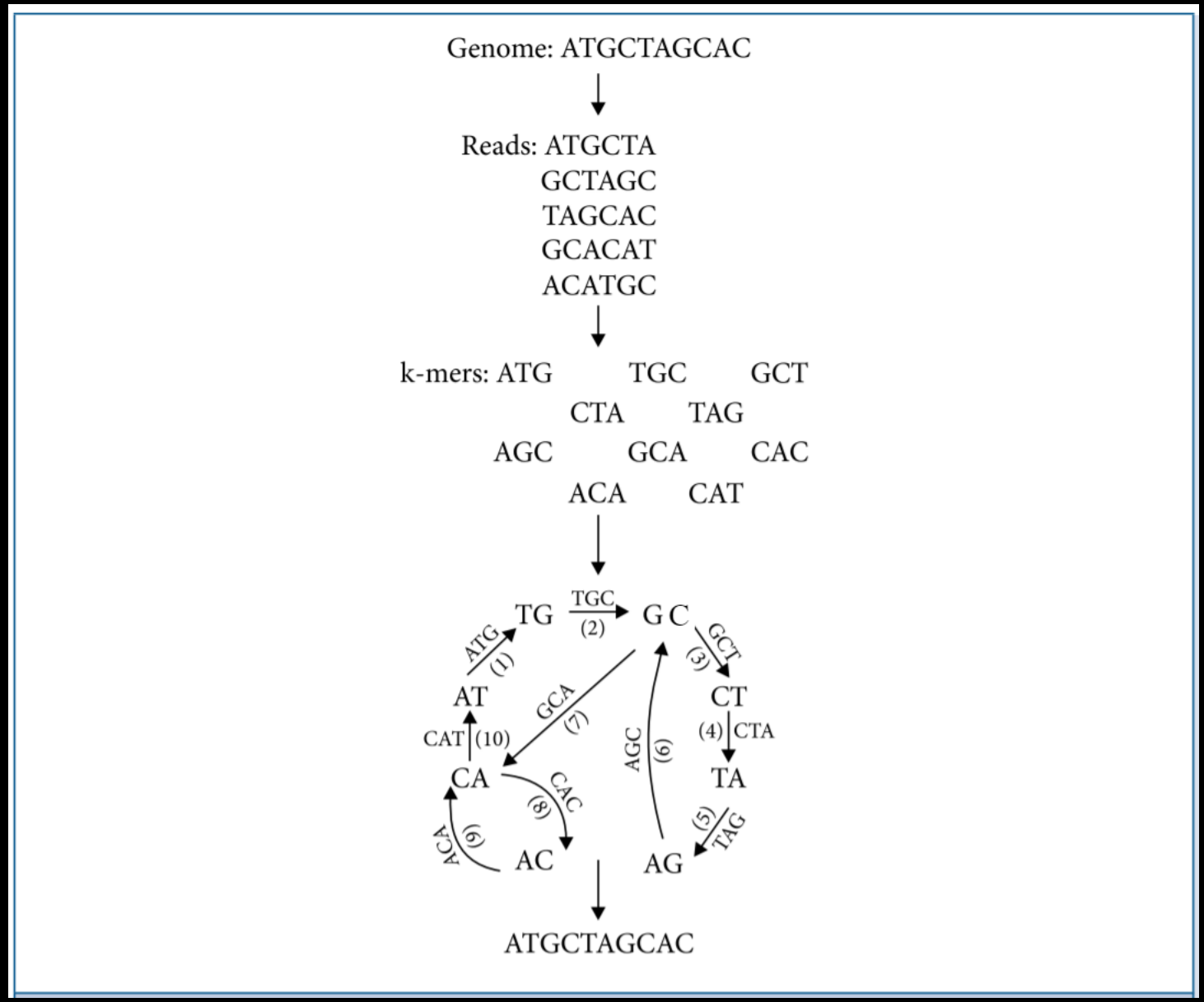
_ed_weak_and

_and_weary

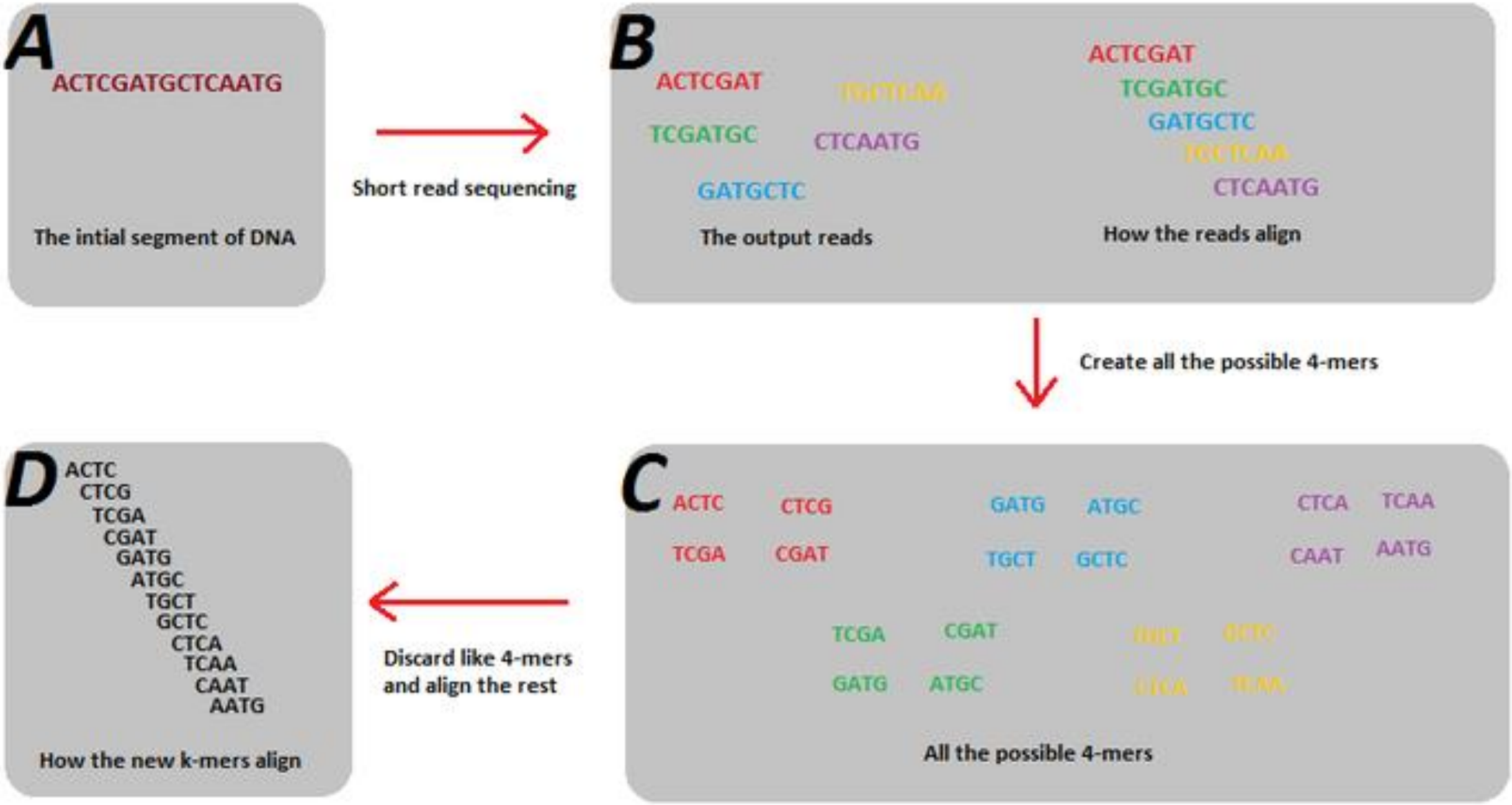
Reconstructed sequence: *Once upon a midnight dreary while I pondered weak and weary*

B. Assembly-based Analysis

- Most assemblies use De Brujn graphs
- Might be able to get full genomes if community is simple
- More often you get contigs (chunks of genomes), longer is better

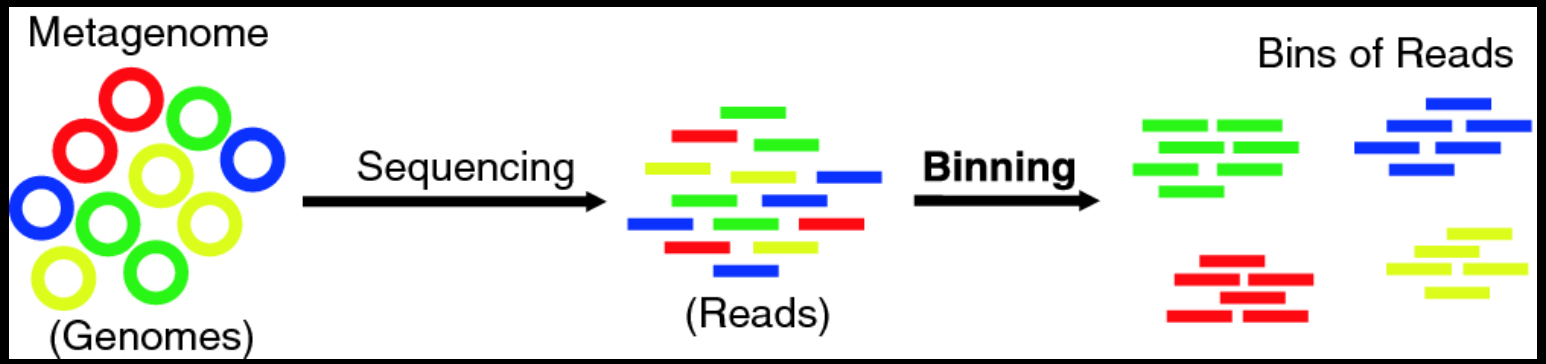


https://towardsdatascience.com/genome-assembly-using-de-bruijn-graphs-69570efcc270?gi=323ee48a7438


A
ACTCGATGCTCAATG
The intial segment of DNA
Short read sequencing
B
ACTCGAT
TGCTCAA
TCGATGC
CTCAATG
GATGCTC
The output reads
ACTCGAT
TCGATGC
GATGCTC
TGCTCAA
CTCAATG
How the reads align
Create all the possible 4-mers
C
ACTC
CTCG
GATG
ATGC
CTCA
TCAA
TCGA
CGAT
TGCT
GCTC
CAAT
AATG
TCGA
CGAT
GATG
ATGC
All the possible 4-mers
Discard like 4-mers
and align the rest
D
ACTC
CTCG
TCGA
CGAT
GATG
ATGC
TGCT
GCTC
CTCA
TCAA
CAAT
AATG
How the new k-mers align

B. Binning Shotgun Reads

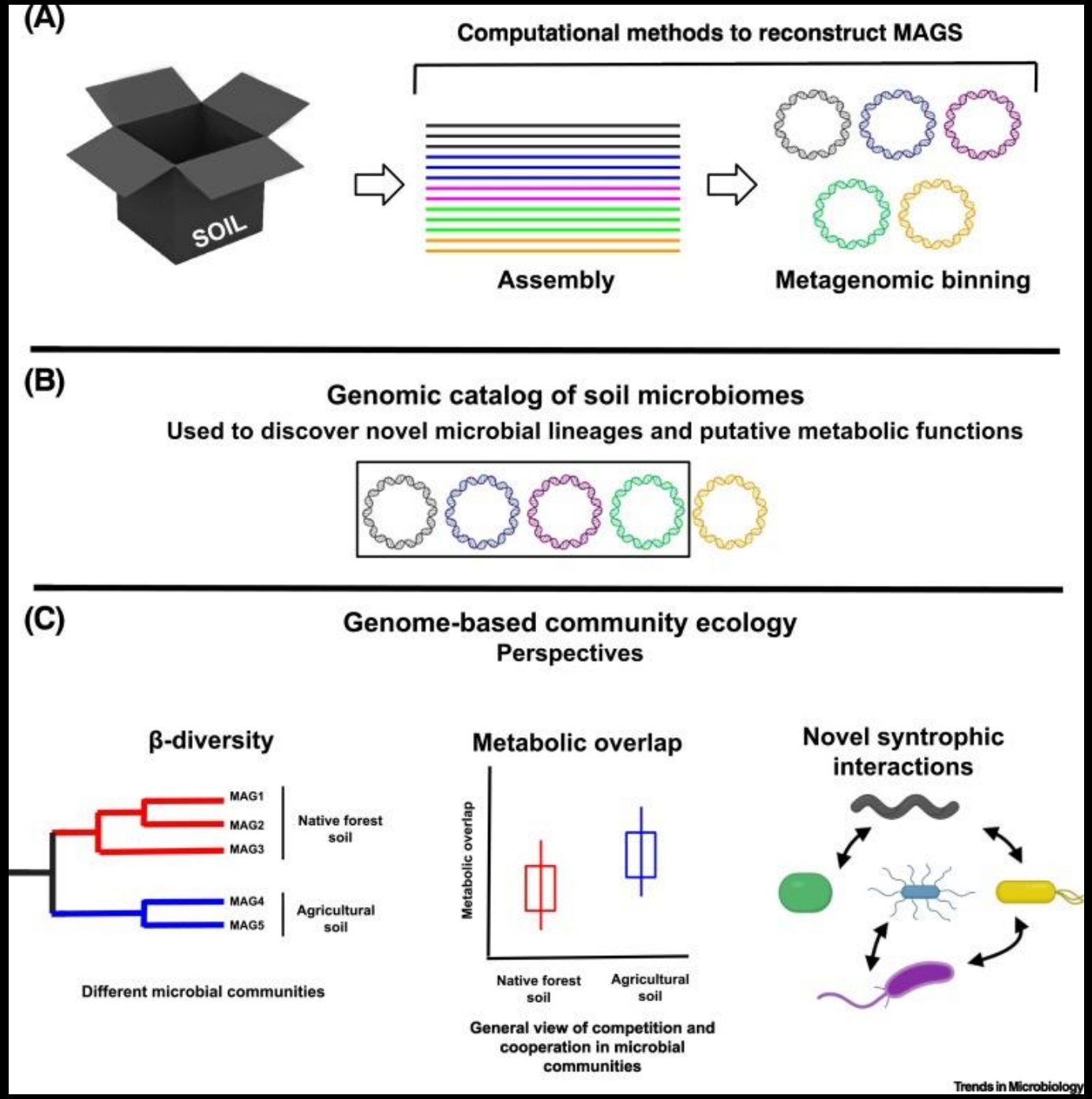
- One of the biggest challenges to assembly is the generation of chimeras, where two sequences from different genomes or parts of the genome are incorrectly merged due to similar sequence composition.
- This is often mitigated by performing a **binning** step, assigning each metagenomic sequence to a taxonomic group and then assembling each bin independently. This helps reduce data complexity and the chance of chimeras.



B. Gene Ontology and Functional Predictions

- Once assembled, genes can be predicted and functionally annotated:

1. De novo gene prediction
2. Protein family classification
    - Sequence or hidden Markov model (HMM) databases
        - SEED
        - KEGG
        - MetaCyc
        - EggNOG
3. Fragment recruitment (binning)



Lemos, L. N., Mendes, L. W., Baldrian, P., & Pylro, V. S. (2021). Genome-resolved metagenomics is essential for unlocking the microbial black box of the soil. *Trends in Microbiology*, *29*(4), 279-282.

B. Challenges of Metagenome Assembly

1. Uneven abundances = highly non-uniform read coverage across different genomes
2. Challenging to correctly place repetitive DNA segments in a genome in mixed organism samples
3. Typical issues with defining a microbial “species” (don’t ask me, I don’t know)
4. Most microbial genomic databases are heavily biased to medically relevant bacteria (not natural history)

B. Consider Computing Power Available to You!

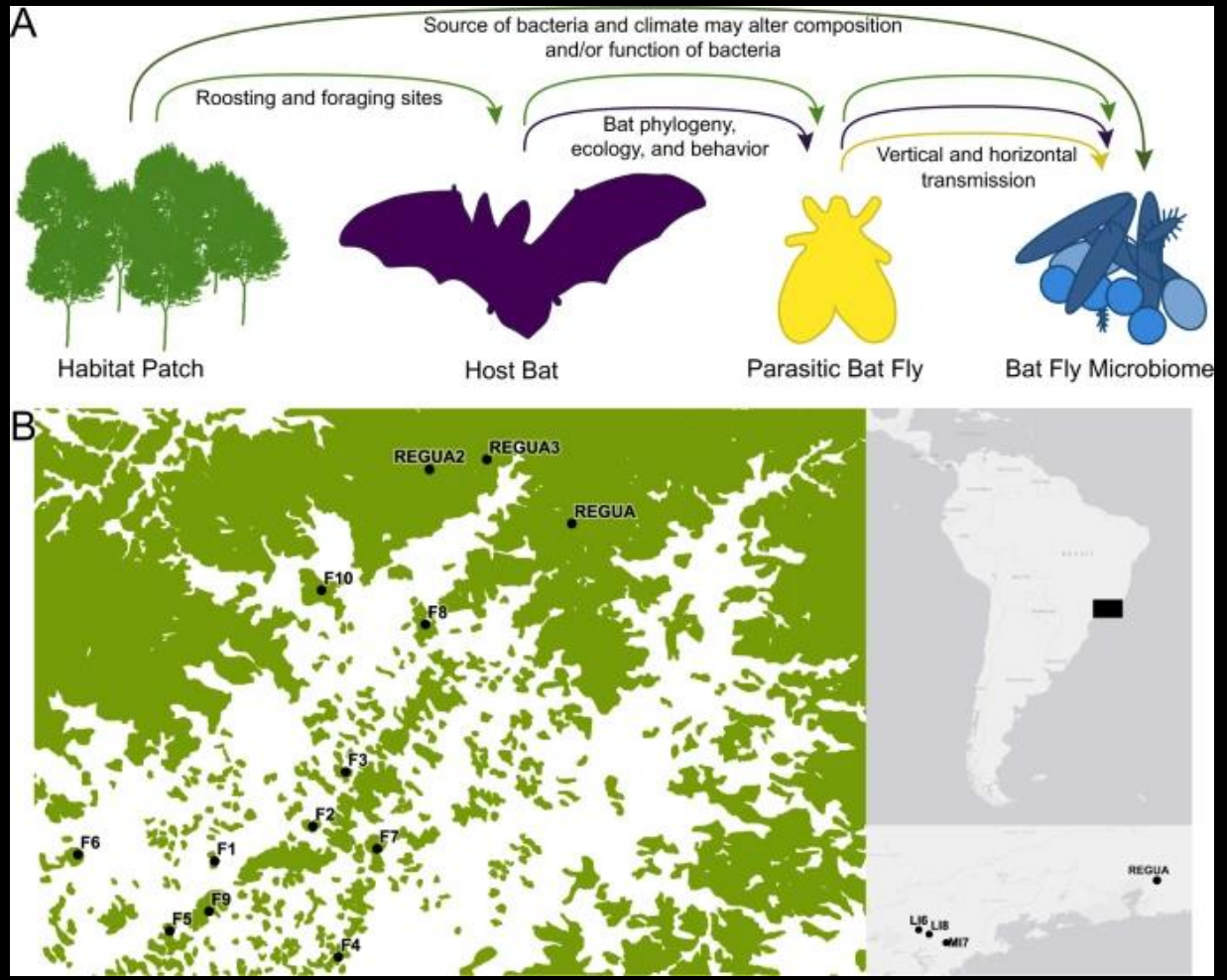
- A standard laptop lacks the CPU power to assemble even a small metagenome, but you can do small 16S studies on local machines
- Raw data files are large (mine were > 5 G)
- There are streamlined pipelines that are less computationally intensive, but you will still require HPC access
  - E.g., MetaWrap
  - 16S data is memory intensive but not I/O intensive. Shotgun data are both.
  - Reference genome databases take up a LOT of memory alone

Part III: Applications to Ecology & Evolutionary Biology

How can we use microbiomes to understand host ecology?

- Effects of habitat loss cascade through multiple scales of biological organization (bat hosts → bat flies → bat fly microbes (Speer et al. 2022)

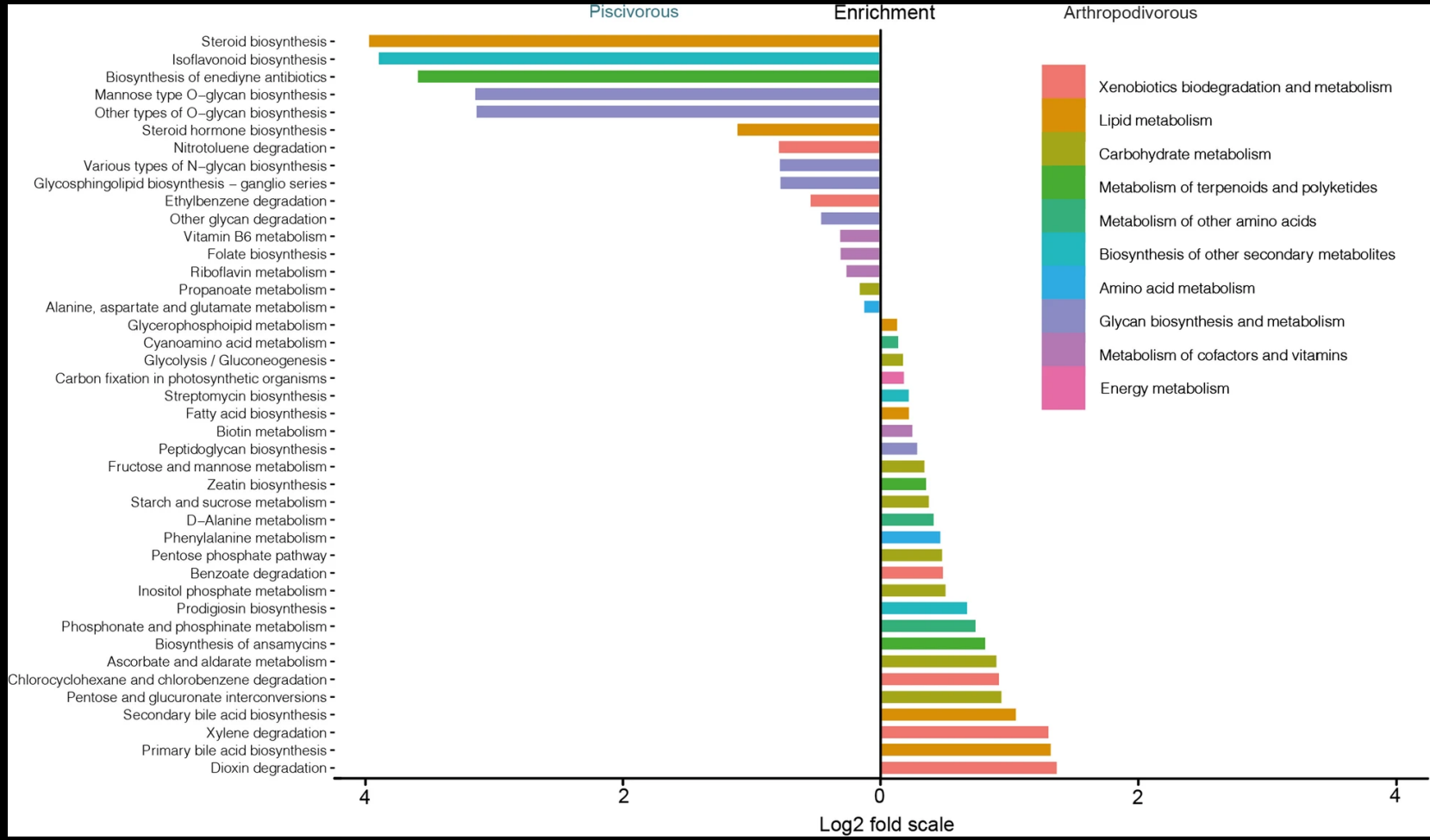
- Implications for diseases vectored by blood-feeding arthropods



Speer, K.A., Teixeira, T.S.M., Brown, A.M. *et al.* Cascading effects of habitat loss on ectoparasite-associated bacterial microbiomes. *ISME COMMUN.* **2**, 67 (2022). https://doi.org/10.1038/s43705-022-00153-0

How can we use microbiomes to understand host ecology?

- We can assess the role of microbes in facilitating animal dietary ecology
- Fishing bats have gut microbes that help break down OMEGA fatty acids found in fish
- Implications for host evolution & adaptation

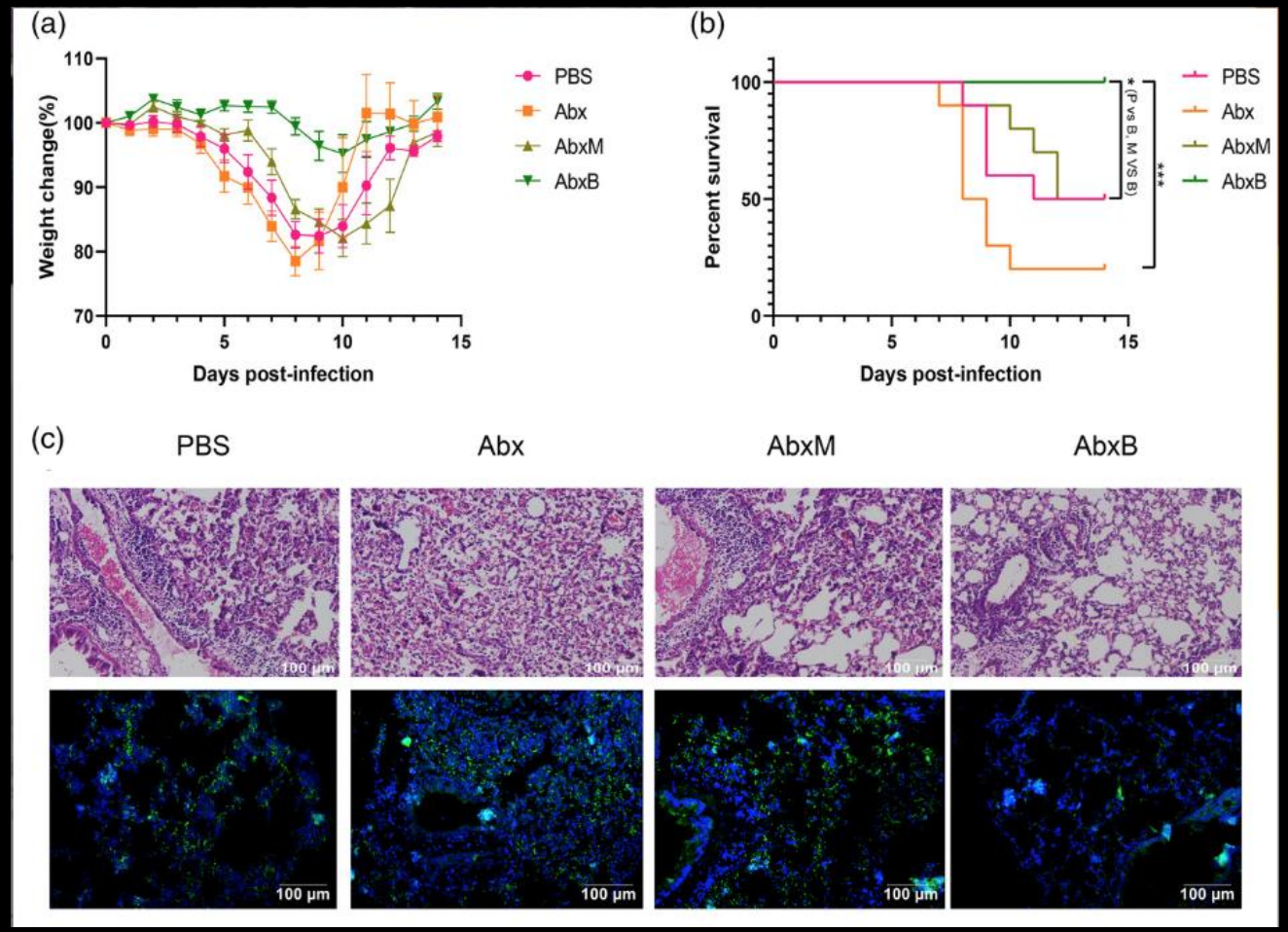


Aizpurua, O., Nyholm, L., Morris, E. *et al.* The role of the gut microbiota in the dietary niche expansion of fishing bats. *Anim Microbiome* **3**, 76 (2021). https://doi.org/10.1186/s42523-021-00137-w

How can we use microbiomes to understand host immunity?

- Inoculation of mice with *Hipposideros armiger* microbiota decreased H1N1 infection and abrogates lung damage

- Paired metabolomics showed bat-inoculated mice produce more flavonoids and isoflavones (regulate immunity via cytokine secretion)



Liu, B., Chen, X., Zhou, L., Li, J., Wang, D., Yang, W., ... & Jiang, T. (2022). The gut microbiota of bats confers tolerance to influenza virus (H1N1) infection in mice. *Transboundary and Emerging Diseases*, *69*(5), e1469-e1487.

In summary

- Microbiome research in wildlife is a growing field; come join us!

- *Bats in Habitats, Bats as Habitats* working group (GBatNet) meets virtually, all are welcome
  - *Please email m.ingala@fdu.edu if interested!!*