MMBL Symbiont Metabarcoding Workshop 2023

1. 16S rRNA sequence analysis (QIIME2)

2. ITS2 rRNA sequence analysis (SymPortal)

Analysis of 16S rRNA sequences for microbial diversity studies

MMBL Symbiont Metabarcoding Workshop
November 11-12, 2023

Jake Ivan P. Baquiran
John Bennedick Quijano
Dr. Cecilia Conaco

Outline

- 16S rRNA gene
- Sequencing platforms and protocol
- 16S rRNA sequence analysis pipeline
- Diversity measures
  - Alpha diversity
  - Beta diversity
- Ordination
- Demonstration

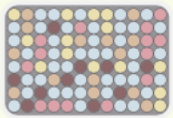
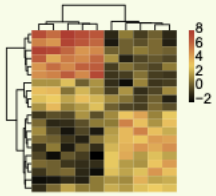
	Method	Advantages	Limitations
	Culturome	• High-throughput • Targeted selection • Provides microbial isolates	• Expensive • Laborious • Influenced by media and the environment
	Amplicon (16S/18S/ITS)	• Quick analysis • Low-biomass requirement • Applicable to samples contaminated by host DNA	• PCR and primer biases • Resolution limited to genus level • False positive in low-biomass samples
	Metagenome	• Taxonomic resolution to species or strain level • Functional potential • Uncultured microbial genome	• Expensive • Time-consuming in analysis • Host-derived contamination

Fig 1. Advantages and limitations of common HTS methods used in microbiome research.

Liu et al. (2020). *Protein & Cell* 12(5):315-330.

16S rRNA gene

- Marker gene for bacterial phylogeny and taxonomy
- most well-studied and characterized gene
- present in almost all bacteria
- 1,500 bp in length
- comprised of conserved and variable regions

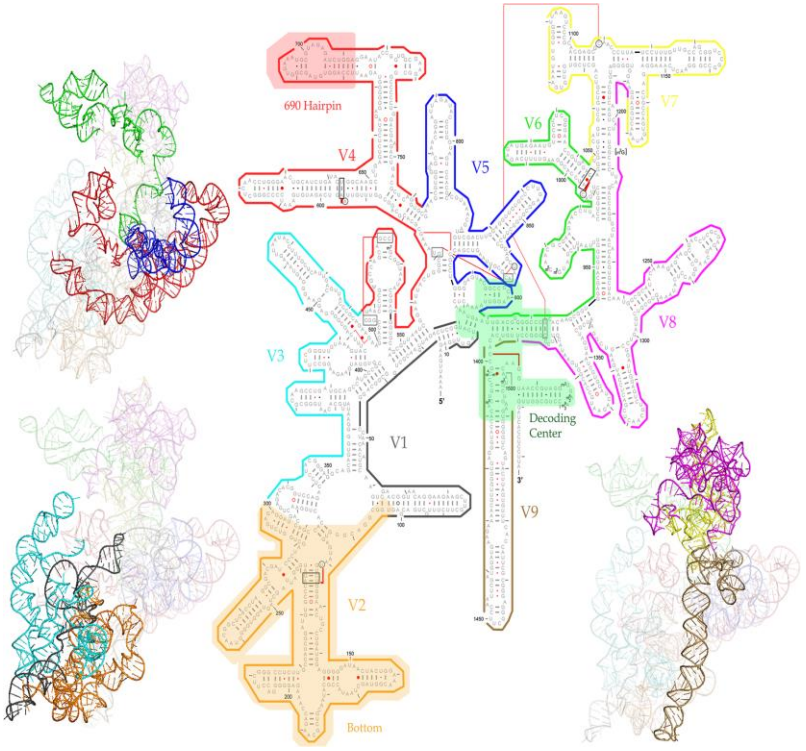


Fig 2. The 2D-3D structures of the 16S rRNA gene.

Yang et al. (2016) . BMC Bioinformatics 17:135.

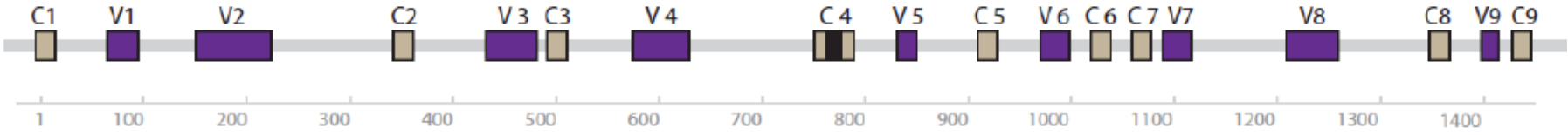


Fig. 3. Schematic representation of the 16S rRNA gene. Location of variable (purple) and conserved (brown) regions in a bacterial 16S rRNA.

conserved regions- allow primers to be designed to target all bacteria

variable regions- allow for determination of various species

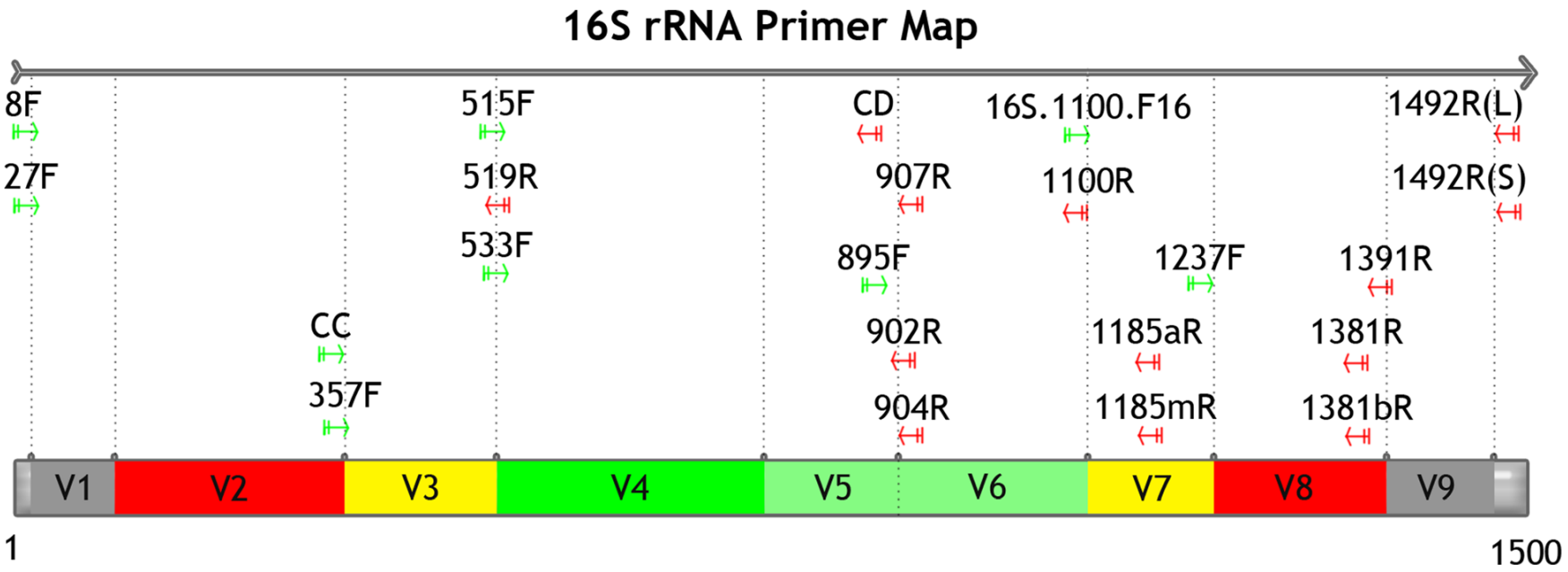


Fig. 4. Illustration of different variable regions.

Yang et al. BMC Bioinformatics (2016) 17:135

Table 1. Pros and cons of different primer pairs

	V4–5	V4	V3–4	V1–2
Populations biases: consensus from mock communities and field samples				
Populations under-represented[b]	SAR11 Deep 1 *Pseudospirillum*	ZD0405 (field)	*Euryarchaeota* (field) *Thaumarchaeota* SAR11 Surface 1 SAR11 Deep 1 some SAR116	SAR11 Deep 1 *Roseobacter* DC5–80-3 *Roseobacter* OCT
Populations over-represented[b]		*Euryarchaeota* *Thaumarchaeota*	*Euryarchaeota* (mock)	*Flavobacteria*: NS5
Populations not detected	*Roseobacter* OCT	some SAR116		*Euryarchaeota*[c] *Thaumarchaeota*[c] *Roseobacter* NAC11-7 some SAR116
Clades with poor classification[d]	*Rhodobacteraceae*			*Rhodobacteraceae*

Wear et al. (2018) Environmental Microbiology **2**0(8),7009—020726

Sample	T1						
Primer Pairs	27F – 338R	27F – 534R	341F – 785R	515F – 806R	515F – 944R	939F – 1378R	1115F – 1492R
Actinobacteria	■	■	■	■	■	■	■
Bacteroidetes	■	■	■	■	■	■	■
Firmicutes	■	■	■	■	■	■	■
Lentisphaerae							
Proteobacteria	■	■	■	■	■	■	■
Tenericutes							
Verrucomicrobia			■	■		■	■
unclassified *Bacteria*							
unknown *Bacteria*							

Sample	T7						
Primer Pairs	27F – 338R	27F – 534R	341F – 785R	515F – 806R	515F – 944R	939F – 1378R	1115F – 1492R
Actinobacteria	■	■	■	■	■	■	■
Bacteroidetes	■	■	■	■		■	■
Firmicutes	■	■	■	■	■	■	■
Lentisphaerae							
Proteobacteria	■	■	■	■	■	■	■
Tenericutes							
Verrucomicrobia							
unclassified *Bacteria*				■			
unknown *Bacteria*							

Sample	T17						
Primer Pairs	27F – 338R	27F – 534R	341F – 785R	515F – 806R	515F – 944R	939F – 1378R	1115F – 1492R
Actinobacteria	■	■	■	■	■	■	■
Bacteroidetes	■	■	■	■		■	■
Firmicutes	■	■	■	■	■	■	■
Lentisphaerae							
Proteobacteria	■	■	■	■	■	■	■
Tenericutes							
Verrucomicrobia							
unclassified *Bacteria*							
unknown *Bacteria*							

Sample	T19						
Primer Pairs	27F – 338R	27F – 534R	341F – 785R	515F – 806R	515F – 944R	939F – 1378R	1115F – 1492R
Actinobacteria			■	■		■	■
Bacteroidetes	■	■	■	■		■	■
Firmicutes	■	■	■	■	■	■	■
Lentisphaerae							
Proteobacteria	■	■	■	■	■	■	■
Tenericutes					■		
Verrucomicrobia			■	■		■	■
unclassified *Bacteria*							
unknown *Bacteria*							

Sample	T20						
Primer Pairs	27F – 338R	27F – 534R	341F – 785R	515F – 806R	515F – 944R	939F – 1378R	1115F – 1492R
Actinobacteria	■	■	■	■		■	■
Bacteroidetes	■	■	■	■	■	■	■
Firmicutes	■	■	■	■	■	■	■
Lentisphaerae				■		■	■
Proteobacteria	■	■	■	■	■	■	■
Tenericutes							
Verrucomicrobia			■	■		■	■
unclassified *Bacteria*				■			
unknown *Bacteria*				■	■		

Sample	T22						
Primer Pairs	27F – 338R	27F – 534R	341F – 785R	515F – 806R	515F – 944R	939F – 1378R	1115F – 1492R
Actinobacteria	■	■	■	■	■	■	■
Bacteroidetes	■	■	■	■	■	■	■
Firmicutes	■	■	■	■	■	■	■
Lentisphaerae				■			
Proteobacteria	■	■	■	■	■	■	■
Tenericutes							
Verrucomicrobia							
unclassified *Bacteria*				■			
unknown *Bacteria*							

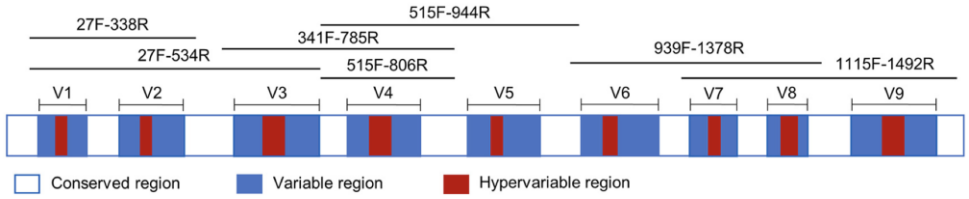


Fig. 5. “Presence-and-absence” map of human samples on phylum level for different V-regions.

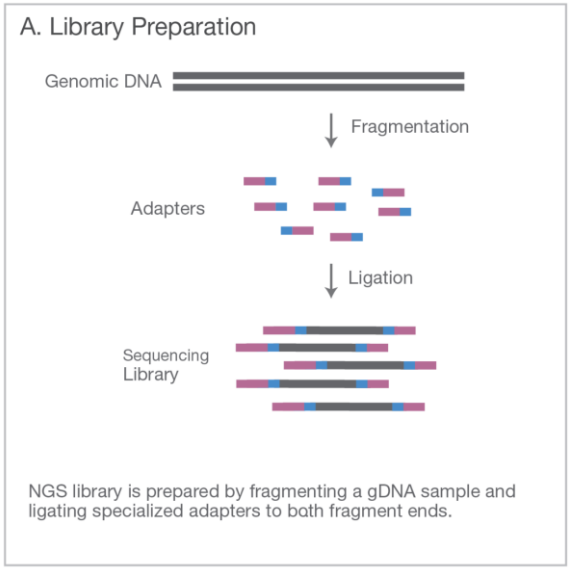
Abelan-Scheneyder et al. (2021) mSphere 6:e01202-20

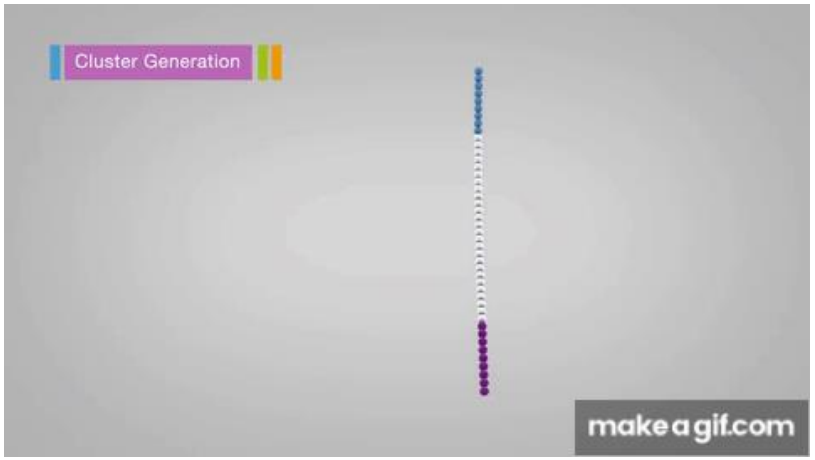
Sequencing platforms

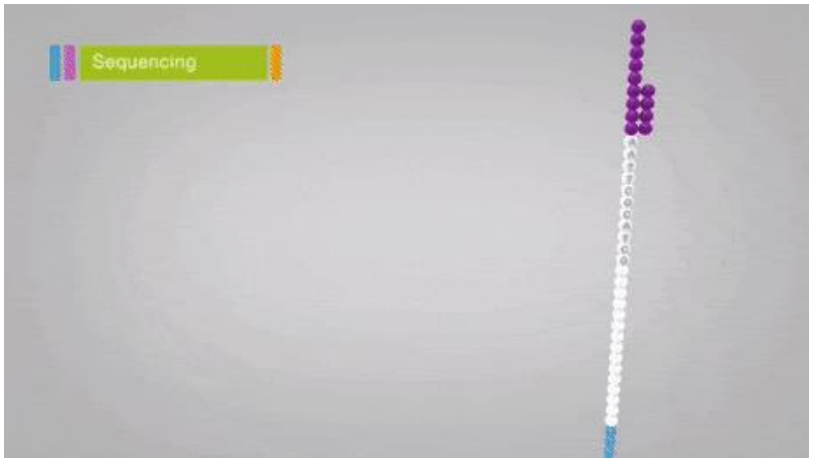
Table 2. Different platforms comparisons.

	Roche 454	Ion Torrent	Illumina MiSeq
Sequencing Kit	GS FLX Titanium XLR70	PGM 400 Sequencing	MiSeq Reagent Kits v2
Expected Read Length	Up to 600 bp	Up to 400 bp	MiSeq Reagent Kit v2: Up to 2 × 250 bp
Typical Throughput	450 Mb	Ion 314™ Chip v2: Up to 100 Mb Ion 316™ Chip v2: Up to 1 Gb Ion 318™ Chip v2: Up to 2 Gb	Up to 8.5 Gb
Reads per Run	~1000,000 shotgun, ~700,000 amplicon	Ion 314™ Chip v2: 400–550 thousand Ion 316™ Chip v2: 2–3 millions Ion 318™ Chip v2: 4–5.5 millions	~15 million reads
Consensus Accuracy	99.995%	99%	99%
Run Time	10 h	Ion 314™ Chip v2: 2.3 to 3.7 h Ion 316™ Chip v2: 3.0 to 4.9 h Ion 318™ Chip v2: 4.4 to 7.3 h	4 h and approximately 39 h depending on the number of cycles
Sample Input	gDNA, cDNA, or amplicons (PCR products)	gDNA, cDNA, or amplicons (PCR products)	gDNA, cDNA, or amplicons (PCR products) Small genome, amplicon, and targeted gene panel sequencing
Weight	532 lbs. (242 kg)	65 lbs. (30 kg)	120 lbs. (54.5 kg)
Instrument cost	~$500 K	~ $80 k	~ $125 k

Allali et al. (2017) BMC Microbiology 17:194







Illumina MiSeq protocol

16S pipeline

Quantitative Insights Into Microbial Ecology version 2
(https://docs.qiime2.org)

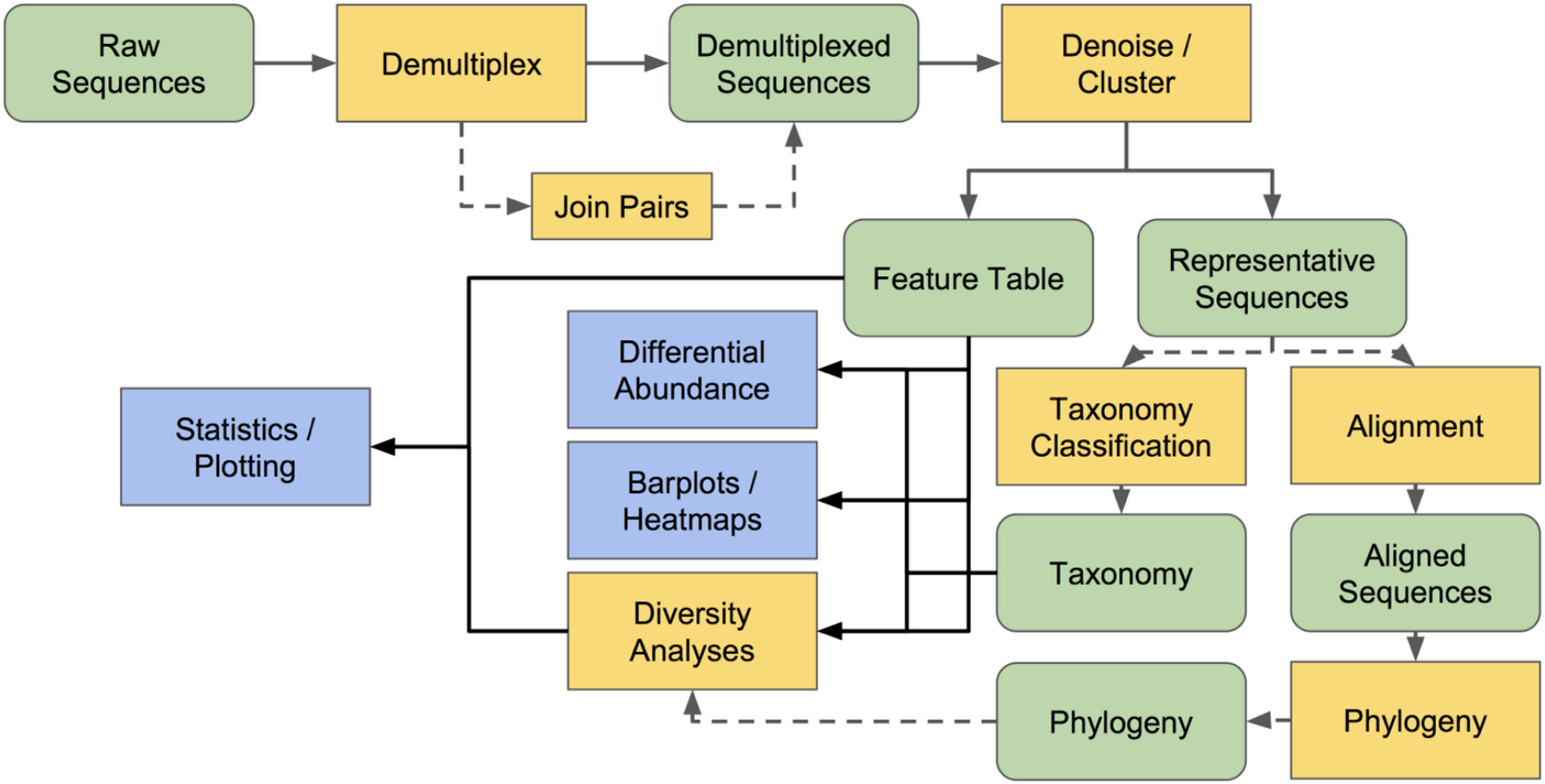
Mothur
(https://www.mothur.org)

SILVAngs
(https://www.arb-silva.de/silvangs)

and others

Abelan-Scheneyder et al. (2021) mSphere 6:e01202-20

QIIME2 pipeline overview



https://qiime2.org

Demultiplexing

Multiplexing:
During sequencing, samples were mixed together or multiplex to minimize sequencing cost.

How do we know which sample each read came from?

This is typically done by appending a **unique barcode** (a.k.a. index or tag) sequence to one or both ends of each sequence. Detecting these barcode sequences and mapping them back to the samples they belong allows us to *demultiplex* sequences.

Demultiplexing

Adapter Barcode Overhang adapter

Overhang adapter Barcode Adapter

Sample ID	Barcode sequence
Sample 1	AAGAGGCAGTAAGGAG
Sample 2	GTAGAGGAACTGCATA
Sample 3	GCTCATGAAAGGAGTA
Sample 4	ATCTCAGGCTAAGCCT
Sample 5	ACTCGCTACGTCTAAT
Sample 6	GGAGCTACTCTCTCCG

Removing non-biological sequences

For example:

16S Amplicon PCR Forward Primer
TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG

16S Amplicon PCR Reverse Primer
GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC

Available tools to remove non-biological sequences:

- Trimmomatic
- Trim Galore
- Sickle
- Cutadapt - available in Qiime2 workflow

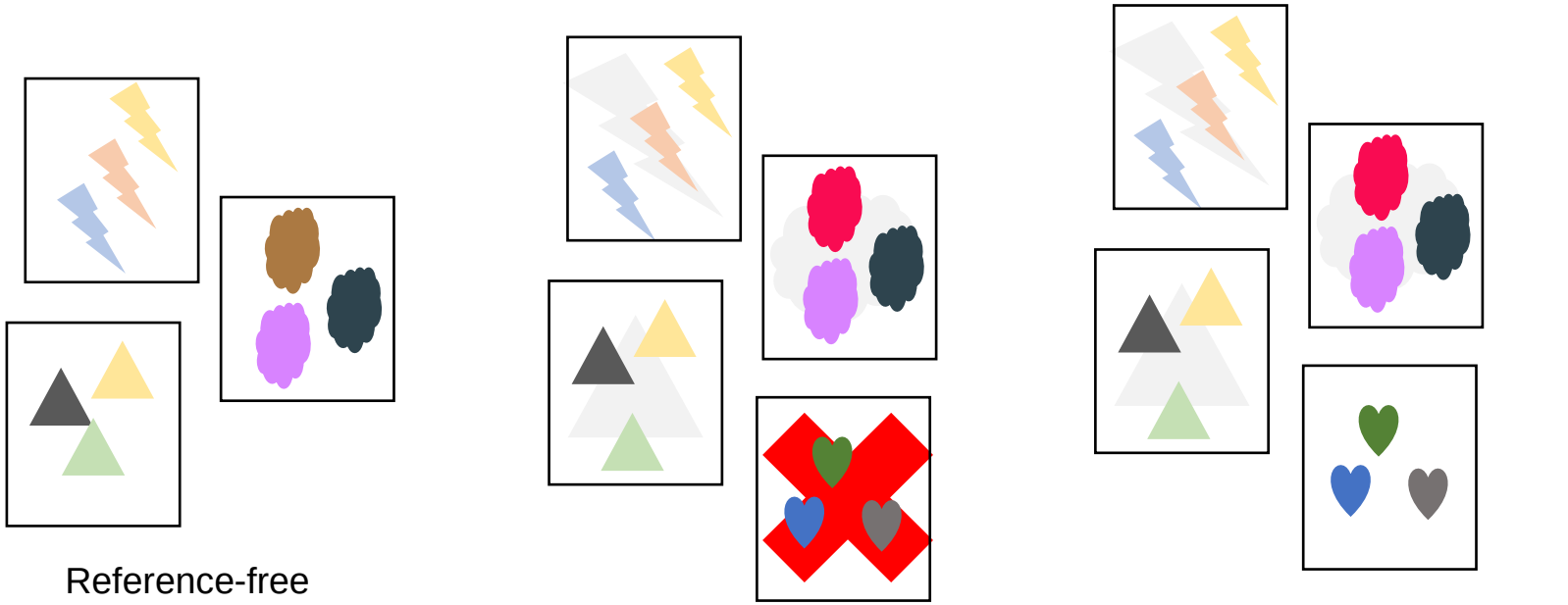
Minimizing the effects of targeted sequencing error

Strategies:

a. OTU clustering- based upon the idea that related/similar organisms will have similar target gene sequences and that rare sequencing errors will have a trivial contribution

- clusters often being generated using a similarity threshold of 97% sequence identity

- Reference-free (De novo) OTU Clustering and Reference-based OTU Clustering (Open and Closed).



Reference-free

Reference-based (Closed)

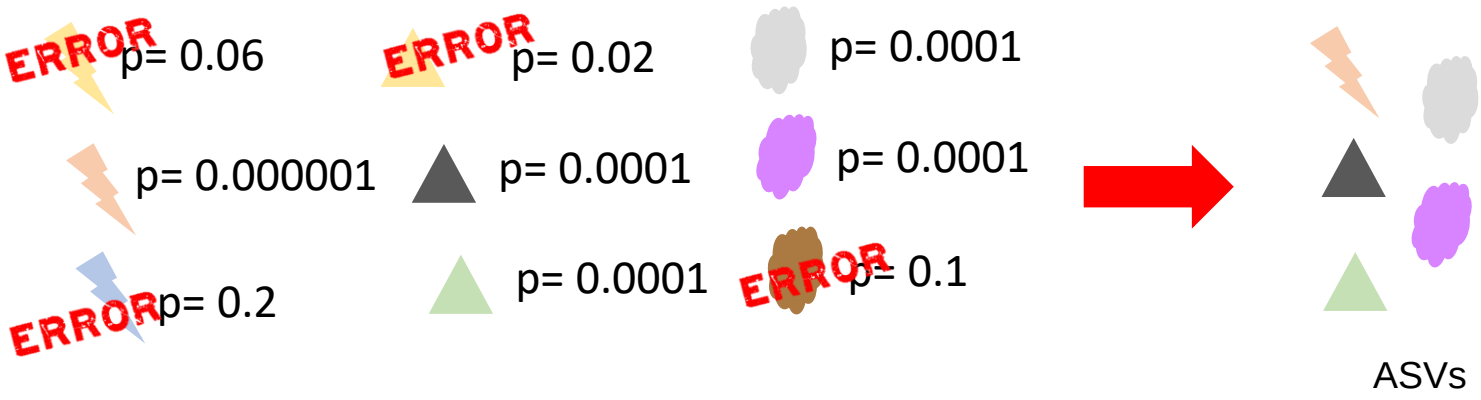
Reference-based (Open)

Nearing et al. (2018) *PeerJ*6:e5364

Minimizing the effects of targeted sequencing error

Strategies:

b. Amplicon sequence variant approach- determines which exact sequences were read and how many times each exact sequence was read
-even a single base difference in the sequence will result in a unique ASV
-data will be combined with an an error model for the sequencing run, enabling the comparison of similar reads to determine the probability that a given read at a given frequency is not due to sequencer error
-sometimes called exact sequence variant or zero-radius OTU (zOTU)



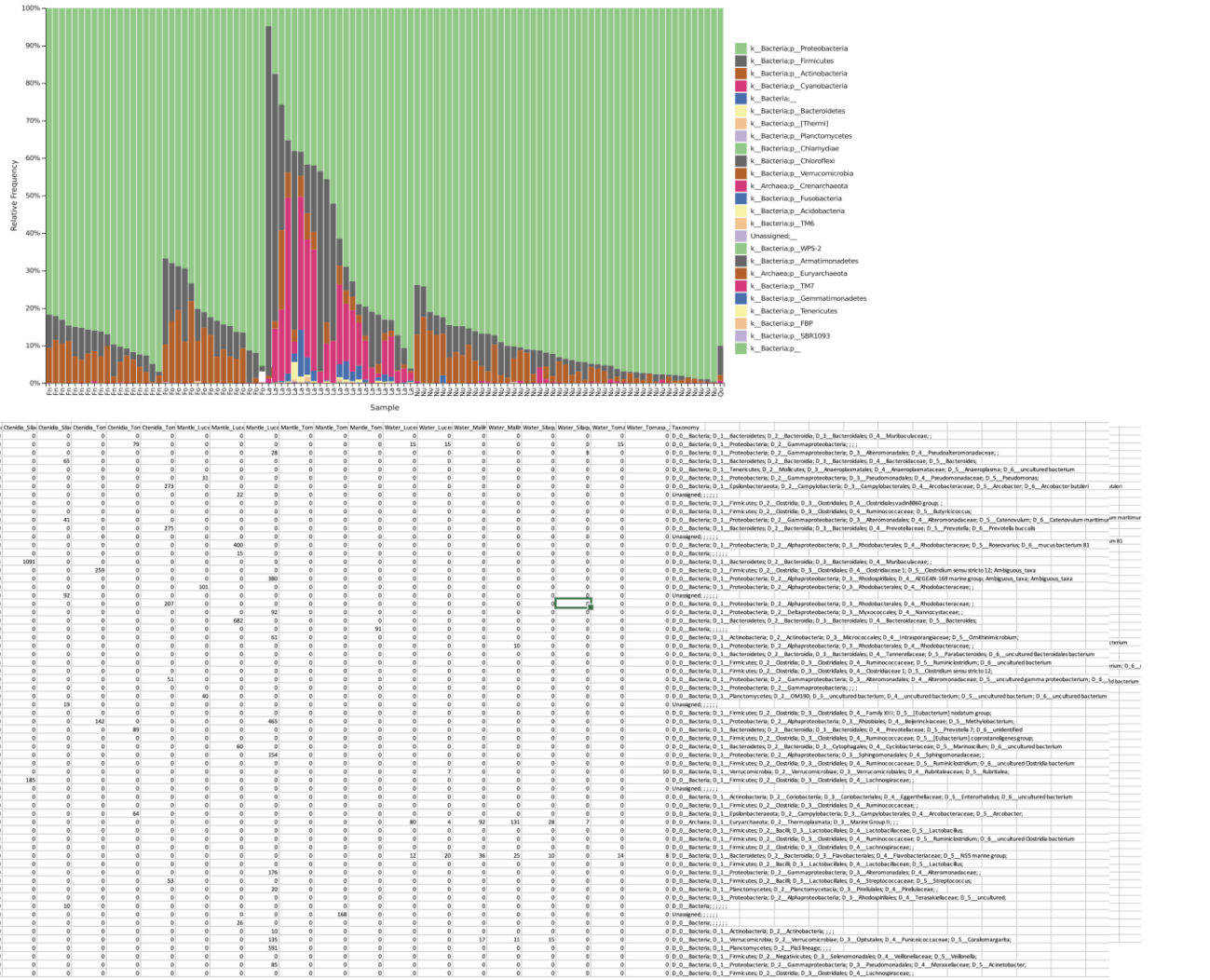
Nearing et al. (2018) *PeerJ*6:e5364

Taxonomic assignment

Table 3. Overview of different taxonomic classifications

Taxonomy	Domains	Lowest rank	Latest release	Sources
SILVA	Bacteria, Archaea, Eukarya	Species	August 2020	Bergey's Taxonomic Outlines; List of Prokaryotic Names with Standing in Nomenclature; International Society of Protistologists
Ribosomal Database Project (RDP)	Bacteria, Archaea, Fungi	Genus	August 2020	International Nucleotide Sequence Database Collaboration
Greengenes (GG)	Bacteria, Archaea	Species	October 2022	Mainly NCBI
NCBI	All organisms	Species	Today	Catalog of Life; the Encyclopedia of Life; Name- Bank; WikiSpecies
Open tree of life taxonomy (OTT)	All organisms	Species	September 2023	IndexFungorum; SILVA, NCBI, Global Biodiversity Information Facility ; Interim Register of Marine and Nonmarine Genera

View the taxonomic classification and raw count table



#OTU ID	Ctenidia_Lucens_1	Ctenidia_Lucens_2	Ctenidia_Luc	Ctenidia_Mal	Ctenidia_Mal	Ctenidia_Mal	Ctenidia_Silv	Ctenidia_Silv	Ctenidia_Tor	Ctenidia_Tor	Mantle_Luce	Mantle_Luce	Mantle_Luce	Mantle_Tom	Mantle_Tom	Mantle_Tom	Water_Luce	Water_Luce	Water_Malli	Water_Malli	Water_Silva	Water_Silva	Water_Tom	Water_Tomasus	Taxonomy
1214ec7ec3359fe8b93	0	272	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Bacteroidetes; D_2__Bacteroidia; D_3__Bacteroidales; D_4__Muribaculaceae; ;
c60f0a8363c9cc1a15f3	0	0	0	0	0	0	0	0	79	0	0	0	0	0	0	0	15	15	0	0	0	0	15	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Gammaproteobacteria; ; ; ;
1f785257fdb712df9aaec	0	0	0	0	0	0	0	0	0	0	0	0	28	0	0	0	0	0	0	0	8	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Gammaproteobacteria; D_3__Alteromonadales; D_4__Pseudoalteromonadaceae; ;
1e392135ac1fe0e04c0f	0	0	0	0	0	0	0	65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Bacteroidetes; D_2__Bacteroidia; D_3__Bacteroidales; D_4__Bacteroidaceae; D_5__Bacteroides;
3fa98463997568baaae63	0	0	0	0	0	114	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Tenericutes; D_2__Mollicutes; D_3__Anaeroplasmatales; D_4__Anaeroplasmataceae; D_5__Anaeroplasma; D_6__uncultured bacterium
78a7bfd551495d29122f	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Gammaproteobacteria; D_3__Pseudomonadales; D_4__Pseudomonadaceae; D_5__Pseudomonas;
0fd8c1f13020dcc6541f	0	0	0	0	0	0	0	0	273	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Epsilonbacteraeota; D_2__Campylobacteria; D_3__Campylobacterales; D_4__Arcobacteraceae; D_5__Arcobacter; D_6__Arcobacter butzleri
297820efddu2e430ffcc	0	0	0	0	0	0	0	0	0	0	0	22	0	0	0	0	0	0	0	0	0	0	0	0	Unassigned; ; ; ; ; ;
7337c0a747eb7a3976c	0	0	105	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Clostridiales vadinBB60 group; ;
2dfc1ba9b4c4eb41c60	0	0	0	0	102	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Bacteroidetes; D_2__Bacteroidia; D_3__Bacteroidales; D_4__Rikenellaceae; D_5__Alistipes;
e254c004cbc23e36b5f6	0	0	0	0	0	0	0	0	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Gammaproteobacteria; D_3__Alteromonadales; D_4__Alteromonadaceae; D_5__Catenovulum; D_6__Catenovulum maritimum
d09f90dd04d4093eb6f6	0	0	9	0	0	0	0	0	0	0	275	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Bacteroidetes; D_2__Bacteroidia; D_3__Bacteroidales; D_4__Prevotellaceae; D_5__Prevotella; D_6__Prevotella buccalis
182eb2b449f919cdbc95	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Unassigned; ; ; ; ; ;
453d53188aa38e0664d	0	0	0	0	0	0	0	0	0	0	0	400	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Alphaproteobacteria; D_3__Rhodobacterales; D_4__Rhodobacteraceae; D_5__Roseovarius; D_6__mucus bacterium 81
bf8c0df1c3b058b15014	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; ; ; ; ; ;
a2bde3fd7aa0a931273ad	0	0	225	79	0	181	84	1091	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Bacteroidetes; D_2__Bacteroidia; D_3__Bacteroidales; D_4__Muribaculaceae; ;
8c070bb0e226fbe50f8	0	323	0	0	0	29	0	0	259	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Clostridiaceae 1; D_5__Clostridium sensu stricto 12; Ambiguous_taxa
9f51d0184397f8aa7ae6	0	0	0	0	0	0	0	0	0	0	0	0	0	380	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Alphaproteobacteria; D_3__Rhodospirillales; D_4__AEGEAN-169 marine group; Ambiguous_taxa; Ambiguous_taxa
4edc24c94aaa344d13b	0	0	0	0	0	0	0	0	0	0	101	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Alphaproteobacteria; D_3__Rhodobacterales; D_4__Rhodobacteraceae; ;
5bc742eac3ac1b26935	0	0	0	0	0	0	0	0	92	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Unassigned; ; ; ; ; ;
9eecde0ad736ef4c9863	0	0	0	0	0	0	0	0	0	0	207	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Alphaproteobacteria; D_3__Rhodobacterales; D_4__Rhodobacteraceae; ;
3dbb93e9ac8f9cc9987c	0	0	0	0	0	0	0	0	0	0	0	0	92	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Bacteroidetes; D_2__Deltaproteobacteria; D_3__Myxococcales; D_4__Nannocystaceae; ;
ad44283a88fcc95b08b	0	0	0	0	0	0	0	0	0	0	0	682	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Bacteroidetes; D_2__Bacteroidia; D_3__Bacteroidales; D_4__Bacteroidaceae; D_5__Bacteroides;
3065869baa1b9b8d5c0b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	91	0	0	0	0	0	0	0	D_0__Bacteria; ; ; ; ; ;
6d051ce40962dac434f6	0	0	0	0	0	0	0	0	0	0	0	0	61	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Actinobacteria; D_2__Actinobacteria; D_3__Micrococcales; D_4__Intrasporangiaceae; D_5__Ornithinimicrobium;
8cd42a0fe6670dd05422	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Alphaproteobacteria; D_3__Rhodobacterales; D_4__Rhodobacteraceae; ;
c9328d0ce5d43a8e6924	0	0	154	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Bacteroidetes; D_2__Bacteroidia; D_3__Bacteroidales; D_4__Tannerellaceae; D_5__Parabacteroides; D_6__uncultured Bacteroidales bacterium
5e82a006a7e00d3d28e	0	0	0	0	0	94	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Ruminococcaceae; D_5__Ruminiclostridium; D_6__uncultured bacterium
df37d9f257b2941cdddf	0	586	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Clostridiaceae 1; D_5__Clostridium sensu stricto 12;
b049ac4bca9463a8f7f9	0	0	0	0	0	0	0	0	51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Gammaproteobacteria; D_3__Alteromonadales; D_4__Alteromonadaceae; D_5__uncultured gamma proteobacterium; D_6__
364970de5466c5d1cc	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Gammaproteobacteria; ; ; ;
99b2baf72e76267c413	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Planctomycetes; D_2__OM190; D_3__uncultured bacterium; D_4__uncultured bacterium; D_5__uncultured bacterium; D_6__uncultured bacterium
56ad870fbc5aa7853344	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Unassigned; ; ; ; ; ;
876b2dc9f542bb83279c	0	0	107	0	0	69	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Family XIII; D_5__[Eubacterium] nodatum group;
ce74a7402a6683cddbd	0	0	0	0	0	0	0	0	142	0	0	0	0	461	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Alphaproteobacteria; D_3__Rhodobacterales; D_4__Rhodobacteraceae; D_5__Methylobacterium;
0870ec7b4ae09de14f4f	0	0	0	0	0	0	0	0	0	89	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Bacteroidetes; D_2__Bacteroidia; D_3__Bacteroidales; D_4__Prevotellaceae; D_5__Prevotella 7; D_6__unidentified
0c9bef75eb83417a5045	199	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Ruminococcaceae; D_5__[Eubacterium] coprostanoligenes group;
900f0dc5166e8310e5a5	0	0	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Epsilonbacteraeota; D_2__Campylobacteria; D_3__Campylobacterales; D_4__Cyclobacteriaceae; D_5__Marinoscillum; D_6__uncultured bacterium
ff59836f57d3b81c99fbf	0	0	0	96	0	0	0	0	0	0	0	0	0	154	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Alphaproteobacteria; D_3__Sphingomonadales; D_4__Sphingomonadaceae; ;
d6d80e61631c8facd5d1	0	247	154	0	0	124	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Ruminococcaceae; D_5__Ruminiclostridium; D_6__uncultured Clostridia bacterium
c2fbe8acad5b088cbaac	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	D_0__Bacteria; D_1__Verrucomicrobia; D_2__Verrucomicrobiae; D_3__Verrucomicrobiales; D_4__Rubritaleaceae; D_5__Rubritalea;
cfa539db416f73881ddf	0	0	0	0	0	0	185	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Lachnospiraceae; ;
8fba4c7c703032e6e290	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Unassigned; ; ; ; ; ;
713e0ff3652a4c836128c	0	0	81	0	0	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Actinobacteria; D_2__Coriobacteriia; D_3__Coriobacteriales; D_4__Eggerthellaceae; D_5__Enterorhabdus; D_6__uncultured bacterium
4685e76da73a335f0f0f	0	0	63	0	0	48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Ruminococcaceae; ;
651ad0ef3ff9f008c8af0	0	0	0	0	0	0	0	0	64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Epsilonbacteraeota; D_2__Campylobacteria; D_3__Campylobacterales; D_4__Arcobacteraceae; D_5__Arcobacter;
872fe0455b3831c9ebbb	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	80	4	92	131	28	7	0	0	D_0__Archaea; D_1__Thermoplasmata; D_2__Thermoplasmata; D_3__Marine Group II; ; ;
513eec1b7376f0ffa46c	0	0	53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Bacilli; D_3__Lactobacillales; D_4__Lactobacillaceae; D_5__Lactobacillus;
296fdf8228fe3eb55ec6	0	337	492	0	0	151	123	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Ruminococcaceae; D_5__Ruminiclostridium; D_6__uncultured Clostridia bacterium
0a992cf770cedaa3712f	0	0	0	0	0	132	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Lachnospiraceae; ;
a32e38f1d361b866f6b2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	20	36	25	10	0	14	8	D_0__Bacteria; D_1__Bacteroidetes; D_2__Bacteroidia; D_3__Flavobacteriales; D_4__Flavobacteriaceae; D_5__NS5 marine group;
0d770f1aee40fa9f6f1be	0	0	124	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Bacilli; D_3__Lactobacillales; D_4__Lactobacillaceae; D_5__Lactobacillus;
f1e84383484bf7ec9b43	0	0	0	0	0	0	0	0	0	0	0	0	176	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Gammaproteobacteria; D_3__Alteromonadales; D_4__Alteromonadaceae; ;
cb84ab53c82fe288c81f	0	0	0	0	0	0	0	0	53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Bacilli; D_3__Lactobacillales; D_4__Streptococcaceae; D_5__Streptococcus;
63de1d6c9fce33a1de51	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Planctomycetes; D_2__Planctomycetacia; D_3__Pirellulales; D_4__Pirellulaceae; ;
a017e767bfacb011f33f	0	0	0	0	0	26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Alphaproteobacteria; D_3__Rhodospirillales; D_4__Terasakiellaceae; D_5__uncultured;
0ac42df98dd60a1e065c	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; ; ; ; ; ;
0c8da1df823e6157099f	0	0	0	0	0	0	0	0	0	0	0	0	168	0	0	0	0	0	0	0	0	0	0	0	Unassigned; ; ; ; ; ;
fe94f67751be3b9d288f	0	0	0	0	0	0	0	0	0	0	26	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; ; ; ; ; ;
1153dc11fc78851fc23a	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Actinobacteria; D_2__Actinobacteria; ; ;
a94d1683bbda73d39d0	0	0	0	0	0	0	0	0	0	0	0	0	135	0	0	0	0	17	11	15	0	0	0	0	D_0__Bacteria; D_1__Verrucomicrobia; D_2__Verrucomicrobiae; D_3__Opitutales; D_4__Puniceicoccaceae; D_5__Coraliomargarita;
bd09fe0b9d62e1a810e6	0	0	0	0	0	0	0	0	0	0	0	0	591	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Planctomycetes; D_2__Pla3 lineage; ; ;
9802e2b5f3eca117ffc0	0	0	54	0	0	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Negativicutes; D_3__Selenomonadales; D_4__Veillonellaceae; D_5__Veillonella;
17a80d49f80d8baf2fdc	0	0	0	0	0	0	0	0	0	0	0	0	85	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Proteobacteria; D_2__Gammaproteobacteria; D_3__Pseudomonadales; D_4__Moraxellaceae; D_5__Acinetobacter;
742f3ac20a0768f202f7	0	0	0	0	6	123	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	D_0__Bacteria; D_1__Firmicutes; D_2__Clostridia; D_3__Clostridiales; D_4__Lachnospiraceae; ;

https://qiime2.org

Data analysis

- **Rarefaction curve:**estimates covered diversity by available sequencing depth

  -Plateauing means sufficient sequencing effort

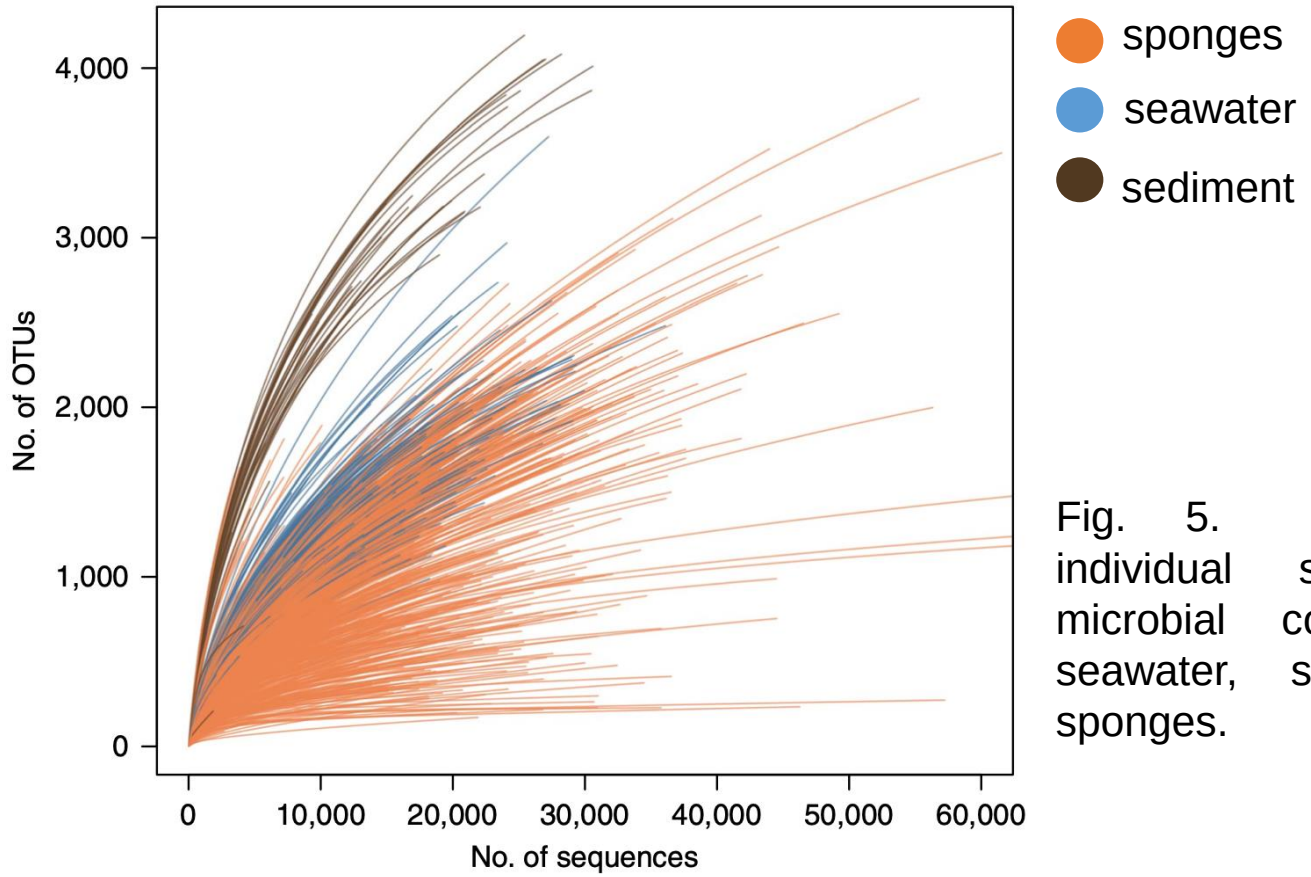


Fig. 5. Richness of individual samples from microbial communities in seawater, sediments and sponges.

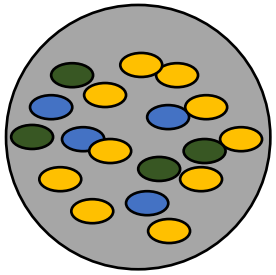
Thomas et al. (2016) Nature Communications 7:11870

Diversity measures

Alpha diversity (within)

What is there? How much is there?

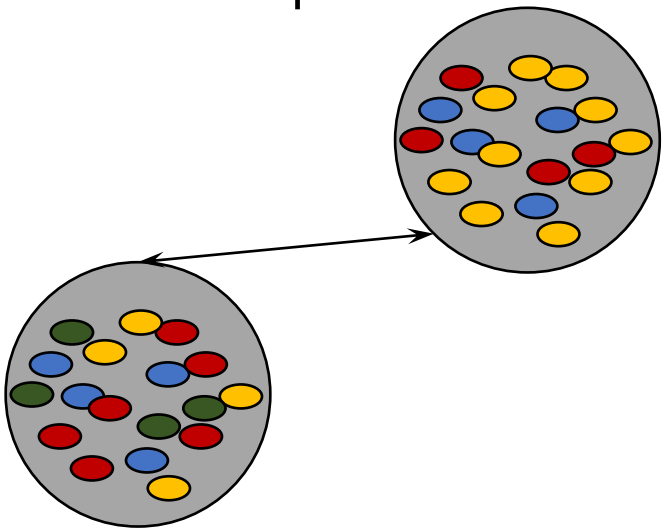
- Diversity within one sample
- Richness
- Evenness



Beta diversity (between)

How similar or different are samples?

- Diversity between samples (comparison)
- Dissimilarity
- Shared species



Alpha diversity

- **Species Richness**- species counts

Sample A
Endozoicomonas acroporae
Endozoicomonas montiporae
Endozoicomonas gorgoniicola

Sample B
Endozoicomonas acroporae
Endozoicomonas montiporae

Sample C
Endozoicomonas acroporae

Richness:
Sample A:3
Sample B:2
Sample C:1

Conclusion:
Sample A is more diverse than Sample B, and B is more diverse than C.

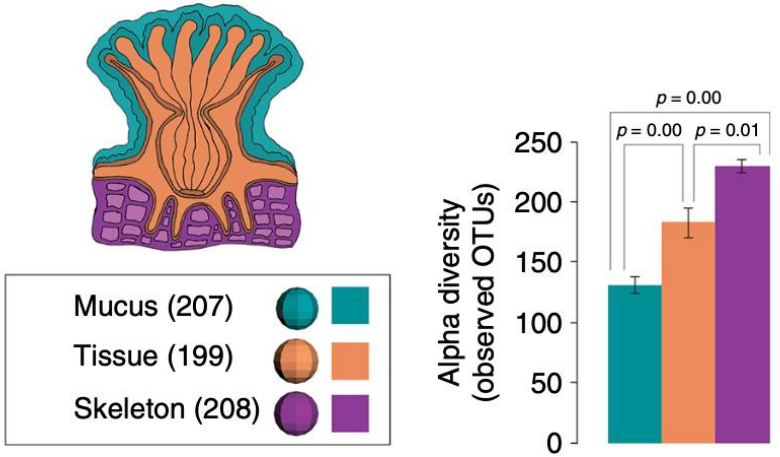


Fig. 6. Richness of microbes from different coral compartments.

Pollock et al. (2018) Nature Communications 9:4921

Alpha diversity

- **Species Richness**- species counts

Sample A

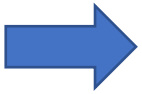
Endozoicomonas acroporae
Endozoicomonas montiporae
Endozoicomonas gorgoniicola

Sample B

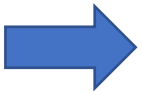
Endozoicomonas acroporae
Endozoicomonas montiporae
Vibrio coralliilyticus

Sample C

Endozoicomonas acroporae
Vibrio coralliilyticus
Bacillus flexus



Richness:
Sample A:3
Sample B:3
Sample C:3



Conclusion:
Samples A, B and C are equally diverse.

*ignores relatedness of the species

Alpha diversity

- **Phylogenetic divesity index (PD)** - based on phylogeny

Sample A

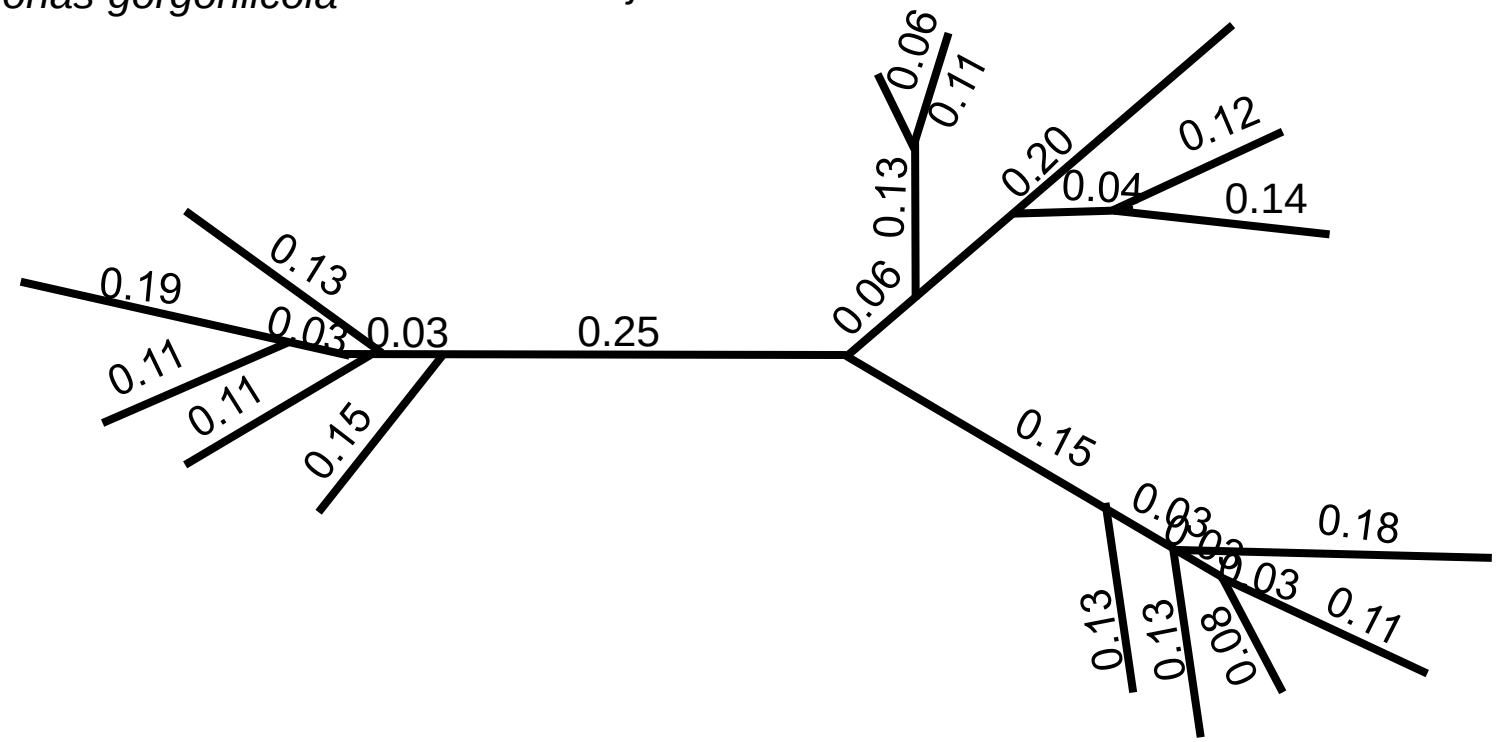
Endozoicomonas acroporae
Endozoicomonas montiporae
Endozoicomonas gorgoniicola

Sample B

Endozoicomonas acroporae
Endozoicomonas montiporae
Hahella chejuensis

Sample C

Endozoicomonas acroporae
Vibrio coralliilyticus
Bacillus flexus

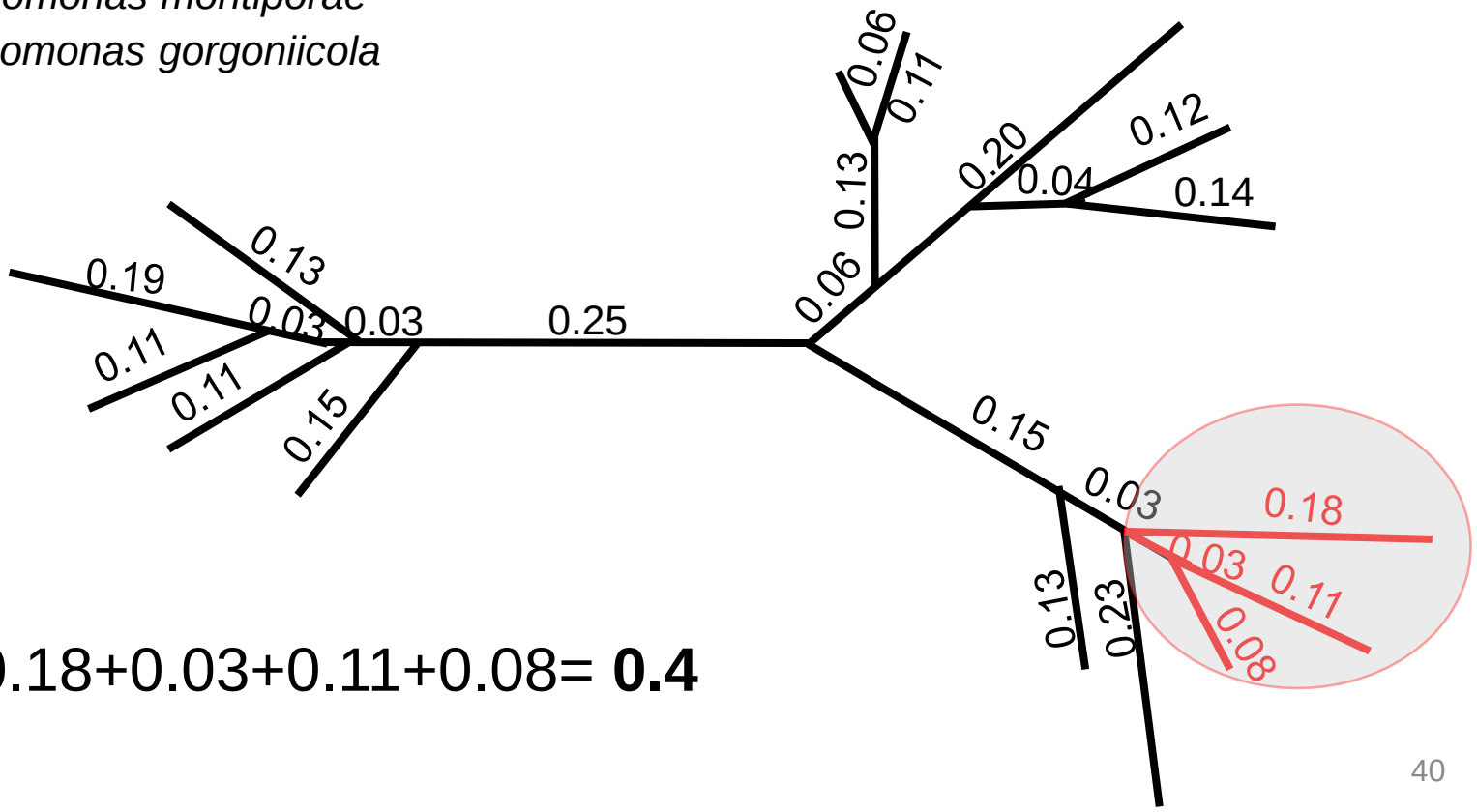


Alpha diversity

- **Phylogenetic divesity index (PD)** - based on phylogeny

Sample A

Endozoicomonas acroporae
Endozoicomonas montiporae
Endozoicomonas gorgoniicola



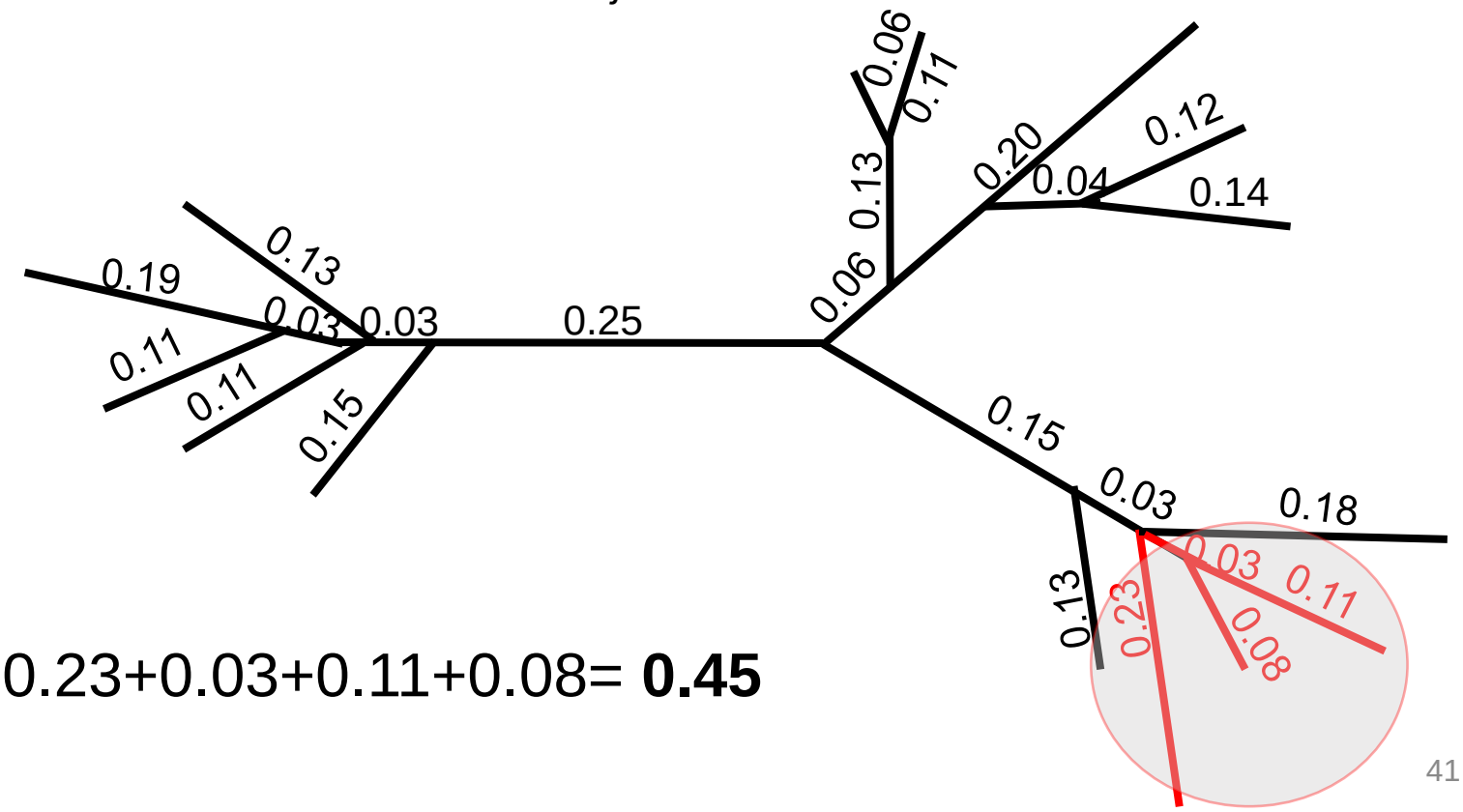
PD= 0.18+0.03+0.11+0.08= **0.4**

Alpha diversity

- **Phylogenetic divesity index (PD)** - based on phylogeny

Sample B

Endozoicomonas acroporae
Endozoicomonas montiporae
Hahella chejuensis



PD= 0.23+0.03+0.11+0.08= **0.45**

Alpha diversity

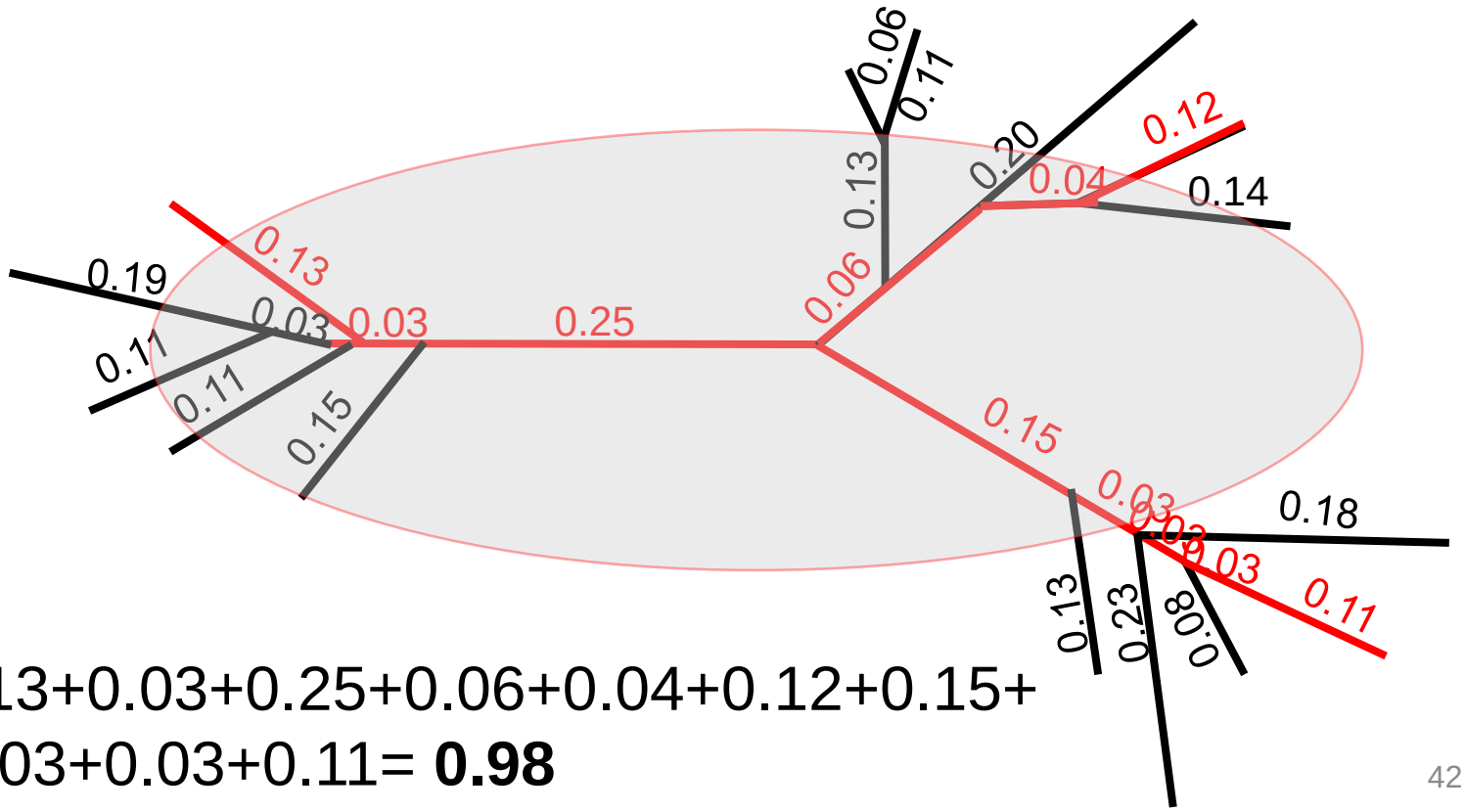
- **Phylogenetic divesity index (PD)** - based on phylogeny

Sample C

Endozoicomonas acroporae

Vibrio coralliilyticus

Bacillus flexus



PD= 0.13+0.03+0.25+0.06+0.04+0.12+0.15+ 0.03+0.03+0.03+0.11= **0.98**

Alpha diversity

- **Phylogenetic divesity index (PD)** - based on phylogeny

Sample A	Sample B	Sample C
Endozoicomonas acroporae	*Endozoicomonas acroporae*	*Endozoicomonas acroporae*
Endozoicomonas montiporae	*Endozoicomonas montiporae*	*Vibrio coralliilyticus*
Endozoicomonas gorgoniicola	*Hahella chejuensis*	*Bacillus flexus*
PD= 0.40	PD= 0.45	PD= 0.98

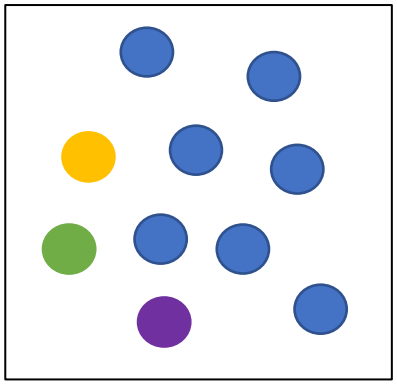
PD= 0.40 < PD= 0.45 < PD= 0.98

Conclusion:
Sample C is more diverse than Sample B,
which is more diverse than Sample A

Alpha diversity

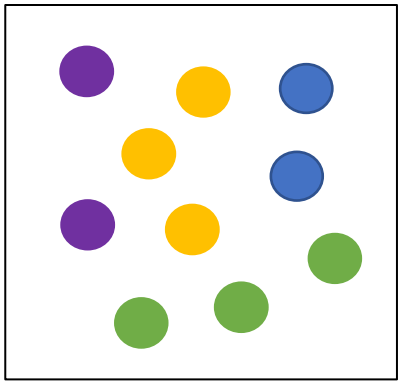
- **Shannon Diversity index (H)-** measures the number of species richness but scales them based on the evenness of the community, more weight on species richness

Sample A



Abundance= 10
Species Richness= 4
Diversity= ?

Sample B



Abundance= 10
Species Richness= 4
Diversity= ?

Kim et al. (2017) J. Microbiol. Biotechnol. 27, 2089–2093

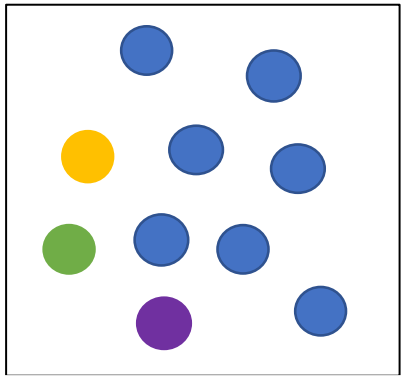
Alpha diversity

- **Shannon Diversity index (H)-** measures the number of species richness but scales them based on the evenness of the community, more weight on species richness

$$H = -\sum(Pi) \times ln(Pi)$$

where Pi is the proportion of individuals in each species

Sample A



Species	Abundance	Pi	Ln (Pi)	Pi x Ln(Pi)
Blue	7	0.70	-0.51	-0.36
Yellow	1	0.10	-2.30	-0.23
Green	1	0.10	-2.30	-0.23
Violet	1	0.10	-2.30	-0.23
Total	10			-1.05

Sample A **H= 1.05**

Kim et al. (2017) J. Microbiol. Biotechnol. 27, 2089–2093

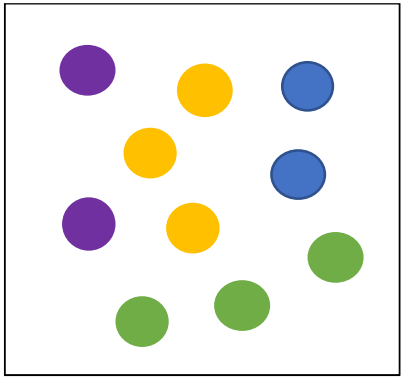
Alpha diversity

- **Shannon Diversity index (H)-** measures the number of species richness but scales them based on the evenness of the community

$$H = -\sum(Pi) \times ln(Pi)$$

where Pi is the proportion of individuals in each species

Sample B



Species	Abundance	Pi	Ln (Pi)	Pi x Ln(Pi)
Blue	2	0.20	-1.61	-0.32
Yellow	3	0.30	-1.20	-0.36
Green	3	0.30	-1.20	-0.36
Violet	2	0.20	-1.61	-0.32
Total	10			-1.36

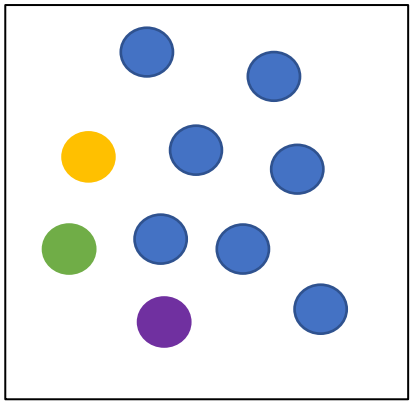
Sample B H= **1.36**

Kim et al. (2017) J. Microbiol. Biotechnol. 27, 2089–2093

Alpha diversity

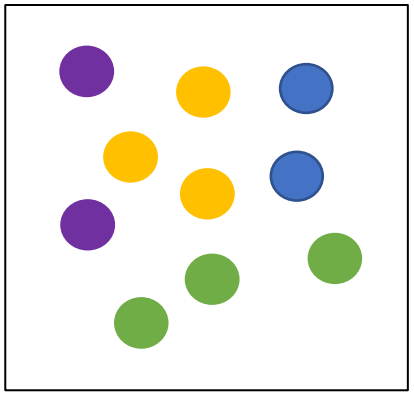
- **Shannon Diversity index (H)-** measures the number of species richness but scales them based on the evenness of the community

Sample A



Abundance= 10
Species Richness= 4
H= 1.00

Sample B



Abundance= 10
Species Richness= 4
H= 1.36

Conclusion:
Sample B is more diverse than Sample A

Kim et al. (2017) J. Microbiol. Biotechnol. 27, 2089–2093

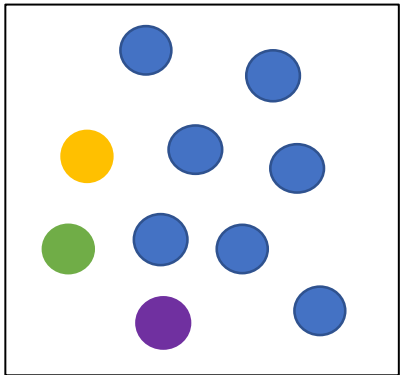
Alpha diversity

- **Inverse Simpson (D)-** measures the number of species richness but scales them based on the evenness of the community, more weight on species evenness

$$D = \frac{N(N-1)}{\sum n(n-1)}$$

N = total number of organisms of all species found
n = number of individuals of a particular species

Sample A



D= 10 (10-1)/ 7(7-1)+1(1-1)+1(1-1)+1(1-1)
D=90/42+0+0+0
D= 2.14

Kim et al. (2017) J. Microbiol. Biotechnol. 27, 2089–2093

Alpha diversity

- **Inverse Simpson (D)-** measures the number of species richness but scales them based on the evenness of the community, more weight on species evenness

$$D = \frac{N(N-1)}{\sum n(n-1)}$$

N = total number of organisms of all species found
n = number of individuals of a particular species

Sample B

D= 10 (10-1)/ 3(3-1)+3(3-1)+2(2-1)+2(2-1)
D=90/6+6+2+2
D= 90/16
D=5.62

Conclusion:
Sample B is more diverse than Sample A

Kim et al. (2017) J. Microbiol. Biotechnol. 27, 2089–2093

Alpha diversity

- **Chao1-** considers rare species

$$S_1 = S_{obs} + \frac{F_1^2}{2F_2}$$

where F_1 and F_2 are the count of singletons and doubletons, respectively, and S is the number of observed species.

Kim et al. (2017) J. Microbiol. Biotechnol. 27, 2089–2093

Beta diversity

Measure of overall change between samples

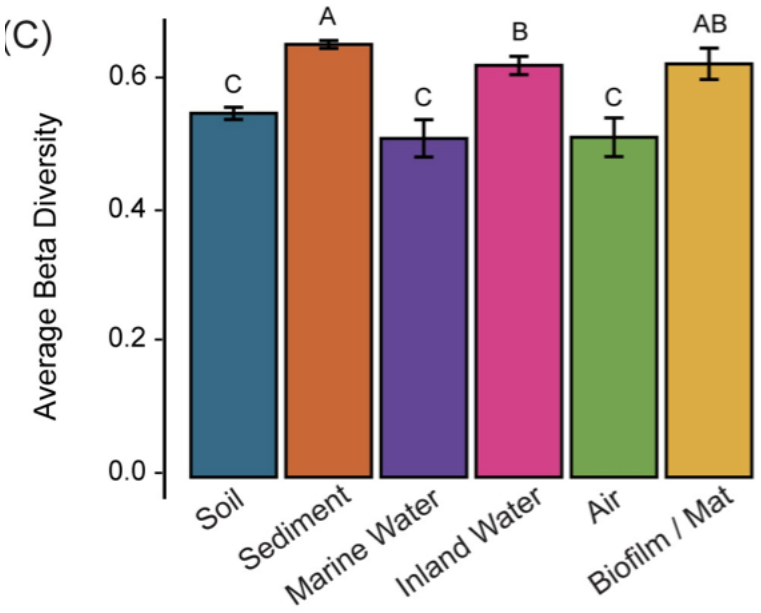


Fig. 7. Beta-diversity pattern of microbial communities from different environments.

Walters and Martiny (2020) PLoS ONE 15(9): e0233872

Beta diversity

- **Bray-Curtis dissimilarity-**based on abundance or read count data

$$BC_{ij} = 1 - \frac{2C_{ij}}{S_i + S_j}$$

i & j are the two samples,
S_i is the total number of all individuals counted on Sample A,
S_j is the total number of all individuals counted on Sample B,
C_{ij} is the sum of only the lesser counts for each species found in both sites.

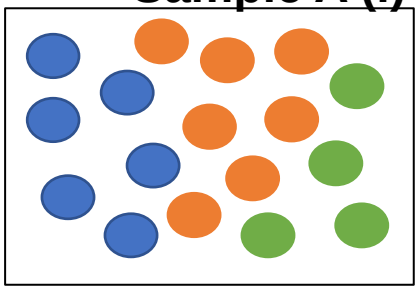
values are from 0 to 1
0 means both samples share the same species at exact the same abundances
1 means both samples have complete different species abundances

Beta diversity

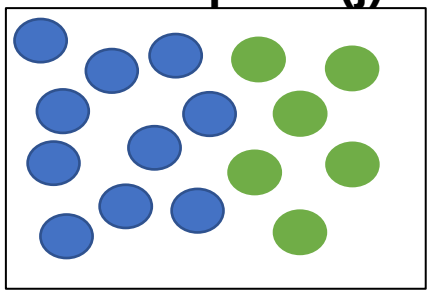
- **Bray-Curtis dissimilarity-**based on abundance or read count data

$$BC_{ij} = 1 - \frac{2C_{ij}}{S_i + S_j}$$

Sample A (i)



Sample B (j)



C_{ij} = 6+4 =10

Si= 6+7+4=17

Sj 10+6=16

BC_{ij} = 1 – (2 * 10) / (17 + 16),
=1 – 0.61
= 0.39

Conclusion:
Samples A and B are 39% dissimilar to each other

Beta diversity

- **Jaccard index-**based on presence or absence of species (does not include abundance information)

Jaccard Index = (the number of species in both sets) / (the number species in either set) * 100

or

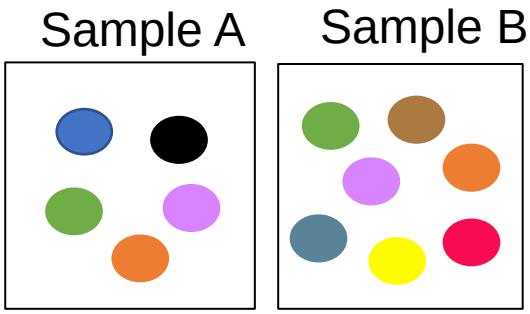
J(A,B) = |A∩B| / |A∪B|

values are from 0% to 100%
0% means both samples share exact the same species
1 00% means both samples have no species in common

Beta diversity

- **Jaccard index**-based on presence or absence of species (does not include abundance information)

J(A,B) = |A∩B| / |A∪B| *100



Solution: J(A,B) = |A∩B| / |A∪B|

= |  |/

| |

= 3/9

= 0.33*100

=33%

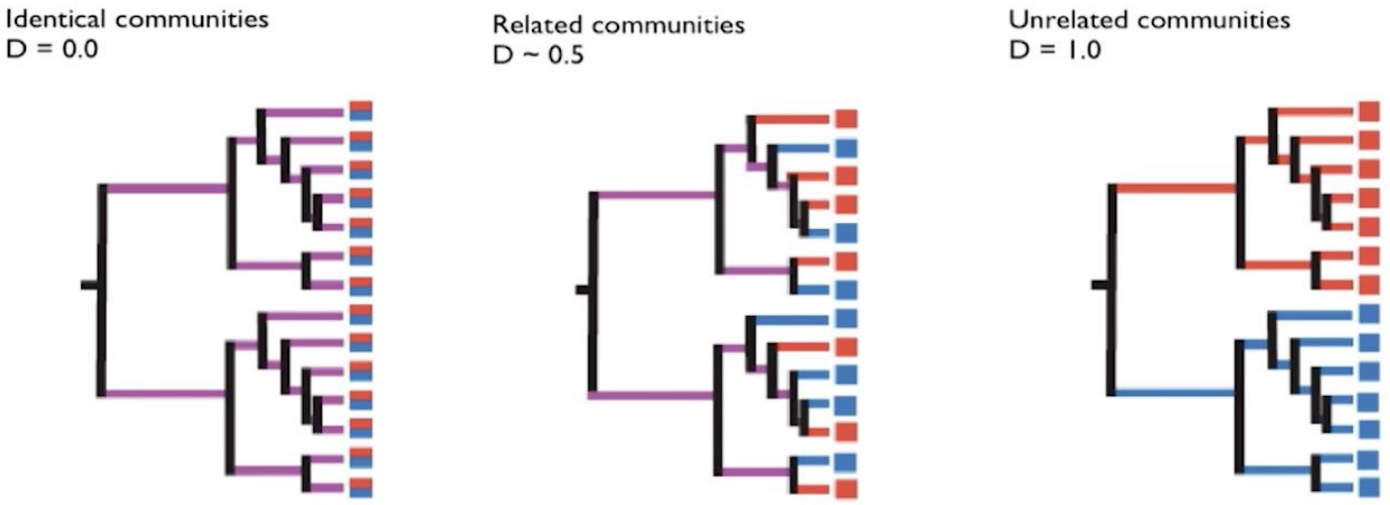
Conclusion:

Samples A and B are 33% similar.

Beta diversity

- **UniFrac-** sequence distances (phylogenetic tree)

  -based on the fraction of branch length that is shared between two samples or unique to one or the other sample



weighted UniFrac: branch lengths are weighted by relative abundances (includes both sequence and abundance information); emphasizes the dominant species

unweighted UniFrac: purely based on sequence distances (does not include abundance information); emphasizes the minor species

Ordination

-summarizes community data by producing low-dimensional ordination space; similar species and samples are plotted close together, and dissimilar species and samples are placed far apart

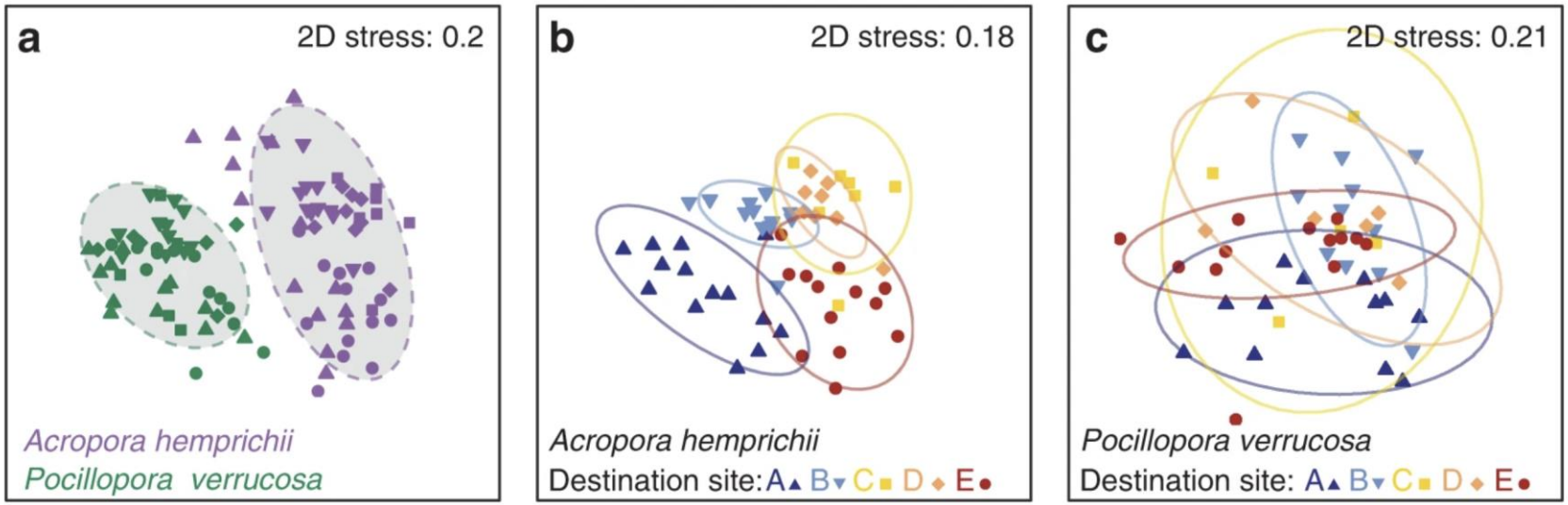
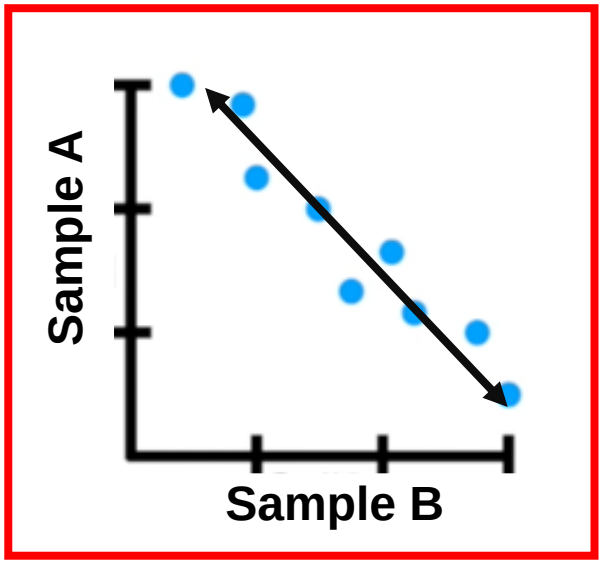


Fig. 8. Bacterial community structure and relative dispersion of the coral species *A*. *hemprichii* and *P*. *verrucosa.*

Ziegler et al. (2019) *Nat Commun* **10,** 3092

Ordination

- **Principal component analysis (PCA)-** converts the correlations (or lack thereof) among all the samples into 2-D graph

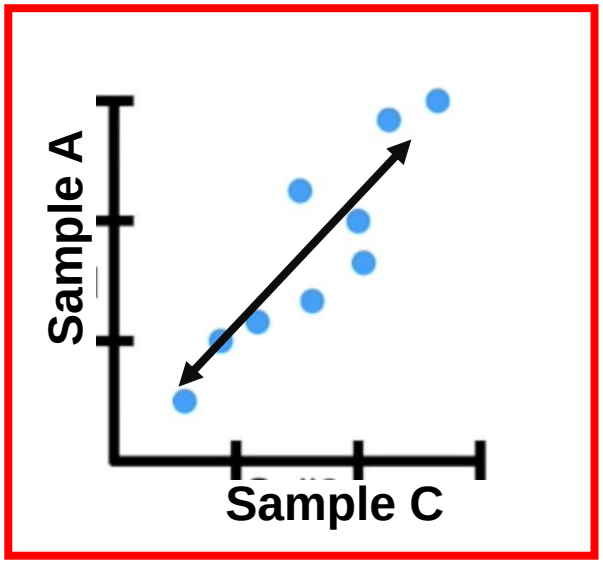


	Sample A	Sample B	Sample C
Species 1	3	0.25	2.8
Species 2	2.9	0.8	2.2
Species 3	2.2	1	1.5
Species 4	2	1.4	2
Species 5	1.3	1.6	1.6
Species 6	1.5	2	2.1
Species 7	1.1	2.2	1.2
Species 8	1	2.7	0.9
Species 9	0.4	3	0.6

Inversely correlated= dissimilar samples

Ordination

- **Principal component analysis (PCA)-** converts the correlations (or lack thereof) among all the samples into 2-D graph



positively correlated= similar samples

	Sample A	Sample B	Sample C
Species 1	3	0.25	2.8
Species 2	2.9	0.8	2.2
Species 3	2.2	1	1.5
Species 4	2	1.4	2
Species 5	1.3	1.6	1.6
Species 6	1.5	2	2.1
Species 7	1.1	2.2	1.2
Species 8	1	2.7	0.9
Species 9	0.4	3	0.6

Ordination

- **Principal coordinate analysis (PCoA)-** converts distances among all the samples into 2-D graph

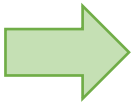
$$BC_{ij} = 1 - \frac{2C_{ij}}{S_i + S_j}$$

	Sample A	Sample B	Sample C
Species 1	3	0.25	2.8
Species 2	2.9	0.8	2.2
Species 3	2.2	1	1.5
Species 4	2	1.4	2
Species 5	1.3	1.6	1.6
Species 6	1.5	2	2.1
Species 7	1.1	2.2	1.2
Species 8	1	2.7	0.9
Species 9	0.4	3	0.6

Ordination

- **Non metric multidimensional analysis (nMDS)-** relies on rank orders (distances) for ordination to identify similarity in a data set

Raw data	Sp1	Sp2	Sp3	Sp4	Sp5
Sample A	3	5	7	1	2
Sample B	4	4	10	2	2



Sp1	Sp2	Sp3	Sp4	Sp5
3 =5	5 =8	7 =9	1 =1	2 =2
4 =6	4 =7	10 =10	2 =3	2 =4

FOR "2":
2+3+4=9
9/3= 3

FOR "4":
6+7=13
13/2= 6.5

Ranked data	Sp1	Sp2	Sp3	Sp4	Sp5
Sample A	5.00	8.00	9.00	1.00	3.00
Sample B	6.50	6.50	10.00	3.00	3.00

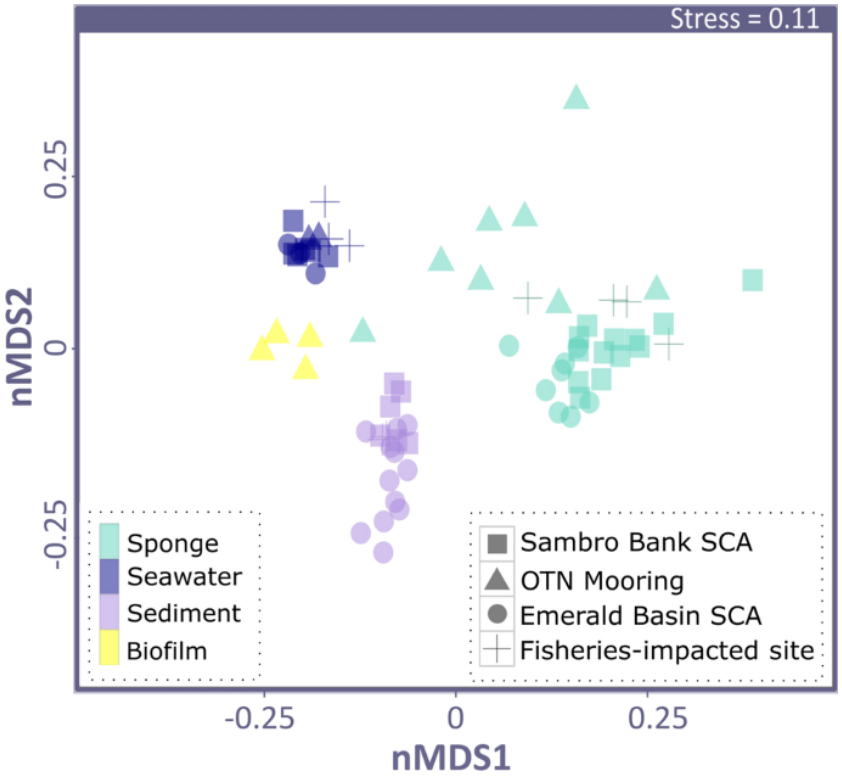
Busch et al. (2020) Conservation Genetics 21, 1001–1010

Ordination

- **Non metric multidimensional analysis (nMDS)-** relies on rank orders (distances) for ordination to identify similarity in a data set

Stress – indicator of goodness-of-fit

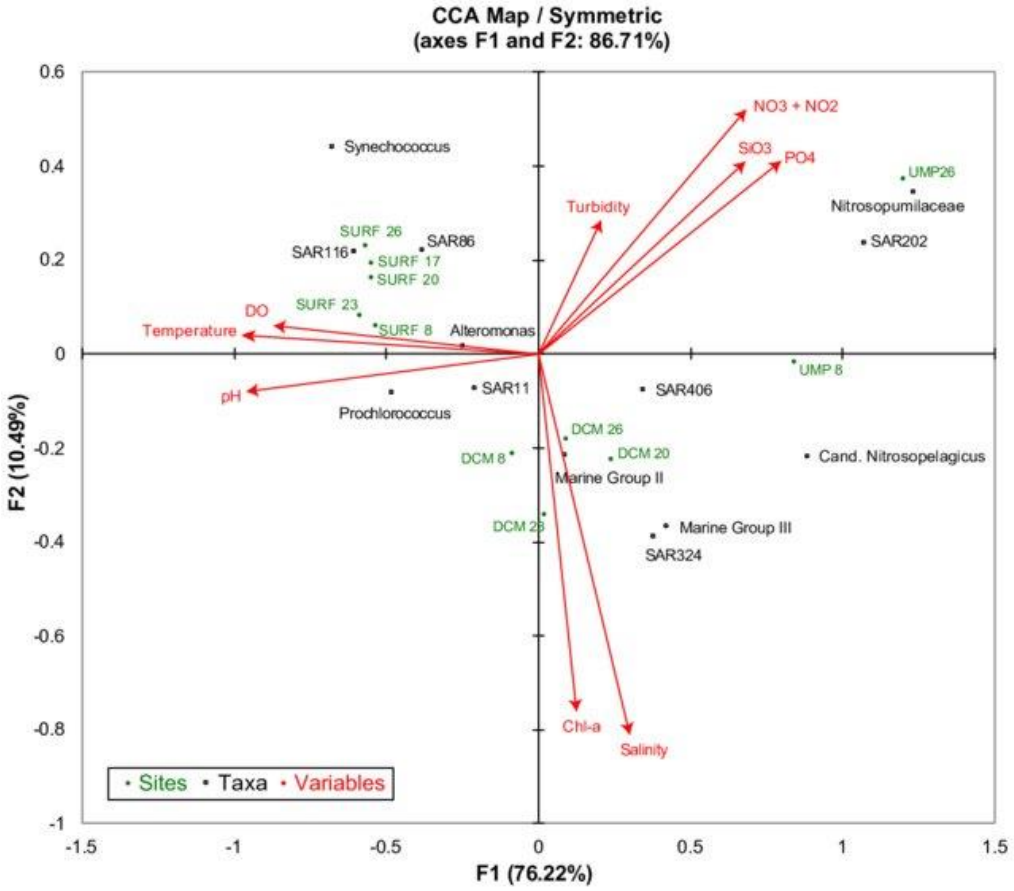
Stress values:
>0.2 are generally poor
<0.1 are good
<0.05 are excellent



Busch et al. (2020) Conservation Genetics 21, 1001–1010

Ordination

- **Canonical Correspondece Analysis (CCA)**-directly relates species to environmental variables



Gajigan et al. (2018) PeerJ 6(3)

Other Multivariate statistical tools

- **Multivariate analysis of variance with permutation (PERMANOVA)** - the sum of squared differences between points and their centroid is equal to the sum of the squared interpoint distances divided by the number of points.

  -based on distance matrices and permutation

  -R-squared = variations between samples (0-1; 0=similar, 1=highly variable)

  -F-ratio= indicates group separation (higher the value=more pronounced separation)

  -p value = tells you whether or not this result was likely a result of chance

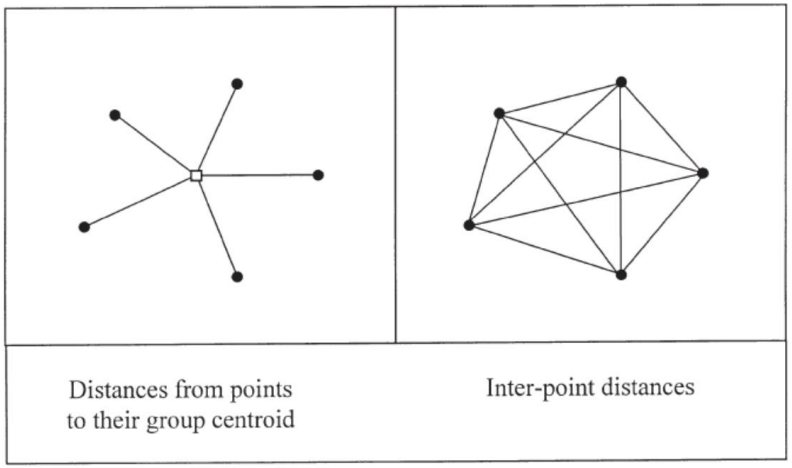


Fig. 2. The sum of squared distances from individual points to their centroid is equal to the sum of squared inter-point distances divided by the number of points.

Other Multivariate statistical tools

Multivariate analysis of dispersion (PERMDISP)- involves calculating the distance from each data point to its group centroid and then testing whether those distances differ among the groups.

-p value = tells you whether or not this result was likely due to differences in location or dispersion

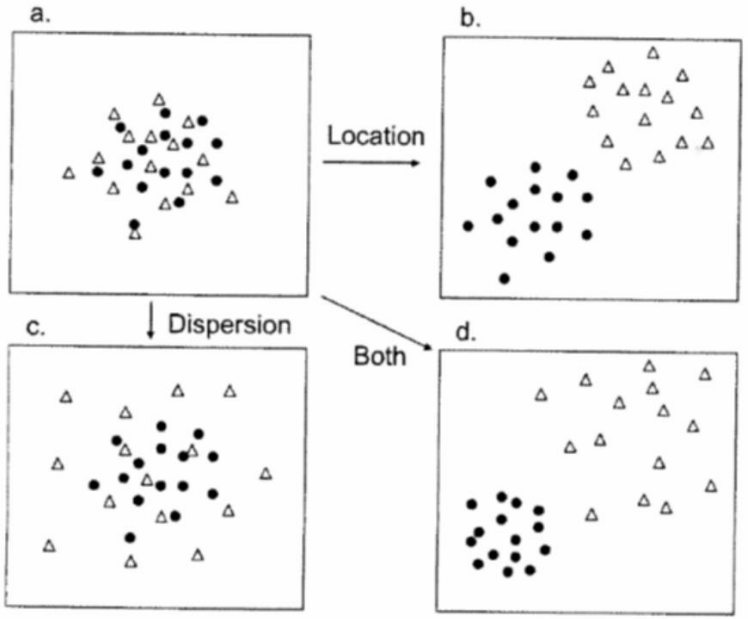


Fig. 2.1. Schematic diagram showing two groups of samples in a bivariate system (two dimensions) that (a) do not differ in either location or dispersion, (b) differ only in their location in multivariate space, (c) differ only in their relative dispersions and (d) differ in both their location and in their relative dispersion.

Other Multivariate statistical tools

- **Analysis of group similarities (ANOSIM)-** used to the mean of ranked dissimilarities between groups to the mean of ranked dissimilarities within groups

  -R=variations between samples (0-1; 0=similar, 1=dissimilar)

  -p value = tells you whether or not this result was likely a result of chance

Helpful resources

- GUide to STatistical Analysis in Microbial Ecology (GUSTA ME)! (https://mb3is.megx.net/gustame)
- https://rachaellappan.github.io/VL-QIIME2-analysis/pre-processing-of-sequence-reads.html
- https://qiime2.org

Demo

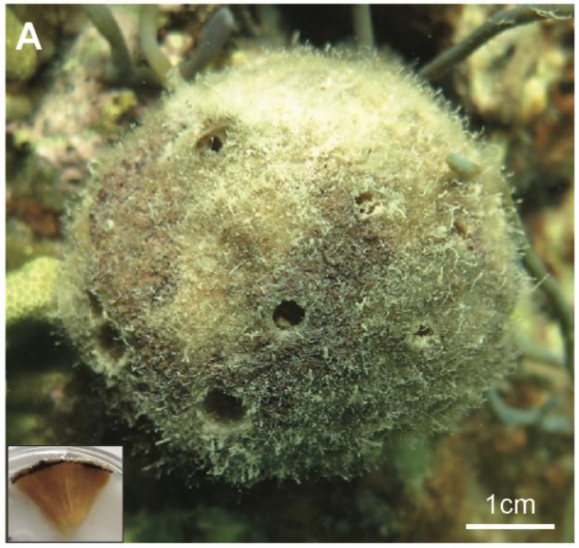
- **Rarefaction curve**
- **Relative abundance plot**
- **Alpha diversity**
- **Beta diversity**
- **Ordination**

Demo

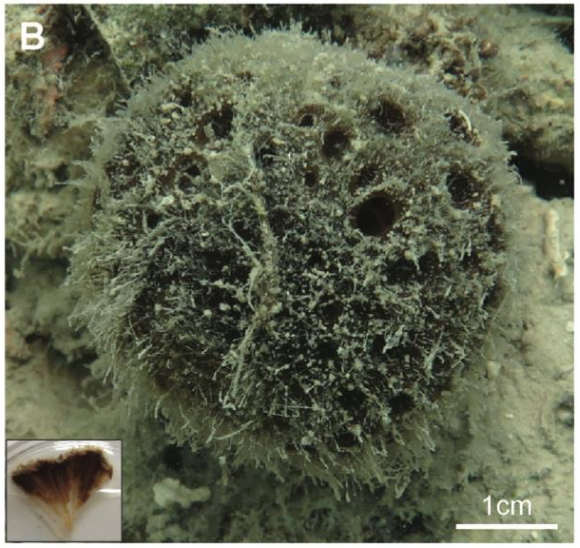
- **Details**

- Microbiome of ball sponges in Bolinao Samples:
  - Cinachyrella sp. (n=3)
  - Paratetilla (n=3)

- 16S rRNA V34 region sequenced



Cinachyrella sp.



Paratetilla sp.

Objective: Examine associated prokaryotic microbial communities of two species of ball sponges from two variable environmental conditions

References

- Abellan-Schneyder I, Matchado MS, Reitmeier S, Sommer A, Sewald Z, Baumbach J, List M, Neuhaus K. 2021. Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. mSphere 6:e01202-20.
- Allali A, Arnold JW, Roach, J, Cadenas MB, Butz, N,Hasaan M, Koci M, Ballou, A, Mendoza M, Ali RizwN, Azacarate-Peril MA. 2017. A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC Microbiology* 17:194.
- Nearing JT, Douglas GM, Comeau AM, Langille MGI. 2018. Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches. *PeerJ*6:e5364
- Kim, B., Shin, J., Guevarra, R. B., Lee, J. H., Kim, D. W., Seol, K., Lee, J., Kim, H. B., & Isaacson, R. E. 2017. Deciphering diversity indices for a better understanding of microbial communities. *Journal of Microbiology and Biotechnology*, *27*(12), 2089-2093.
- Pollock, F.J., McMinds, R., Smith, S. *et al.* 2018. Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nat Commun* **9,** 4921.
- Pollock, J., Glendinning, L., Wisedchanwet, T., & Watson, M. 2018. The madness of microbiome: Attempting to find consensus "Best practice" for 16S microbiome studies. *Applied and Environmental Microbiology*, *84*(7).
- Thomas, T., Moitinho-Silva, L., Lurgi, M. *et al.* 2016. Diversity, structure and convergent evolution of the global sponge microbiome. *Nat Commun* **7,** 11870.
- Wear EK, Wilbanks EG, Nelson CE, Carlson CA. 2018. Primer selection impacts specific population abundances but not community dynamics in a monthly time-series 16S rRNA gene amplicon analysis of coastal marine bacterioplankton. *Environmental Microbiology*, 20, 2709–2726.
- Yang B, Wang Y, Qian P-Y. 2016. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. BMC Bioinformatics 17:135.
- Walters KE, Martiny JBH .2020. Alpha-, beta-, and gamma-diversity of bacteria varies across habitats. PLoS ONE 15(9): e0233872.