

Metabarcoding



Who am I?

**Anthropogenic
biodiversity change**

**Community Ecology &
Biogeography**

**Methods in sedaDNA
and eDNA**

**Metabarcoding &
metagenomics**



Globe Institute

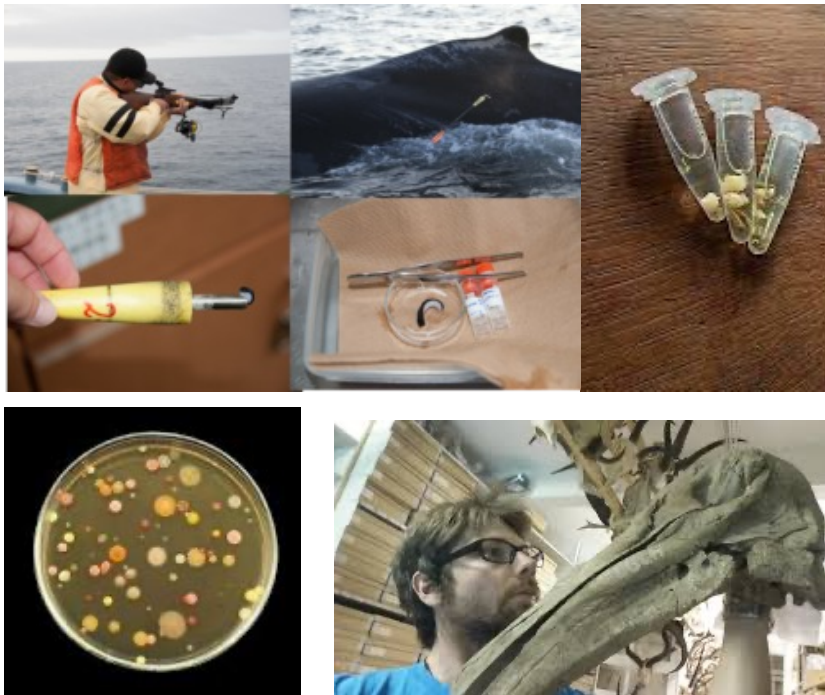


Today

- Sample types
- What is metabarcoding?
- Sequence detective
- Metabarcoding libraries
- Bioinformatics overview

Next Generation Sequencing Sample Sources

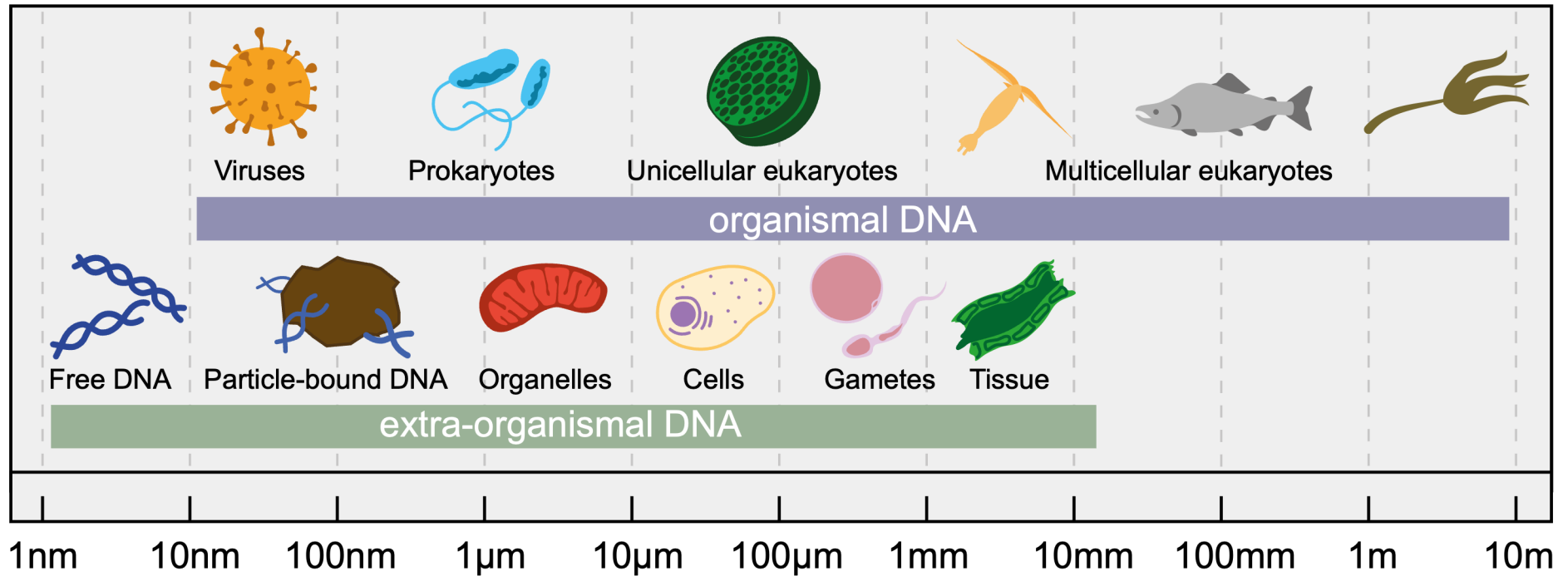
Point Source 



Mixed Source 



a)



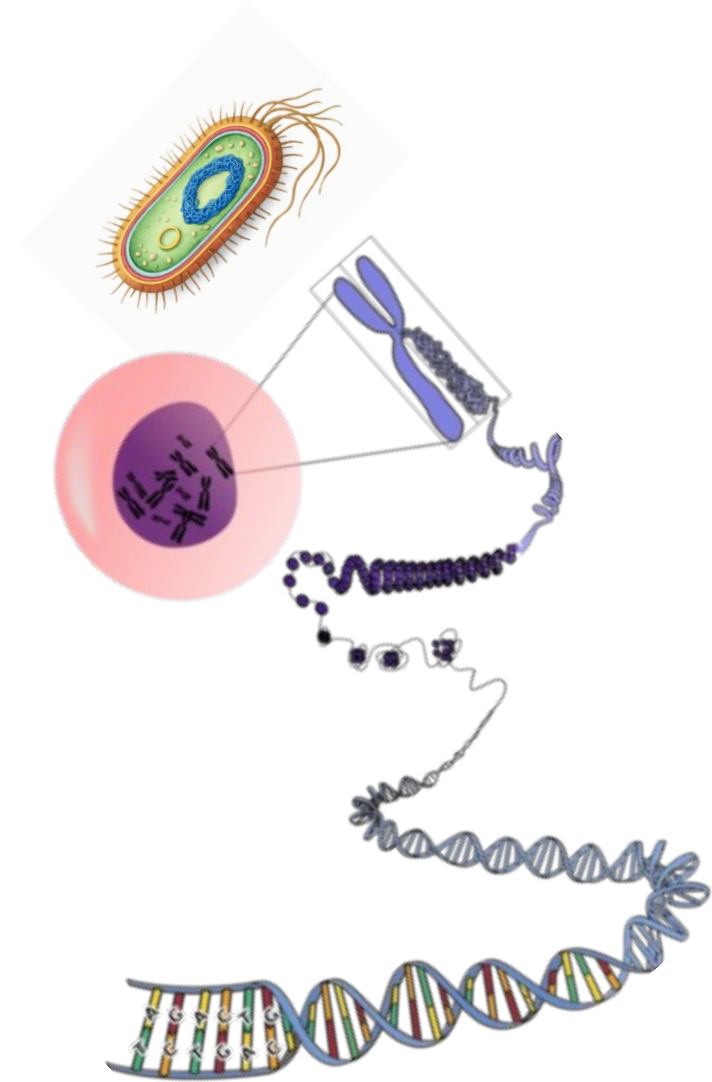
b)



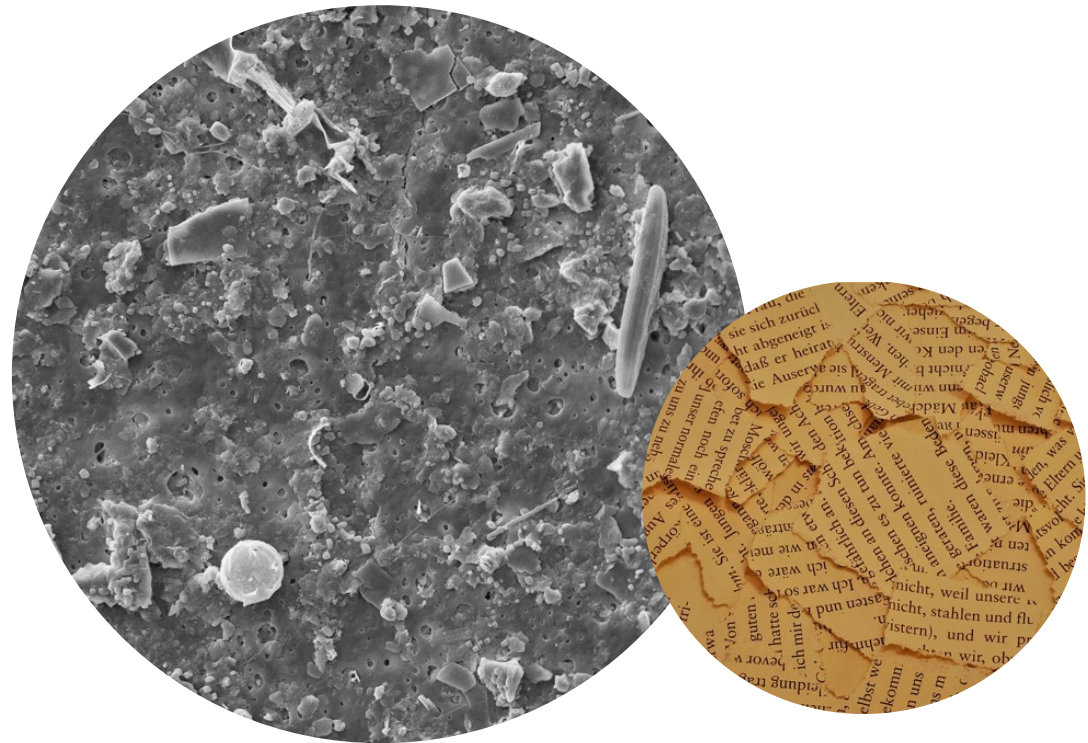




What's in these mixed samples?



What's in these mixed samples?



Animal genomes

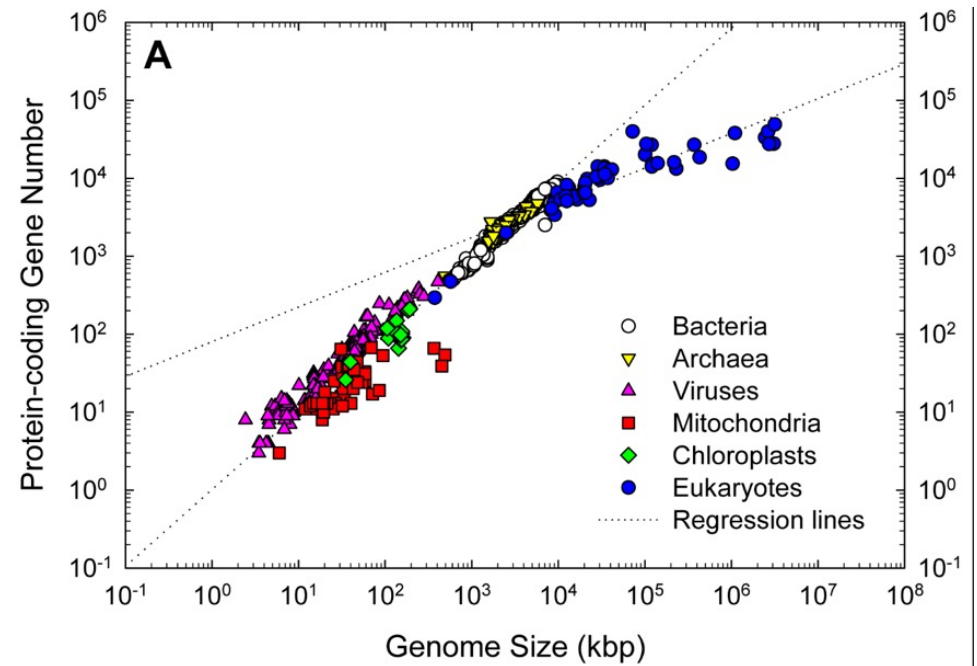
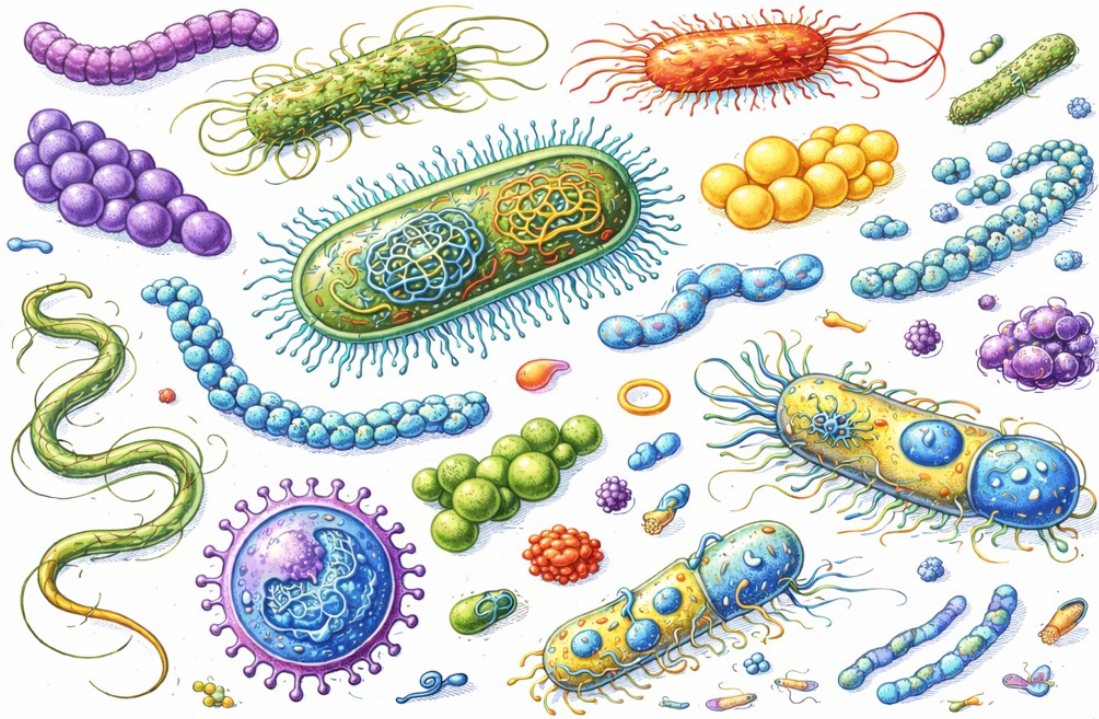
[Mock et al. 2022](#)

A collage of 15 detailed illustrations of various microorganisms. The organisms include: a green, elongated bacterium with a red flagellum; a blue, oval-shaped bacterium with a large blue nucleus; a green, spherical microorganism with many small green dots; a purple, star-shaped microorganism; a green, multi-lobed microorganism with a large blue nucleus; a red, branching, coral-like structure; a brown, oval-shaped microorganism with many small brown dots; a green, oval-shaped microorganism with many small green dots; a green, multi-lobed microorganism with a large blue nucleus; a blue, oval-shaped microorganism with a large blue nucleus; a green, multi-lobed microorganism with a large blue nucleus; a green, multi-lobed microorganism with a large blue nucleus; a green, multi-lobed microorganism with a large blue nucleus; a green, multi-lobed microorganism with a large blue nucleus; and a green, multi-lobed microorganism with a large blue nucleus.



Mock et al. 2022

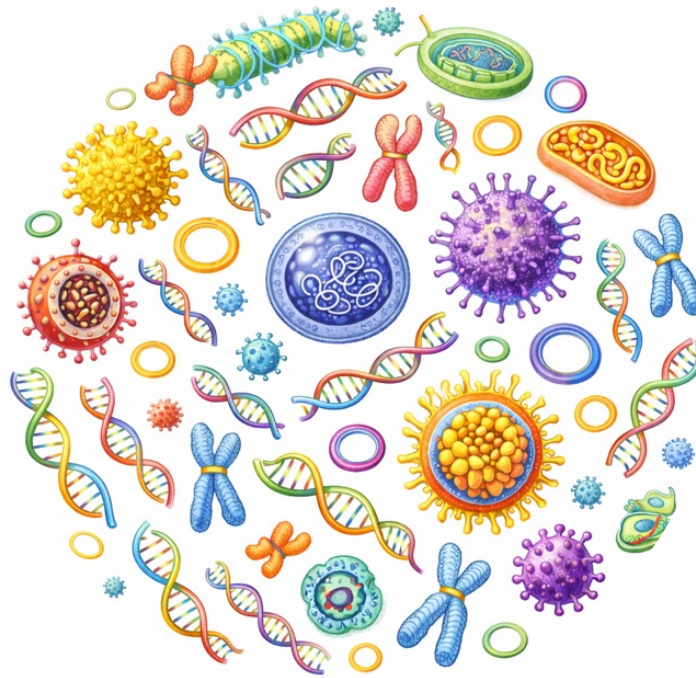
What's in these mixed samples?



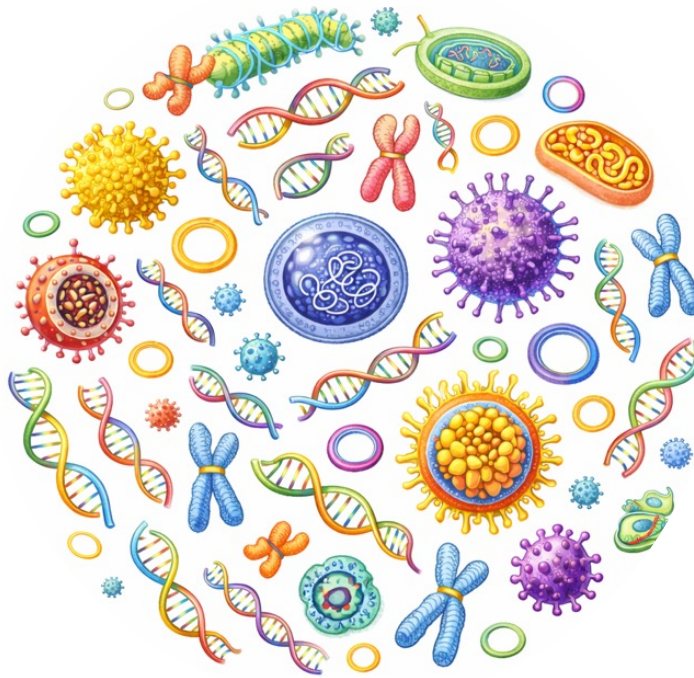
Bacteria & Viruses

Hou & Lin 2009

What taxa are in this sample?

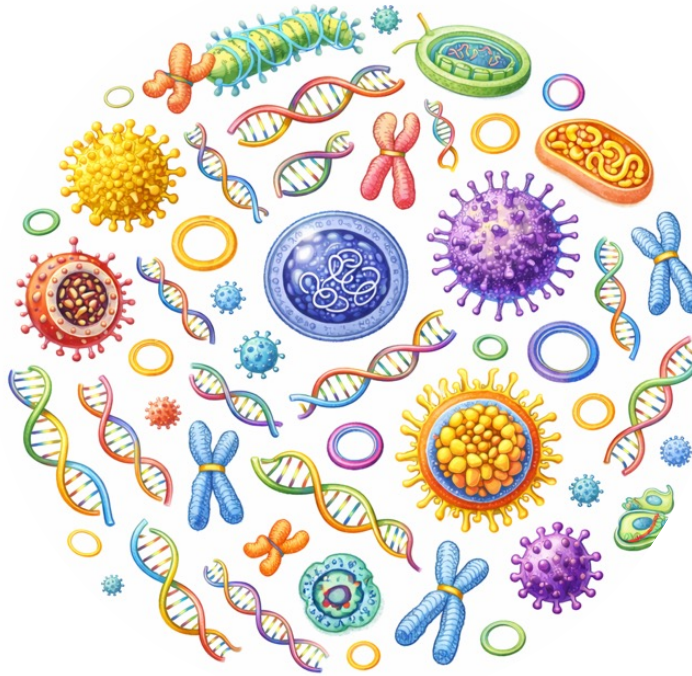


What taxa are in this sample?



So... what shall we do?

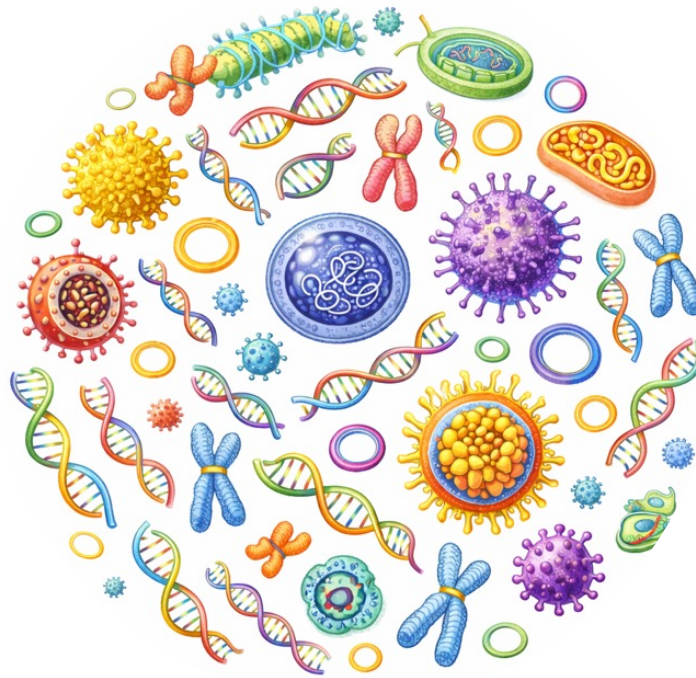
What taxa are in this sample?



So... what shall we do?



What taxa are in this sample?

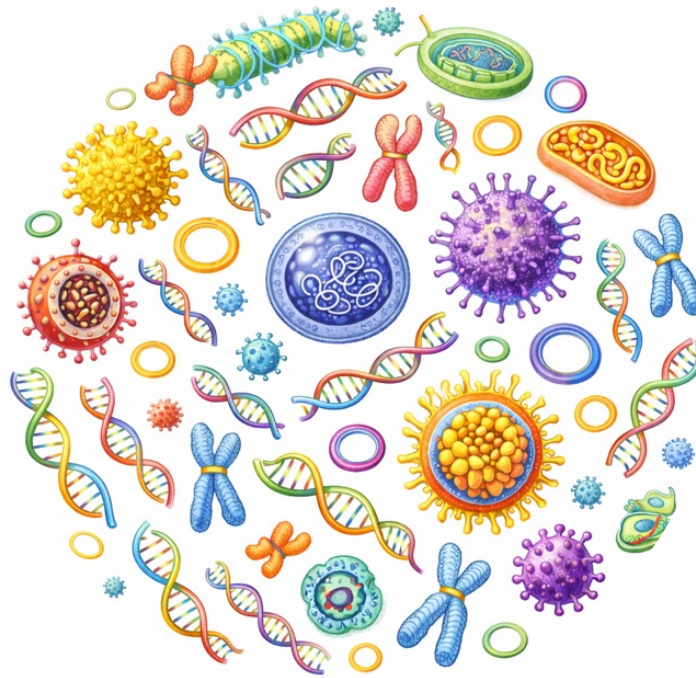


Metagenomics

So... what shall we do?



So... what shall we do?

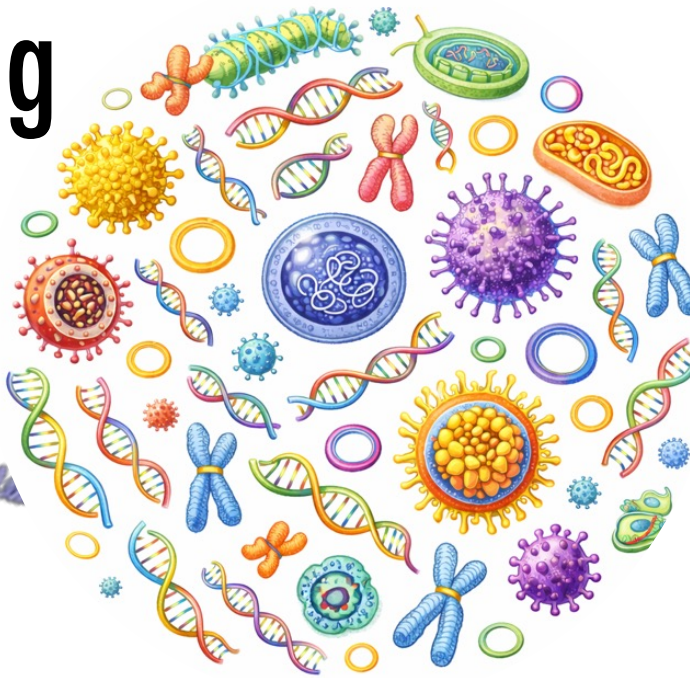
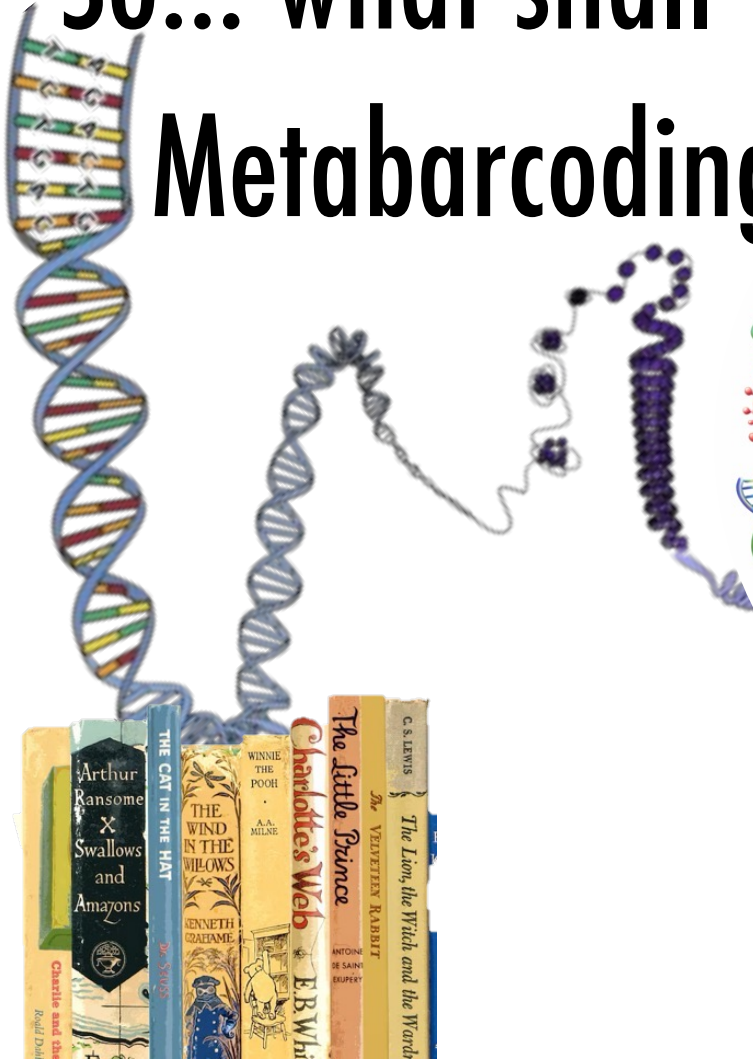


Metagenomics



So... what shall we do?

Metabarcoding



Metagenomics



DNA Barcoding



Ciona intestinalis



Arcuatula senhousia



Oncorhynchus gorbuscha

Molecular barcode – a short (>600bp) DNA sequence, from a uniform locality in the genome used for species identification

THE ROYAL SOCIETY

Received 29 July 2002
Accepted 30 September 2002
Published online 8 January 2003

Biological identifications through DNA barcodes

Paul D. N. Hebert*, Alina Cywinska, Shelley L. Ball
and Jeremy R. deWaard

Department of Zoology, University of Guelph, Guelph, Ontario N1G 2W1, Canada

Although much biological research depends upon species diagnoses, taxonomic expertise is collapsing. We are convinced that the sole prospect for a sustainable identification capability lies in the construction of systems that employ DNA sequences as taxon 'barcodes'. We establish that the mitochondrial gene cytochrome *c* oxidase I (COI) can serve as the core of a global bioidentification system for animals. First, we demonstrate that COI profiles, derived from the low-density sampling of higher taxonomic categories, ordinarily assign newly analysed taxa to the appropriate phylum or order. Second, we demonstrate that species-level assignments can be obtained by creating comprehensive COI profiles. A model COI profile, based upon the analysis of a single individual from each of 200 closely allied species of lepidopterans, was 100% successful in correctly identifying subsequent specimens. When fully developed, a COI identification system will provide a reliable, cost-effective and accessible solution to the current problem of species identification. Its assembly will also generate important new insights into the diversification of life and the rules of molecular evolution.

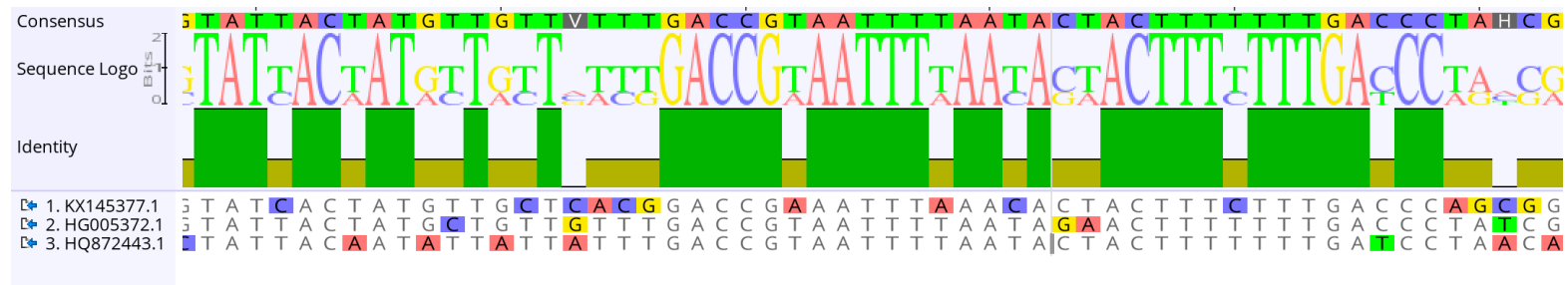
Keywords: molecular taxonomy; mitochondrial DNA; animals; insects; sequence diversity; evolution

1. INTRODUCTION

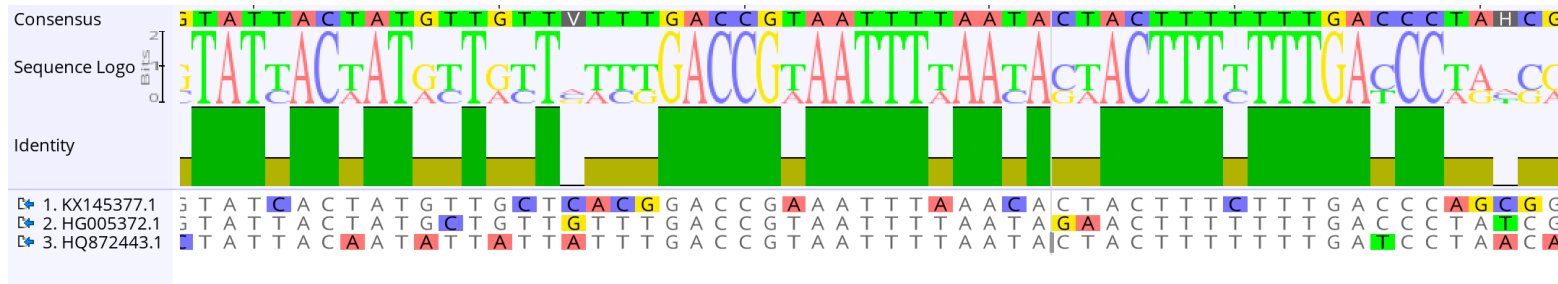
The diversity of life underpins all biological studies, but it is also a harsh burden. Whereas physicists deal with a cosmos assembled from 12 fundamental particles, biologists confront a planet populated by millions of species. Their discrimination is no easy task. In fact, since few taxonomists can critically identify more than 0.01% of the

Allander *et al.* 2001; Hamels *et al.* 2001). However, the problems inherent in morphological taxonomy are general enough to merit the extension of this approach to all life. In fact, there are a growing number of cases in which DNA-based identification systems have been applied to higher organisms (Brown *et al.* 1999; Bucklin *et al.* 1999; Treweek 2000; Vincent *et al.* 2000).

DNA Barcoding



DNA Barcoding



Metazoan – COI



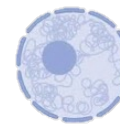
Herbert et al. 2003

Flowering Plants – rbcL



Hollingsworth et al. 2009

Fungi – ITS



Schoch et al. 2011

Eukaryotes – 18S



Amaral-Zettler et al. 2009

DNA Barcoding

Barcoding Projects



Barcode 500k 2010-2015

BioSCAN 2019-2025

<http://www.boldsystems.org>

<http://ibol.org/projects/project-overview/>



Genome Projects



25 High Quality Genomes

<https://www.sanger.ac.uk/science/collaboration/25-genomes-25-years>



Sequence the 66,000 UK species



Sequence all Eukaryotic diversity
Around \$5 billion

<https://www.earthbiogenome.org>

DNA Barcoding

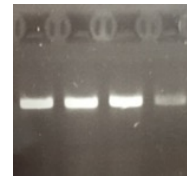
Meta

PCR

Single species sample



A single, clean amplification product



Sanger sequencing

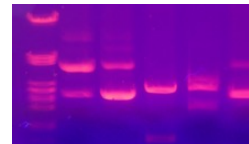
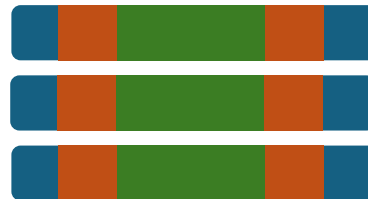


1 seq
(300–1000bp)

Mixed species sample



Multiple products from amplification

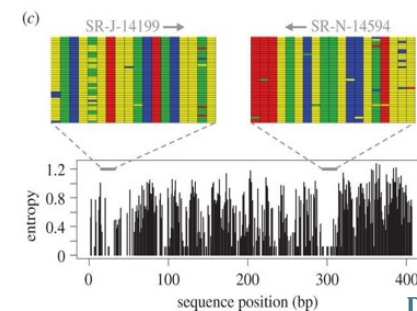
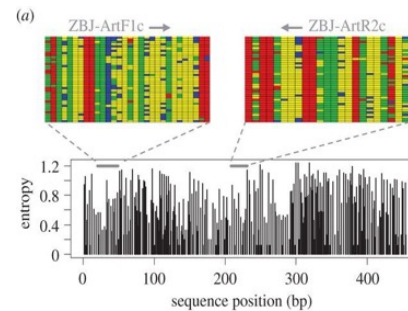


Next-generation Sequencing



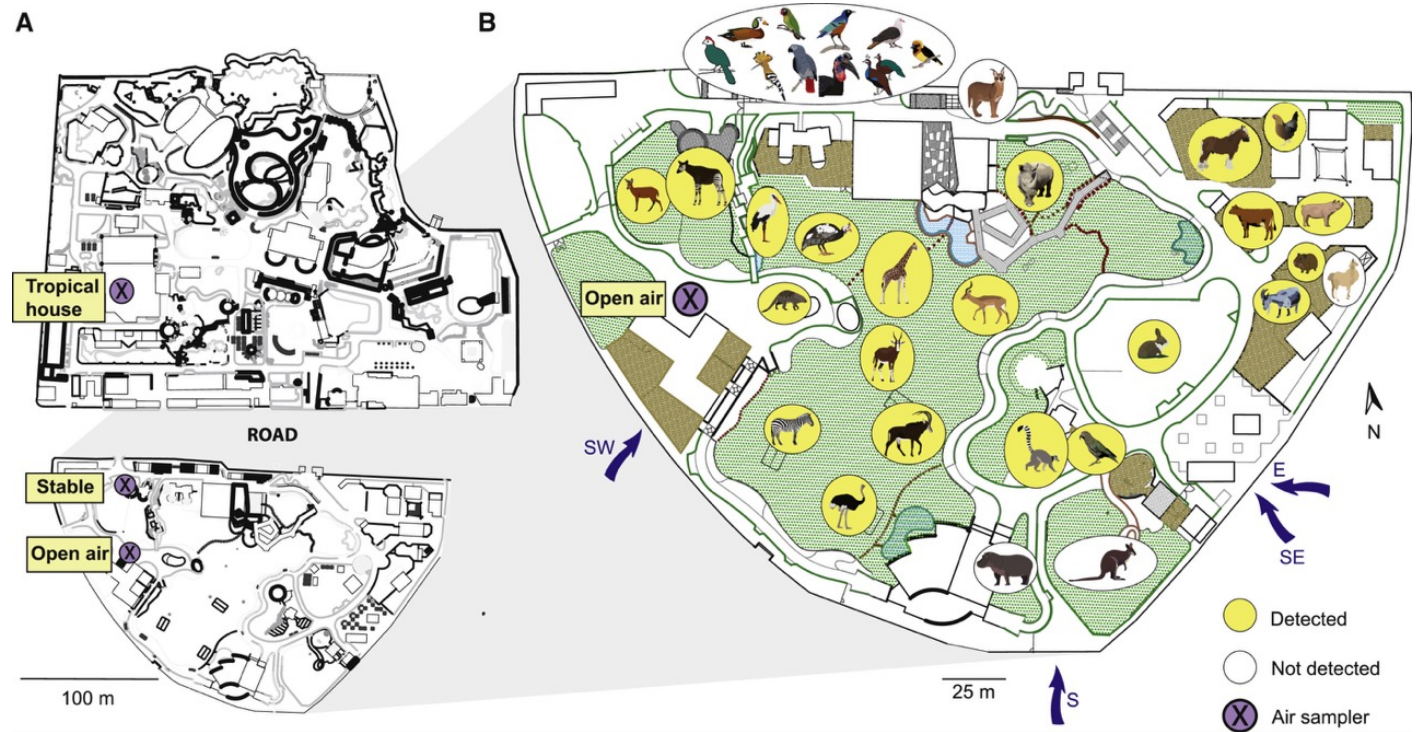
$10^6 - 10^8$ seq
(75–300bp)

Variable region
Conserved region



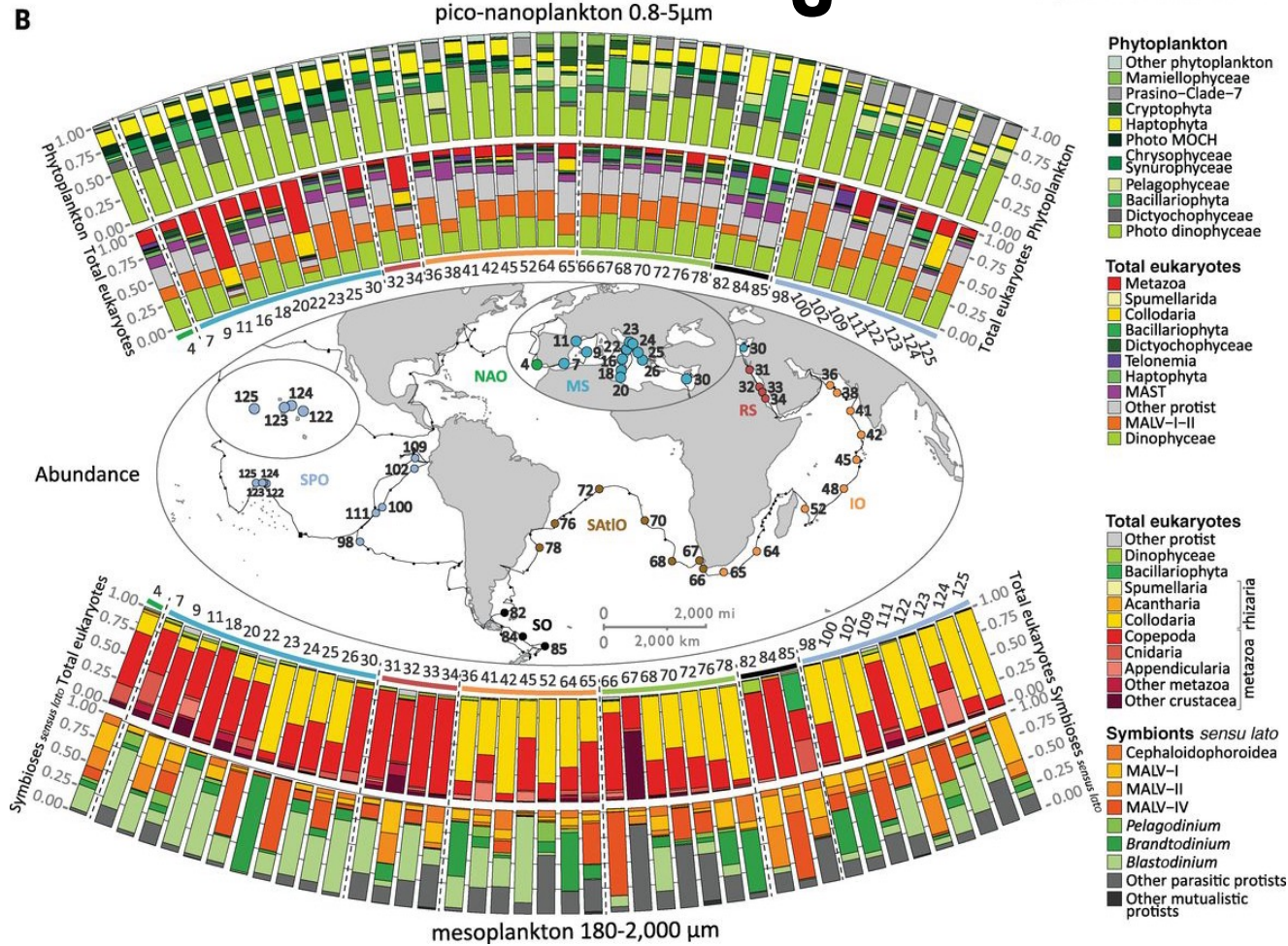
[Deagle et al. 2014](#)

DNA Metabarcoding Studies



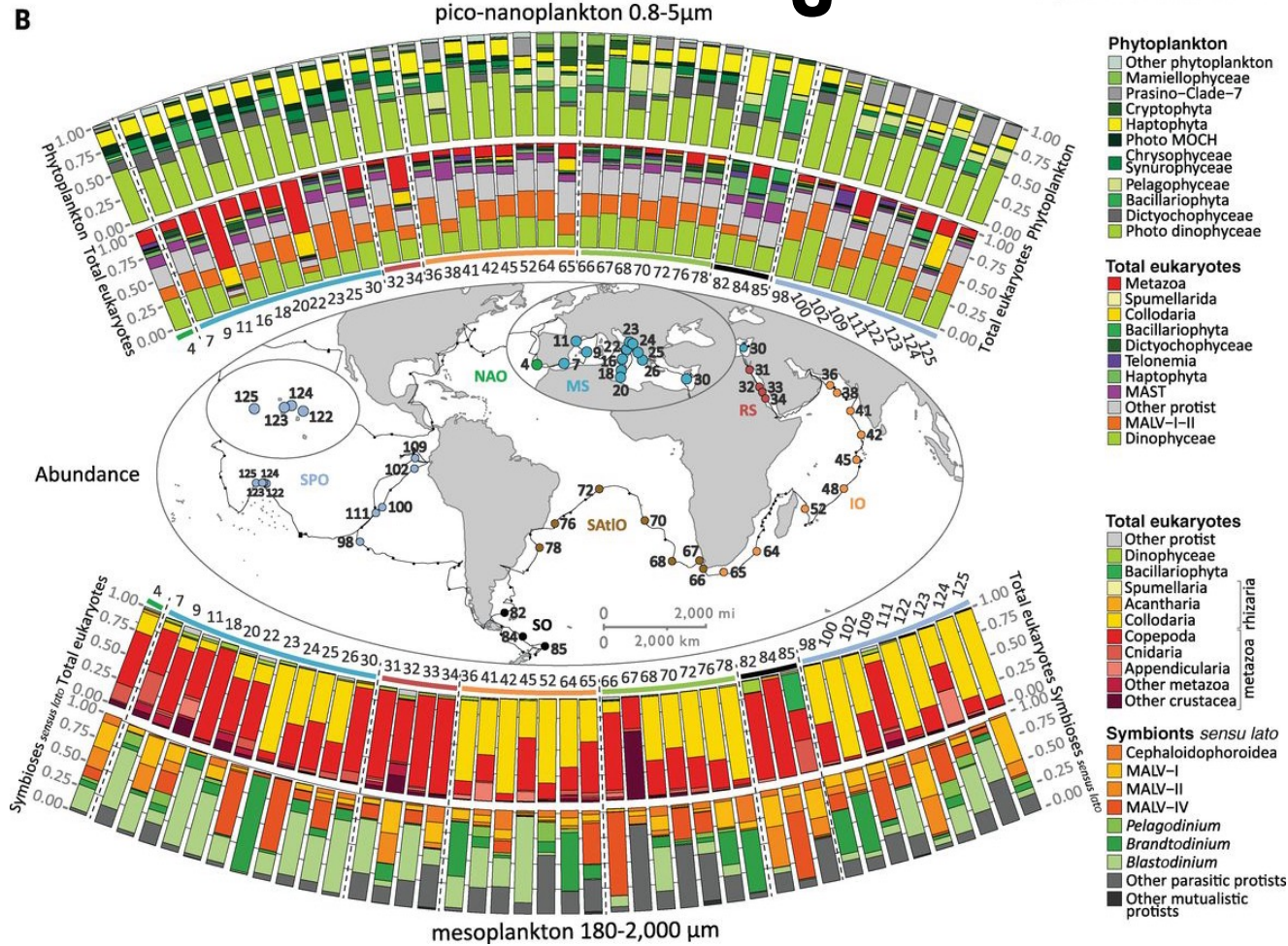
[Lynggaard et al. 2022](#)

DNA Metabarcoding Studies



De Vargas et al. 2015

DNA Metabarcoding Studies



De Vargas et al. 2015

Exercise 1 - Metabarcode detective



ACTATCTAGGAACATTGCACATGCCGGAAGGTCCGT
AGACTTTGCCATCTTTTCGTTACATTAGCCGGAGTTA
GTTCAATTCTTGGGGCAGTAAATTTATTAGGACTTTA
GGAAACCTTCGAGTATTTGGAATATTACTAGATCGCA
TACCTCTGTTTGCATGAGCTGTATTAATTACTGCTGTA
CTATTACTTCTTTCTTTACCTGTTCTAGCTGGGGCTATT
ACCATATTATTAACCGATCGTAATTTGAATACAACATT
TATGATGTGGGGGGAGGAGGAGATCCTATTCTATATC
AACACCTATTC

Exercise 1 - Metabarcoding detective

ACTATCTAGGAACATTGCACATGCCGGAAGGTCCGT
AGACTTTGCCATCTTTTCGTTACATTTAGCCGGAGTTA
GTTCAATTCTTGGGGCAGTAAATTTTATTAGGACTTTA
GGAAACCTTCGAGTATTTGGAATATTACTAGATCGCA
TACCTCTGTTTGCATGAGCTGTATTAATTACTGCTGTA
CTATTACTTCTTTCTTTACCTGTTCTAGCTGGGGCTATT
ACCATATTATTAACCGATCGTAATTTGAATACAACATT
TATGATGTGGGGGGGAGGAGGAGATCCTATTCTATATC
AACACCTATTC

 **National Library of Medicine**
National Center for Biotechnology Information

Log in

BLAST®

Home Recent Results Saved Strategies Help

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance.

[Learn more](#)

NEWS

Mon, 21 Jul 2025
Here are a few highlights in our latest BLAST+ release:
Download BLAST+ 2.17.0 now! [More BLAST news...](#)

Web BLAST

**Nucleotide BLAST**
nucleotide ► nucleotide

blastx
translated nucleotide ► protein

tblastn
protein ► translated nucleotide

**Protein BLAST**
protein ► protein

BLAST Genomes

Search

Human Mouse Rat Microbes

Exercise 1 - Metabarcoding detective

ACTATCTAGGAACATTGCACATGCCGGAAGGTCCGT
AGACTTTGCCATCTTTTCGTTACATTTAGCCGGAGTTA
GTTCAATTCTTGGGGCAGTAAATTTTATTAGGACTTTA
GGAAACCTTCGAGTATTTGGAATATTACTAGATCGCA
TACCTCTGTTTGCATGAGCTGTATTAATTACTGCTGTA
CTATTACTTCTTTCTTTACCTGTTCTAGCTGGGGCTATT
ACCATATTATTAACCGATCGTAATTTGAATACAACATT
TATGATGTGGGGGGGAGGAGGAGATCCTATTCTATATC
AACACCTATTC

NIH National Library of Medicine
National Center for Biotechnology Information

BLAST® Home Recent Results Saved Strategies Help

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

Mon, 21 Jul 2025
Here are a few highlights in our latest BLAST+ release:
Download BLAST+ 2.17.0 now! [More BLAST news...](#)

Web BLAST



Nucleotide BLAST
nucleotide ► nucleotide

blastx
translated nucleotide ► protein

tblastn
protein ► translated nucleotide



Protein BLAST
protein ► protein

BLAST Genomes

Human Mouse Rat Microbes

Exercise 1 - Metabarcode detective

ACTATCTAGGAACATTGCACATGCCGGAAGGTCCGT
AGACTTTGCCATCTTTTCGTTACATTTAGCCGGAGTTA
GTTCAATTCTTGGGGCAGTAAATTTTATTAGGACTTTA
GGAAACCTTCGAGTATTTGGAATATTACTAGATCGCA
TACCTCTGTTTGCATGAGCTGTATTAATTACTGCTGTA
CTATTACTTCTTTCTTTACCTGTTCTAGCTGGGGCTATT
ACCATATTATTAACCGATCGTAATTTGAATACAACATTT
TATGATGTGGGGGGGAGGAGGAGATCCTATTCTATATC
AACACCTATTC

Standard Nucleotide BLAST

blastn blastp blastx tblastn tblastx

BLASTN programs search nucleotide databases using a nucleotide query, more...

Reset page
Bookmark

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [Clear](#) Query subrange [?](#)

TCTAGCTGGTAACCTAGCCACGCGGAGCATCC
GTTGATTAAACAATTTTTCTTCTCATCTAGCAGGT
ATTTCATCAATTTAGGCGCTATTATTTATTACCA
CTATTATTAAATAAAACCACTGCAATTTACAATA
TCAAACCCCACTATTGTTTGAGCCGTTAATAC
AGCCGTCTTCTTCTCTATCCCTACCGTCTTAG
CTGCTGGAATTACAATCTTCTGACAGACCGAAAC
CTAAACACTACCTTTTTGATCCTGCTGCGGTG
GAGATCCTATTCTTTACCAACATTTGTTTC

From
To

Or, upload file No file chosen [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

Choose Search Set

Database ☒ Standard databases (nr etc.): ☐ rRNA/ITS databases ☐ Genomic + transcript databases ☐ Betacoronavirus ☐ Experimental databases

Core nucleotide database (core_nt) [?](#)

Limit SRA by ☐ Run ☐ Experiment [Extended SRA lookup](#)

Organism ☐ exclude [Add Organism](#)

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

Exclude ☐ Models (XM/XP) ☐ Uncultured/environmental sample sequences

Limit to ☐ Sequences from type material

Entrez® Query [YouTube](#) [Create custom database](#)

Enter an Entrez query to limit search [?](#)

Program Selection

Optimize for ☒ Highly similar sequences (megablast)
☐ More dissimilar sequences (discontiguous megablast)
☐ Somewhat similar sequences (blastn)

Choose a BLAST algorithm [?](#)

BLAST

Search database core_nt using Megablast (Optimize for highly similar sequences)

☐ Show results in a new window

Exercise 1 - Metabarcode detective

ACTATCTAGGAACATTGCACATGCCGGAAGGTCCGT
AGACTTTGCCATCTTTTCGTTACATTTAGCCGGAGTTA
GTTCAATTCTTGGGGCAGTAAATTTTATTAGGACTTTA
GGAAACCTTCGAGTATTTGGAATATTACTAGATCGCA
TACCTCTGTTTGCATGAGCTGTATTAATTACTGCTGTA
CTATTACTTCTTTCTTTACCTGTTCTAGCTGGGGCTATT
ACCATATTATTAACCGATCGTAATTTGAATACAACATTT
TATGATGTGGGGGGGAGGAGGAGATCCTATTCTATATC
AACACCTATTC

Standard Nucleotide BLAST

blastn blastp blastx tblastn tblastx

BLASTN programs search nucleotide databases using a nucleotide query. more...

Reset page
Bookmark

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) Clear

Query subrange [?](#)

From

To

TCTAGCTGGTAACTTAGCCACGCGGAGCATCC
GTTGATTAAACAATTTTTCTTCTATCAGCAGGT
ATTTGATCAATTTAGGCGCTATTATTTATTACCA
CTATTATTAATATAAACCCACCTGCAATTCACAATA
TCAAACCCCACTATTGTTGAGCCGATTAAATAC
AGCCGTCCTTTCTCTTATCCCTACCGTCCTAG
CTGCTGGAATTACAATCTCTGACAGACCGAAAC
CTAAACACTACCTTTTTGATCCTGCTGGCGGTG
GAGATCCTATTCTTACCAACATTGTTC

Or, upload file No file chosen [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

Choose Search Set

Database ☒ Standard databases (nr etc.): ☐ rRNA/ITS databases ☐ Genomic + transcript databases ☐ Betacoronavirus ☐ Experimental databases

Core nucleotide database (core_nt) [?](#)

Limit SRA by ☐ Run ☐ Experiment [Extended SRA lookup](#)

Organism [?](#) ☐ exclude [Add Organism](#)

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

☐ Exclude [Optional](#) Models (XM/XP) ☐ Uncultured/environmental sample sequences

☐ Limit to [Optional](#) Sequences from type material

[YouTube](#) [Create custom database](#)

Enter an Entrez query to limit search [?](#)

Program Selection

Optimize for ☒ Highly similar sequences (megablast) ☐ More dissimilar sequences (discontiguous megablast) ☐ Somewhat similar sequences (blastn)

Choose a BLAST algorithm [?](#)

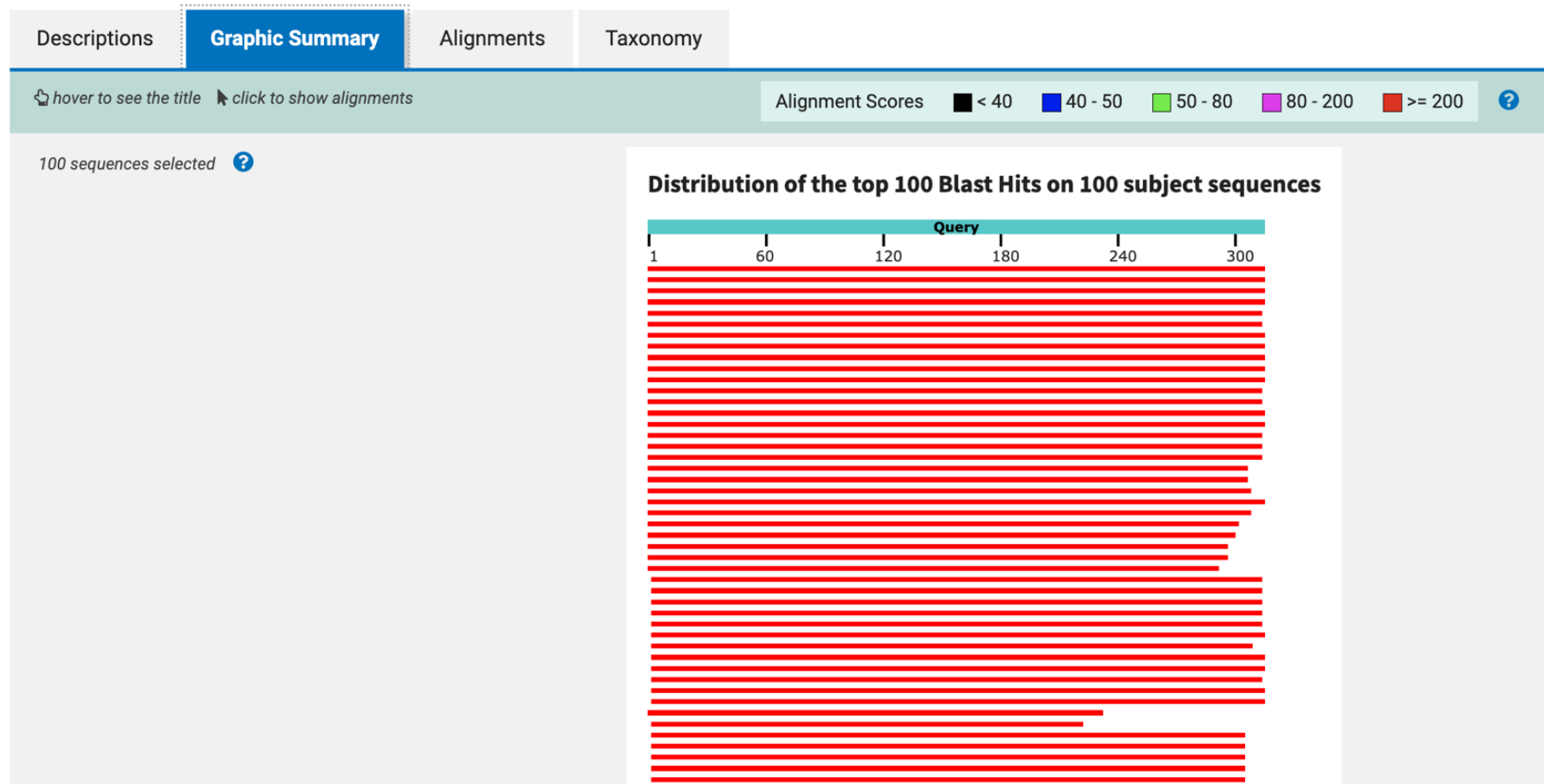
BLAST Search database core_nt using Megablast (Optimize for highly similar sequences)

☐ Show results in a new window

Exercise 1 - Metabarcoding detective

Descriptions	Graphic Summary	Alignments	Taxonomy					
Sequences producing significant alignments								
Download Select columns Show 100 ?								
select all 100 sequences selected								
GenBankGraphicsDistance tree of resultsMSA Viewer								
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Paracalanus parvus voucher P0033 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	579	579	100%	1e-160	100.00%	657	KP861426.1
Paracalanus parvus isolate IT2_9 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	579	579	100%	1e-160	100.00%	644	MH818625.1
Paracalanus parvus voucher V082.NEA1a cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus par...	579	579	100%	1e-160	100.00%	646	KF715875.1
Paracalanus parvus isolate P0244 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	579	579	100%	1e-160	100.00%	644	MH818684.1
Paracalanus parvus isolate MT00828 cytochrome c oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	577	577	100%	5e-160	100.00%	657	JX995216.1
Paracalanus parvus voucher PP0406.NEA cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus par...	577	577	100%	5e-160	100.00%	646	MN764491.1
Paracalanus parvus isolate Bul1_7 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	573	573	100%	6e-159	99.68%	644	MH818639.1
Paracalanus parvus isolate Bul3_2 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	573	573	100%	6e-159	99.68%	644	MH818656.1
Paracalanus parvus isolate Ukr1_4 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	573	573	100%	6e-159	99.68%	644	MH818669.1
Paracalanus parvus isolate IT7_29 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	573	573	100%	6e-159	99.68%	644	MH818637.1
Paracalanus parvus isolate IT2_14 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	573	573	100%	6e-159	99.68%	644	MH818626.1
Paracalanus parvus isolate MT03859 cytochrome c oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	571	571	100%	2e-158	99.68%	657	JX995220.1
Paracalanus parvus voucher PP0400.NEA cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus par...	571	571	100%	2e-158	99.68%	646	MN764495.1
Paracalanus parvus voucher V106.NEA2a cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus par...	568	568	100%	3e-157	99.36%	646	KF715881.1
Paracalanus parvus isolate Bul3_1 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	568	568	100%	3e-157	99.36%	644	MH818655.1

Exercise 1 - Metabarcoding detective



Exercise 1 - Metabarcoding detective

Descriptions

Graphic Summary

Alignments

Taxonomy

Alignment view

Pairwise

☐ CDS feature [Restore defaults](#)

Download

100 sequences selected

Download

GenBank

Graphics

Next

Previous

Descriptions

Paracalanus parvus voucher P0033 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial
Sequence ID: [KP861426.1](#) Length: 657 Number of Matches: 1

Range 1: 345 to 657 [GenBank](#) [Graphics](#)

Next Match Previous Match

Score	Expect	Identities	Gaps	Strand
579 bits(313)	1e-160	313/313(100%)	0/313(0%)	Plus/Plus
Query 1	ACTATCTAGGAACATTGCACATGCCGGAAGGTCCGTAGACTTTGCCATCTTTTCGTTACA	60		
Sbjct 345	ACTATCTAGGAACATTGCACATGCCGGAAGGTCCGTAGACTTTGCCATCTTTTCGTTACA	404		
Query 61	TTTAGCCGGAGTTAGTTCAATTCTTGGGGCAGTAAATTTTATTAGGACTTTAGGAAACCT	120		
Sbjct 405	TTTAGCCGGAGTTAGTTCAATTCTTGGGGCAGTAAATTTTATTAGGACTTTAGGAAACCT	464		
Query 121	TCGAGTATTTGGAATATTACTAGATCGCATACCTCTGTTTGCATGAGCTGTATTAATTAC	180		
Sbjct 465	TCGAGTATTTGGAATATTACTAGATCGCATACCTCTGTTTGCATGAGCTGTATTAATTAC	524		
Query 181	TGCTGTACTATTACTTCTTTCTTTACCTGTTCTAGCTGGGGCTATTACCATAATTATTAAC	240		
Sbjct 525	TGCTGTACTATTACTTCTTTCTTTACCTGTTCTAGCTGGGGCTATTACCATAATTATTAAC	584		
Query 241	CGATCGTAATTTGAATACAACATTTTATGATGTGGGGGAGGAGGAGATCCTATTCTATA	300		
Sbjct 585	CGATCGTAATTTGAATACAACATTTTATGATGTGGGGGAGGAGGAGATCCTATTCTATA	644		

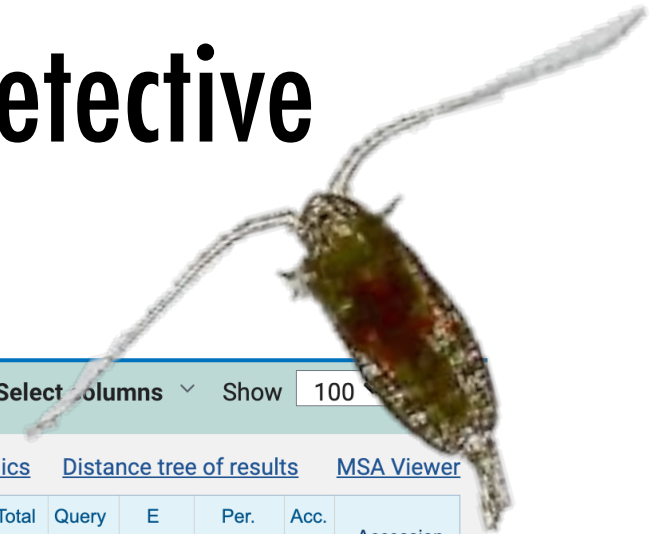
Exercise 1 - Metabarcoding detective

Descriptions Graphic Summary Alignments Taxonomy					
Reports Lineage Organism Taxonomy					
100 sequences selected ?					
Organism	Blast Name	Score	Number of Hits	Description	
Arthropoda	arthropods		191		
• Calanoida	crustaceans		188		
• Paracalanidae	crustaceans		182		
• Paracalanus	crustaceans		164		
• Paracalanus parvus	crustaceans	579	96	Paracalanus parvus hits	
• Paracalanus sp. F AC-2013	crustaceans	457	20	Paracalanus sp. F AC-2013 hits	
• Paracalanus indicus	crustaceans	324	41	Paracalanus indicus hits	
• Paracalanus tropicus	crustaceans	320	3	Paracalanus tropicus hits	
• Paracalanus aff. indicus AC-2013	crustaceans	318	1	Paracalanus aff. indicus AC-2013 hits	
• Paracalanus quasimodo	crustaceans	318	1	Paracalanus quasimodo hits	
• Paracalanus sp. PP0417.GOM	crustaceans	313	1	Paracalanus sp. PP0417.GOM hits	
• Paracalanus parvus environmental sample	crustaceans	296	1	Paracalanus parvus environmental sample hits	
• Acrocalanus monachus	crustaceans	309	2	Acrocalanus monachus hits	
• Acrocalanus gibber	crustaceans	307	2	Acrocalanus gibber hits	

Exercise 1 - Metabarcoding detective

Descriptions	Graphic Summary	Alignments	Taxonomy					
Sequences producing significant alignments								
Download Select columns Show 100 ?								
☒ select all 100 sequences selected								
GenBank Graphics Distance tree of results MSA Viewer								
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☒ Paracalanus parvus voucher P0033 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	579	579	100%	1e-160	100.00%	657	KP861426.1
☒ Paracalanus parvus isolate IT2_9 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	579	579	100%	1e-160	100.00%	644	MH818625.1
☒ Paracalanus parvus voucher V082.NEA1a cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus par...	579	579	100%	1e-160	100.00%	646	KF715875.1
☒ Paracalanus sp. F AC-2013 voucher V001.SEA2 cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus sp...	448	448	99%	4e-121	92.60%	647	KF715968.1
☒ Paracalanus sp. F AC-2013 voucher V420.SEA5a cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus sp...	448	448	99%	4e-121	92.60%	647	KF715973.1
☒ Paracalanus parvus voucher PaPa1 cytochrome c oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	446	446	99%	1e-120	92.60%	625	KC287780.1
☒ Paracalanus sp. F AC-2013 voucher V001.SEA1 cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus sp...	442	442	99%	2e-119	92.28%	647	KF715967.1
☒ Paracalanus sp. F AC-2013 voucher V420.SEA4a cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus sp...	436	436	99%	8e-118	91.96%	647	KF715970.1
☒ Paracalanus parvus isolate IT2_26 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	422	422	74%	2e-113	99.57%	562	MH818634.1
☒ Paracalanus sp. F AC-2013 isolate IT2_29 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus sp...	329	329	70%	1e-85	93.61%	552	MH818800.1
☒ Paracalanus indicus voucher V341.PI2a cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus indi...	324	324	96%	7e-84	86.05%	647	KF715901.1
☒ Paracalanus indicus voucher V071.PI1a cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus indi...	324	324	96%	7e-84	86.05%	647	KF715892.1
☒ Paracalanus parvus isolate Ppar_Ok cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial gene for...	Paracalanus par...	324	324	96%	7e-84	86.05%	656	AF474111.1
☒ Paracalanus parvus voucher PaPa3 cytochrome c oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	324	324	96%	7e-84	86.05%	646	KC287782.1

Exercise 1 - Metabarcoding detective



Descriptions

Graphic Summary

Alignments

Taxonomy

Sequences producing significant alignments

Download

Select columns

Show100

☒ select all

100 sequences selected

GenBank

Graphics

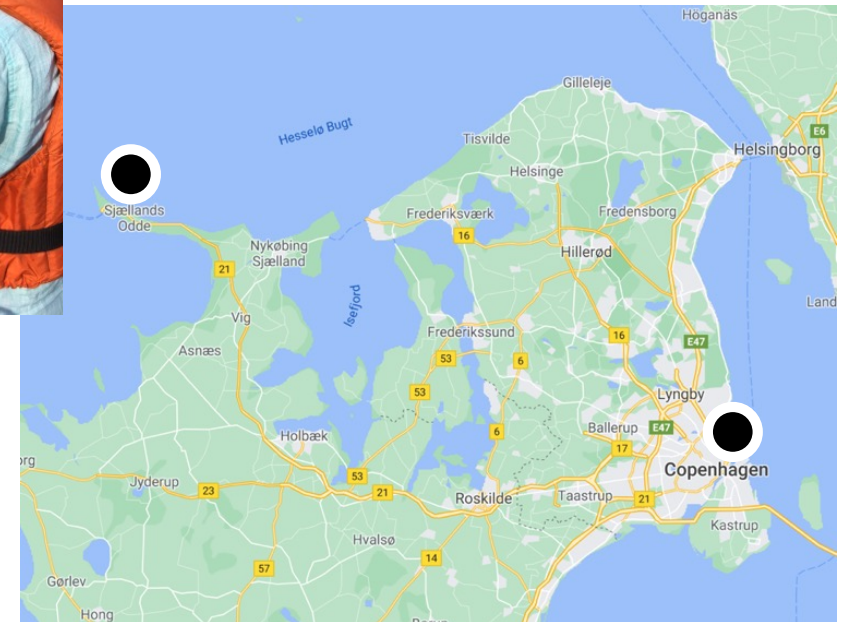
Distance tree of results

MSA Viewer

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
☒	Paracalanus parvus voucher P0033 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	579	579	100%	1e-160	100.00%	657	KP861426.1
☒	Paracalanus parvus isolate IT2_9 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	579	579	100%	1e-160	100.00%	644	MH818625.1
☒	Paracalanus parvus voucher V082.NEA1a cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus par...	579	579	100%	1e-160	100.00%	646	KF715875.1
☒	Paracalanus parvus isolate P0244 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	579	579	100%	1e-160	100.00%	644	MH818684.1
☒	Paracalanus parvus isolate MT00828 cytochrome c oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	577	577	100%	5e-160	100.00%	657	JX995216.1
☒	Paracalanus parvus voucher PP0406.NEA cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus par...	577	577	100%	5e-160	100.00%	646	MN764491.1
☒	Paracalanus parvus isolate Bul1_7 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	573	573	100%	6e-159	99.68%	644	MH818639.1
☒	Paracalanus parvus isolate Bul3_2 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	573	573	100%	6e-159	99.68%	644	MH818656.1
☒	Paracalanus parvus isolate Ukr1_4 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	573	573	100%	6e-159	99.68%	644	MH818669.1
☒	Paracalanus parvus isolate IT7_29 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	573	573	100%	6e-159	99.68%	644	MH818637.1
☒	Paracalanus parvus isolate IT2_14 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	573	573	100%	6e-159	99.68%	644	MH818626.1
☒	Paracalanus parvus isolate MT03859 cytochrome c oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	571	571	100%	2e-158	99.68%	657	JX995220.1
☒	Paracalanus parvus voucher PP0400.NEA cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus par...	571	571	100%	2e-158	99.68%	646	MN764495.1
☒	Paracalanus parvus voucher V106.NEA2a cytochrome oxidase subunit I gene, partial cds; mitochondrial	Paracalanus par...	568	568	100%	3e-157	99.36%	646	KF715881.1
☒	Paracalanus parvus isolate Bul3_1 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial	Paracalanus par...	568	568	100%	3e-157	99.36%	644	MH818655.1

Exercise 1 - Metabarcoding detective

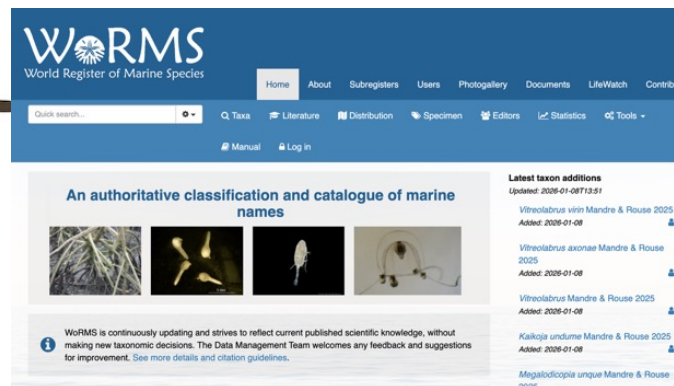
TCTAGCTGGTAACTTAGCCACGCGGGAG
CATCCGTTGATTAAACAATTTTTCTCTTCATC
TAGCAGGTATTTTCATCAATTTTAGGCGCTATT
AATTTATTACCACTATTATTAATATAAAACCA
CCTGCAATTTACAAATATCAAACCCCACTAT
TTGTTTGAGCCGTATTAATTACAGCCGTCCT
TCTTCTCTTATCCCTACCCGTCCTAGCTGCT
GGAATTACAATACTTCTGACAGACCGAAAC
CTAAACACTACCTTTTTTTGATCCTGCTGGCG
GTGGAGATCCTATTCTTTACCAACATTTGTTC



Exercise 1 - Metabarcode detective



ACTATCTAGGAACATTGCACATGCCGG
AAGGTCCGTAGACTTTGCCATCTTTTCG
TTACATTTAGCCGGAGTTAGTTCAATTCT
TGGGGCAGTAAATTTTATTAGGACTTTAG
GAAACCTTCGAGTATTTGGAATATTACTA
GATCGCATACCTCTGTTTGCATGAGCTG
TATTAATTACTGCTGTACTATTACTTCTTT
CTTTACCTGTTCTAGCTGGGGGCTATTAC
CATATTATTAACCGATCGTAATTTGAATA
CAACATTTTATGATGTGGGGGGGAGGAG
GAGATCCTATTCTATATCAACACCTATTC



A collage of various cartoon-style illustrations of animals and plants on a green background. The illustrations include a yellow and grey frog, three red tomatoes, a brown and white bird, a blue shark, two blue whales, a small brown lizard, a panda's head, a green and red cactus, a brown lizard, a green plant with red flowers, a brown and white striped caterpillar, and a white and orange flower. A large black circle is centered over the collage, containing the word "Results" in white text.

Results

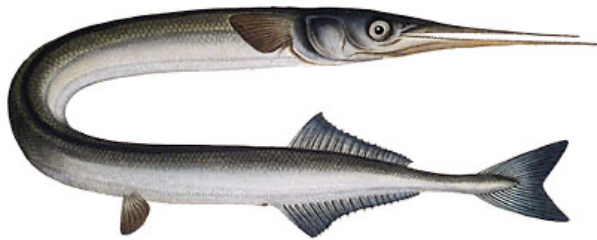
Seq 1



Seq 1

Belone belone

Hornfisk/Garfish



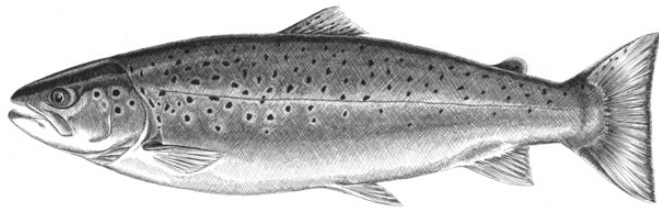
Seq 2



Seq 2

Salmo trutta

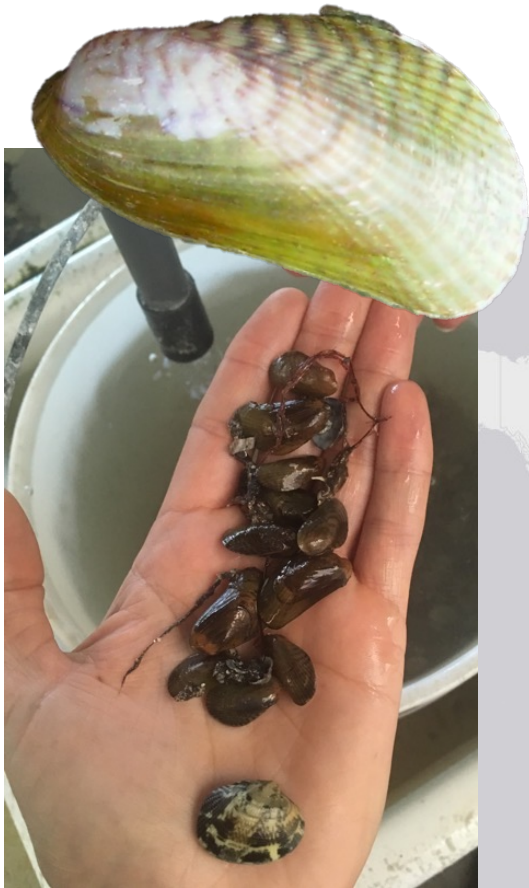
Sea trout/Brown trout



Seq 3



Seq 3 *Arcuatula senhousia*



Seq 3 Arcuatula senhousia



WoRMS
World Register of Marine Species

Home About Subregisters Users Photogallery Documents LifeWatch Contribute

Quick search...

WoRMS name details

★ ***Musculista senhousia* (Benson, 1842)**

AphiaID 140470 (urn:lsid:marinespecies.org:taxname:140470)

Classification Biota > ★ *Animalia* (Kingdom) > ★ *Mollusca* (Phylum) > ★ *Bivalvia* (Class) > ★ *Pteriomorpha* (Subclass) > ★ *Mytilida* (Order) > ★ *Mytilioidea* (Superfamily) > ★ *Mytilidae* (Family) > ★ *Musculista* (Genus) > ★ *Musculista senhousia* (Species)

Status ✖ **unaccepted**

Accepted Name ★ *Arcuatula senhousia* (Benson, 1842)

Rank Species

Parent ★ *Musculista* Yamamoto & Habe, 1958 accepted as ★ *Arcuatula* Jousseaume in Lamy, 1919

Orig. name ★ *Modiola senhousia* Benson, 1842

Environment marine

Original description (of ★ *Modiola senhousia* Benson, 1842) Benson W.H. (1842). Mollusca. in: T. Cantor, General features of Chusan, with remarks on the flora and fauna of that island. *Annals and Magazine of Natural History*. ser. 1, 9: 486-490., available online at <http://www.biodiversitylibrary.org/item/61919#page/500/mode/1up> page(s): 489 [details]

Taxonomic citation MolluscaBase (2019). MolluscaBase. *Musculista senhousia* (Benson, 1842). Accessed

Seq 3 Arcuatula senhousia

- Taxonomy is a dynamic discipline
- Use the taxonomy matching tool in WoRMS (if marine) or similar databases



WoRMS
World Register of Marine Species

marine only ☐ extant only ☒

Home About Subregisters Users Photogallery Documents LifeWatch Contribute

Quick search... Q Taxa Literature Distribution Specimen Editors Statistics Tools

★ *Musculista senhousia* (Benson, 1842)

AphiaID 140470 (urn:lsid:marinespecies.org:taxname:140470)

Classification Biota > ★ [Animalia](#) (Kingdom) > ★ [Mollusca](#) (Phylum) > ★ [Pteriomorpha](#) (Subclass) > ★ [Mytilida](#) (Order) > ★ [Mytilidae](#) (Family) > ★ [Musculista](#) (Genus) > ★ [Musculista](#) (Species)

Status ✖ **unaccepted**

Accepted Name ★ [Arcuatula senhousia](#) (Benson, 1842)

Parent	★ Musculista Yamamoto & Habe, 1958 accepted as ★ Arcuatula Jousseaume in Lamy, 1919
Orig. name	★ Modiola senhousia Benson, 1842
Environment	marine
Original description	(of ★ Modiola senhousia Benson, 1842) Benson W.H. (1842). Mollusca. in: T. Cantor, General features of Chusan, with remarks on the flora and fauna of that island. <i>Annals and Magazine of Natural History</i> . ser. 1, 9: 486-490., available online at http://www.biodiversitylibrary.org/item/61919#page/500/mode/1up page(s): 489 [details]
Taxonomic citation	MolluscaBase (2019). MolluscaBase. <i>Musculista senhousia</i> (Benson, 1842). Accessed

Seq 4



Seq 4 *Caenus pusilla*



**-eDNA contains many
non-target organisms**

**-Think critically
about transport**



Seq 4 *Caenus pusilla*



-eDNA contains many non-target organisms

-Think critically about transport

-Use taxonomy to subset data by species or groups of interest.



Seq 5



Seq 5 *Mytilus* spp.

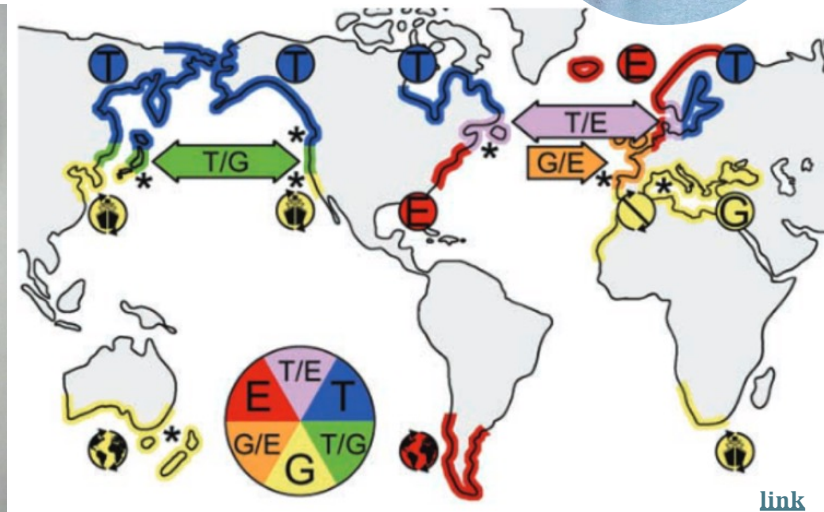


Mytilus edulis

Mytilus trossulus

Mytilus galloprovincialis

Seq 5 *Mytilus* spp.



Mytilus edulis
Mytilus trossulus
Mytilus galloprovincialis

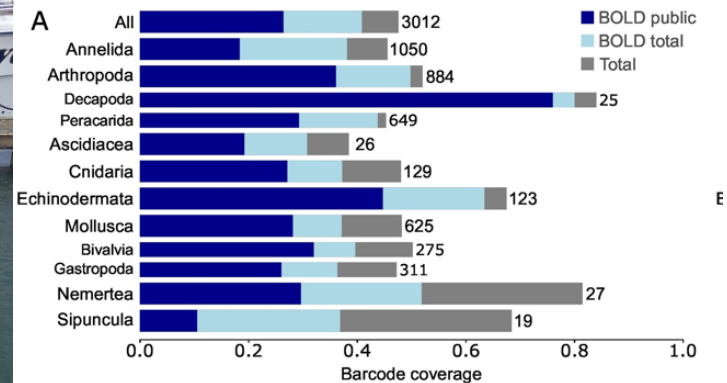
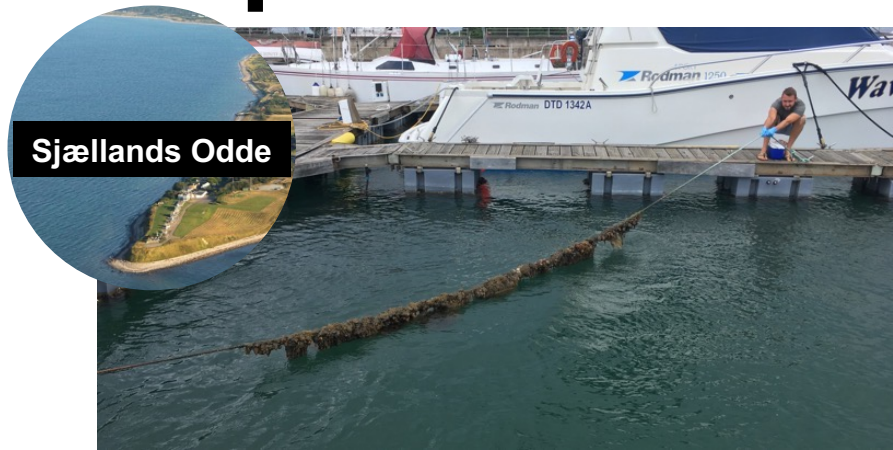
**–eDNA often blind to hybrids
and species complexes**

Seq 6



Sjællands Odde

Seq 6 Unknown botryllid ascidian



Weigand et al. 2019

-Less than 50% coverage
of barcodes in Europe!

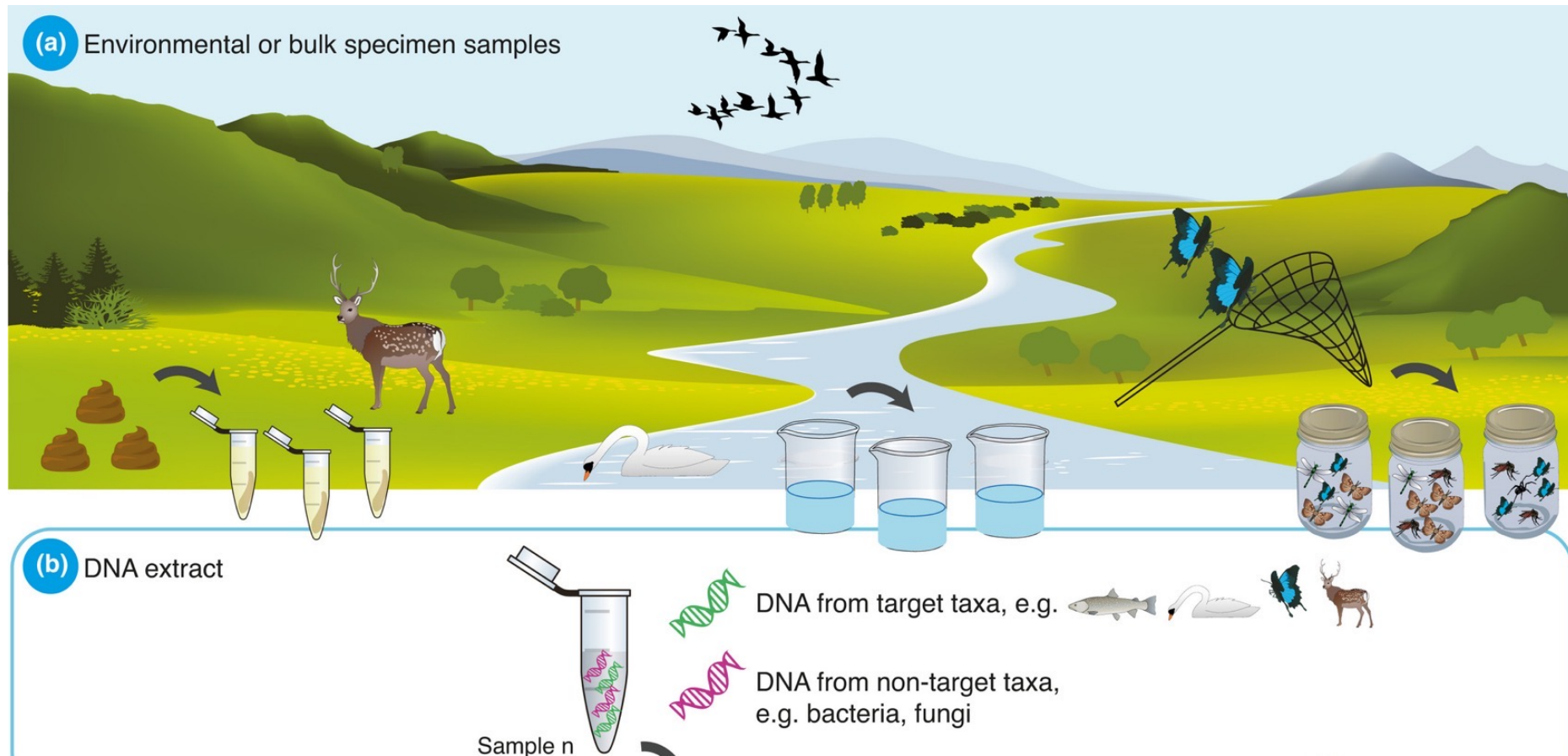
-38% in Ascidacea ☹

Results

How do we do metabarcoding?



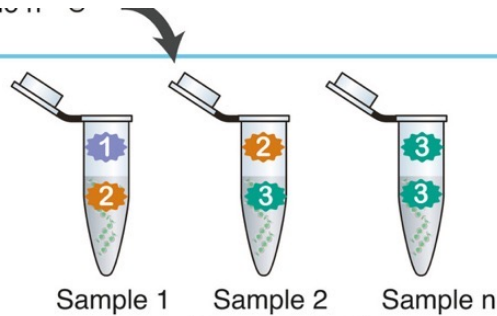
How do we do metabarcoding?



How do we do metabarcoding?

(c) Metabarcoding

PCR amplification with primers for a taxonomically informative marker for a taxonomic group, e.g. generic vertebrate, plant, and insect primers



Unique identifiers attached to each PCR product

- 5' nucleotide tags and/or library indices

(d) High-throughput sequencing

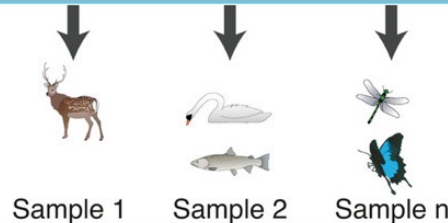


(e) Tracing sequences back to samples

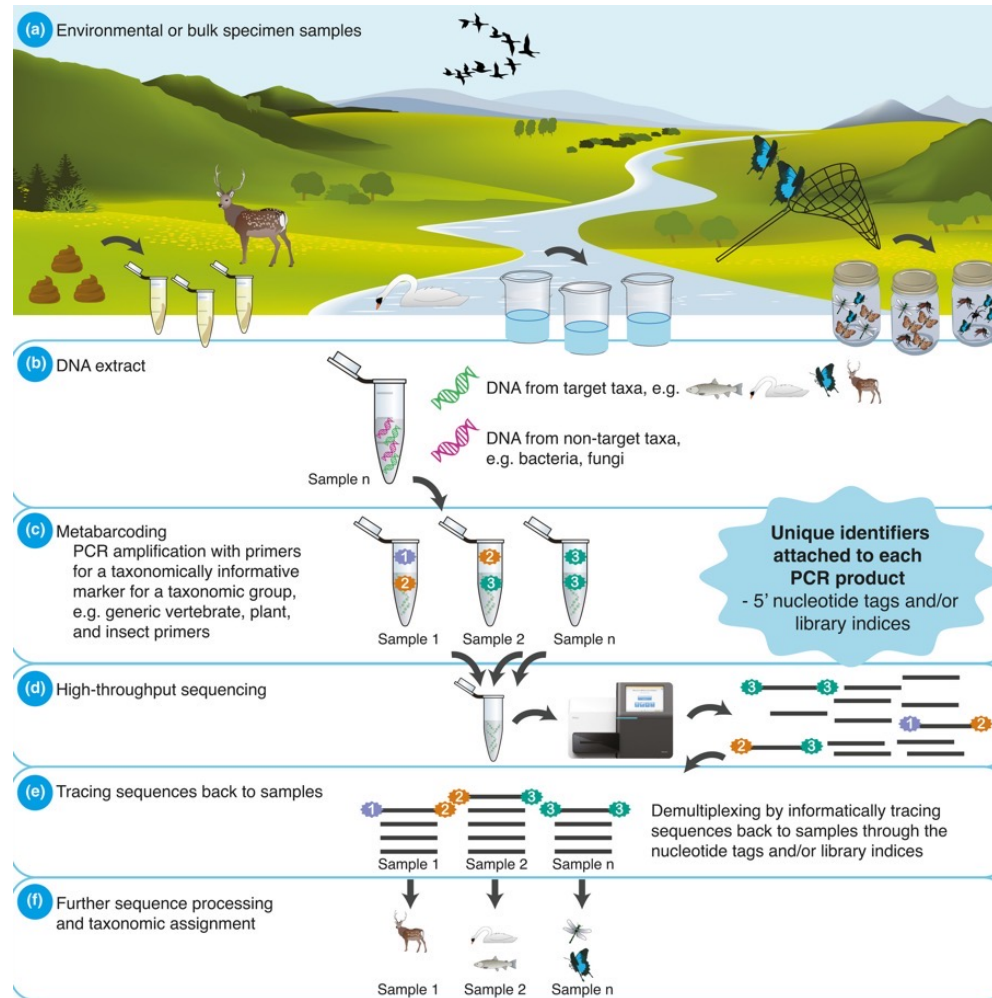


Demultiplexing by informatically tracing sequences back to samples through the nucleotide tags and/or library indices

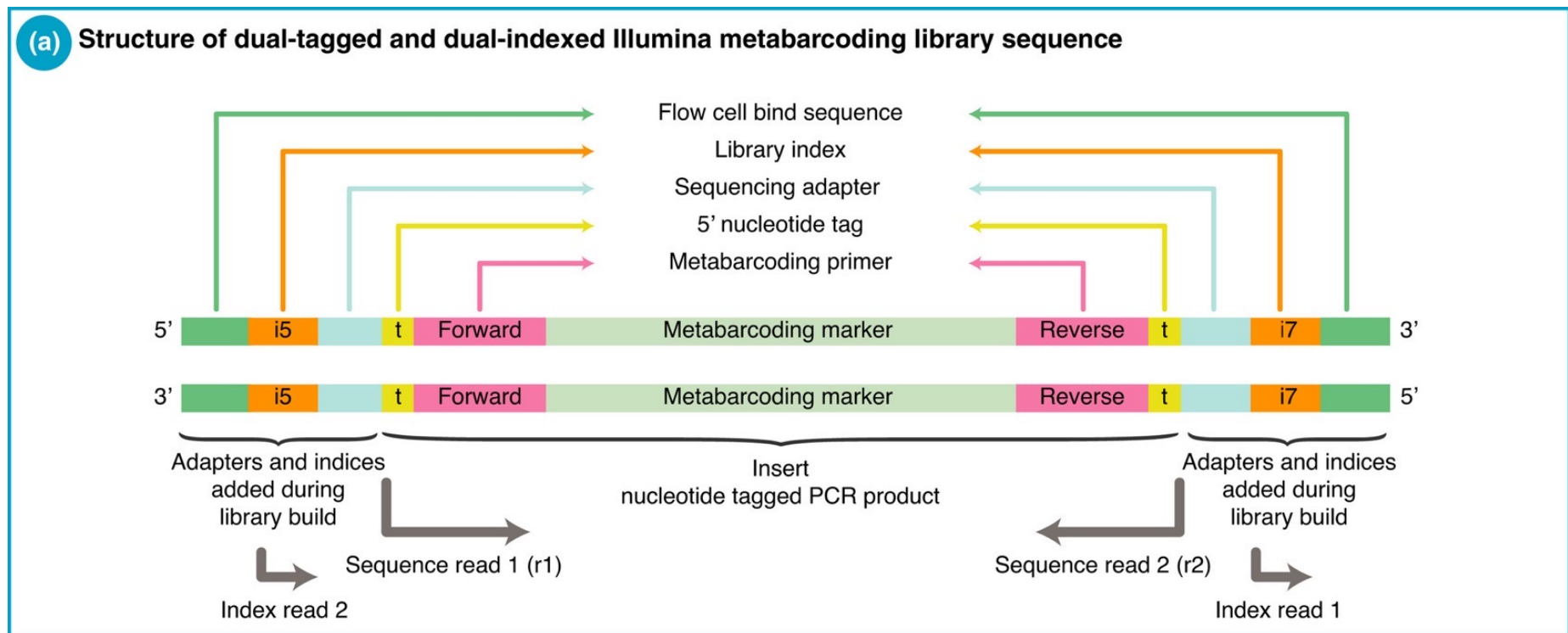
(f) Further sequence processing and taxonomic assignment



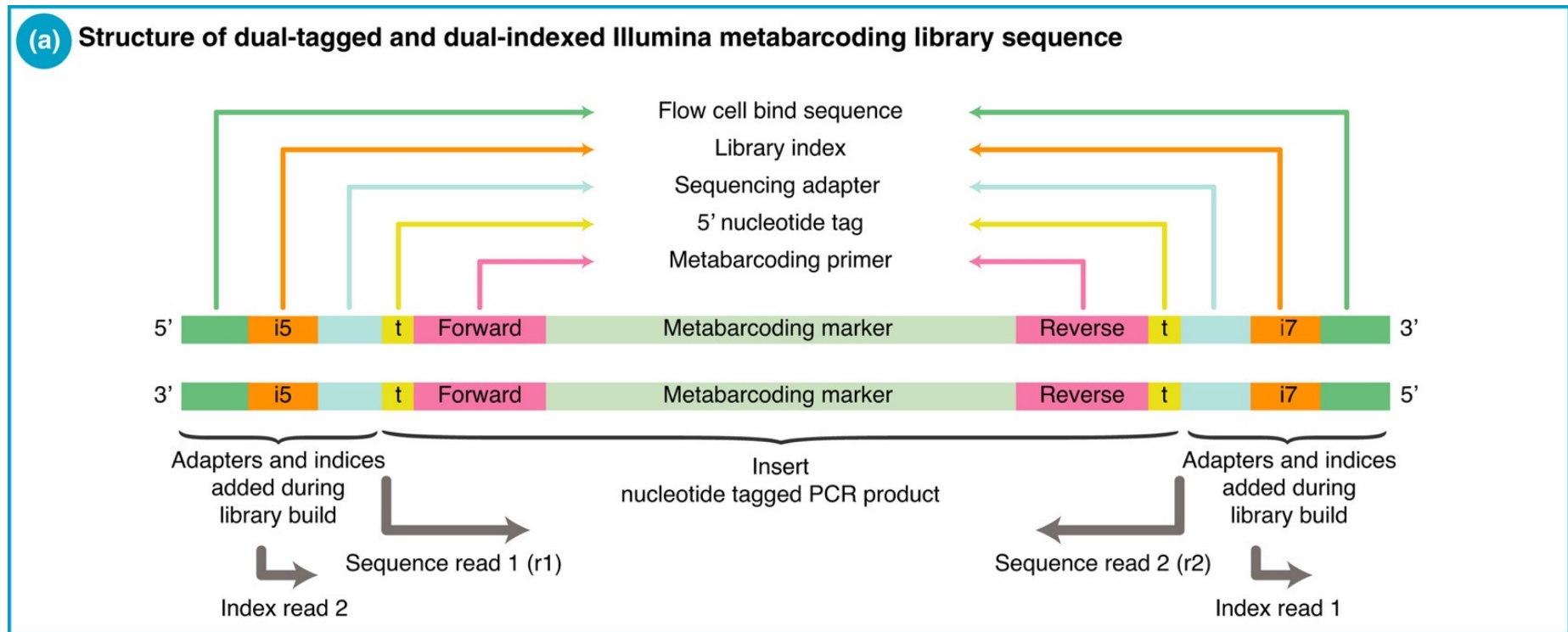
How do we do metabarcoding?



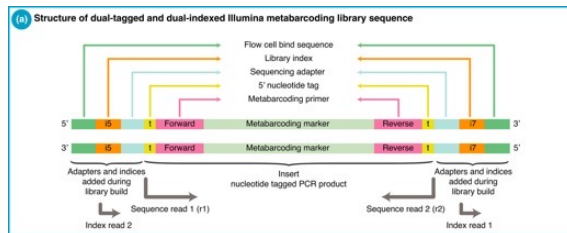
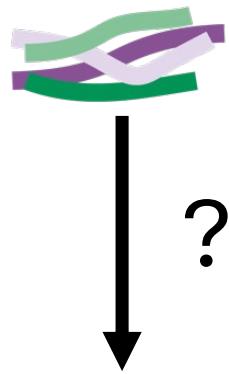
How do we do metabarcoding?



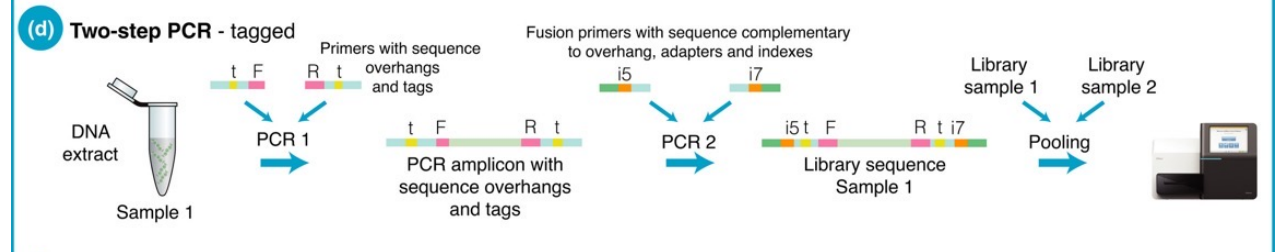
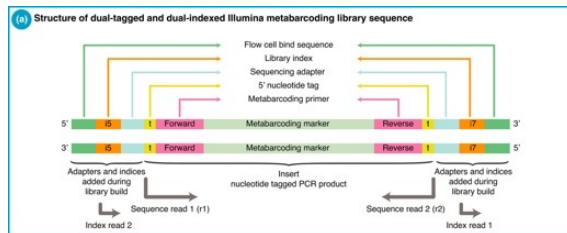
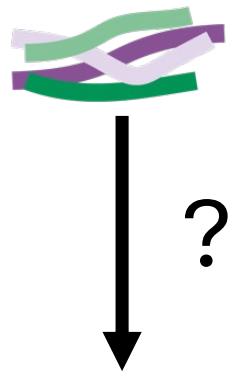
How do we do metabarcoding?



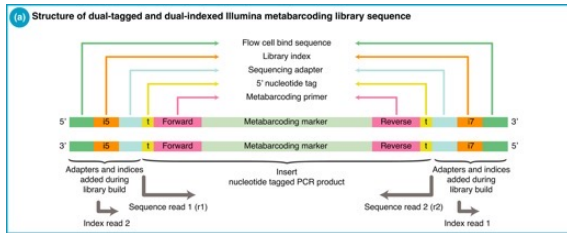
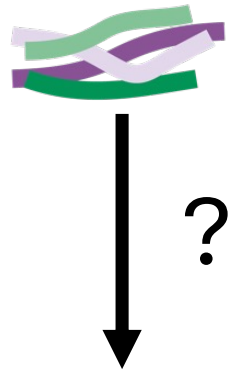
How do we do metabarcoding?



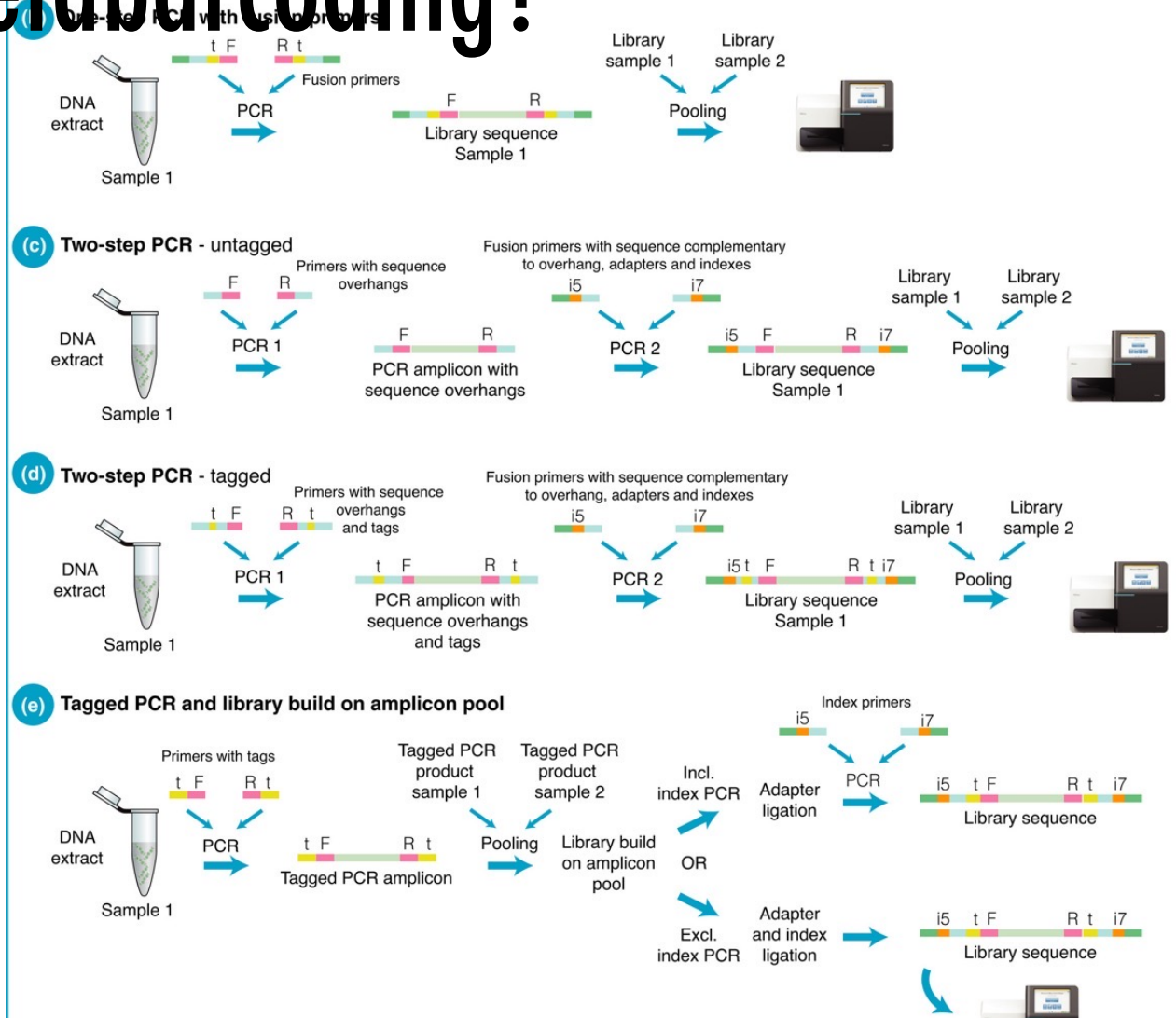
How do we do metabarcoding?



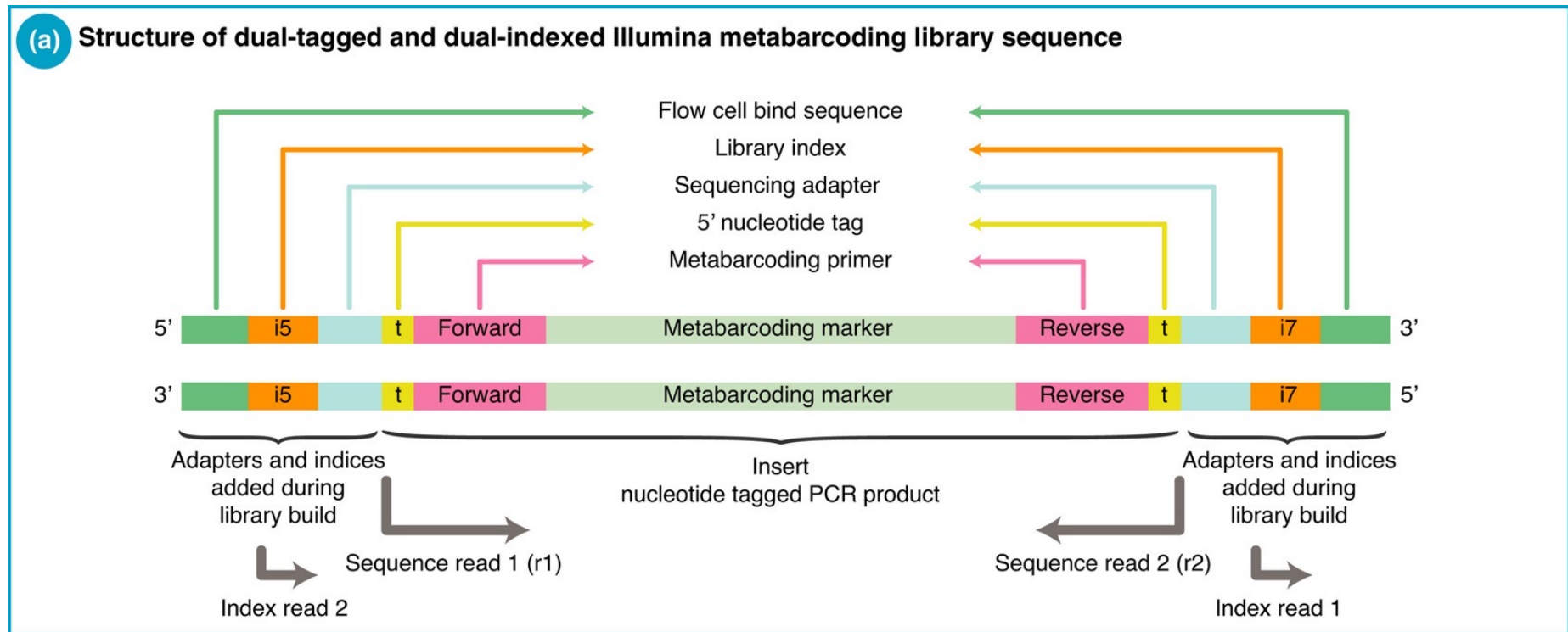
How do we do metabarcoding?



Bohmann et al. 2021



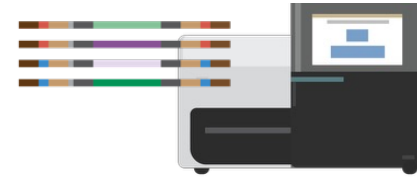
How do we do metabarcoding?



eDNA Metabarcoding Bioinformatics

AIM: To cluster or filter sequences into biologically relevant groupings assigned to samples while minimizing the effect of biases

Sequencing Library



Raw Reads



Quality Control & dereplication



Merging pairs



Sequence by Sample Table

	A	B	C	D
1		Sample_1	Sample_2	Sample_N
2	OTU_1	0	20	0
3	OTU_2	500	20	0
4	OTU_N	100	20	0

Metabarcoding Pipelines

Simple



Complicated



Metabarcoding Pipelines

Simple



Constrained

Complicated



Adaptable

Metabarcoding Pipelines

Simple



Constrained

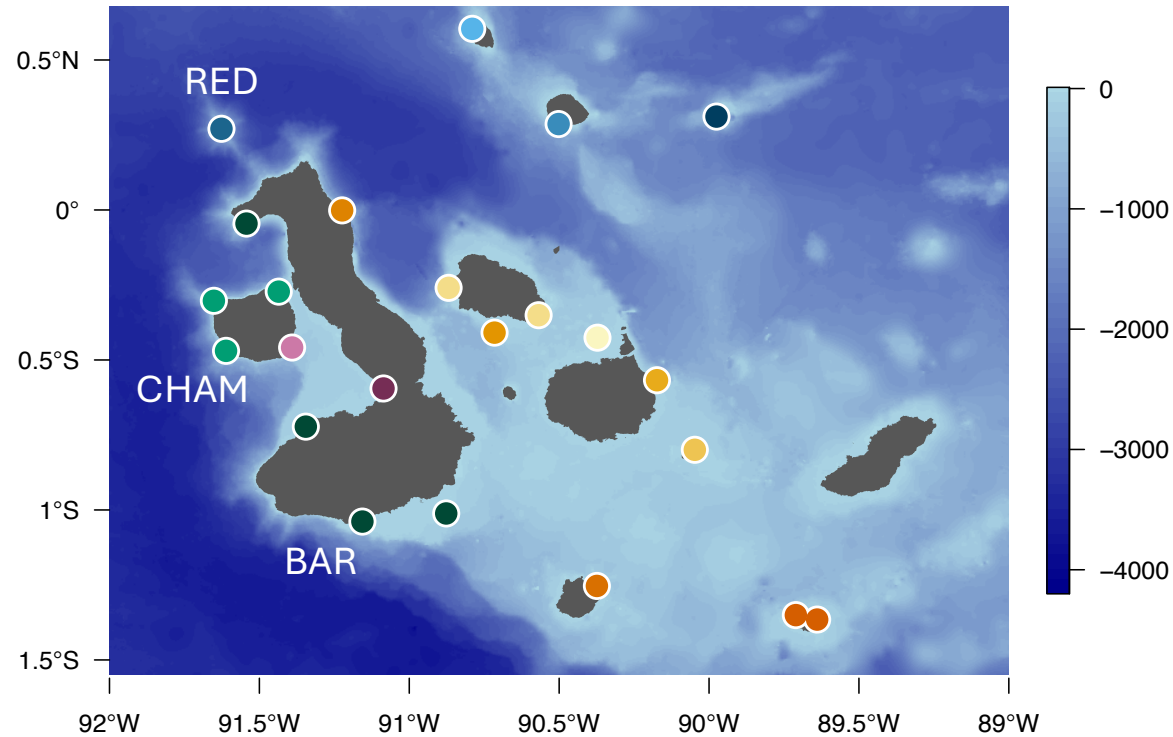


Complicated



Adaptable

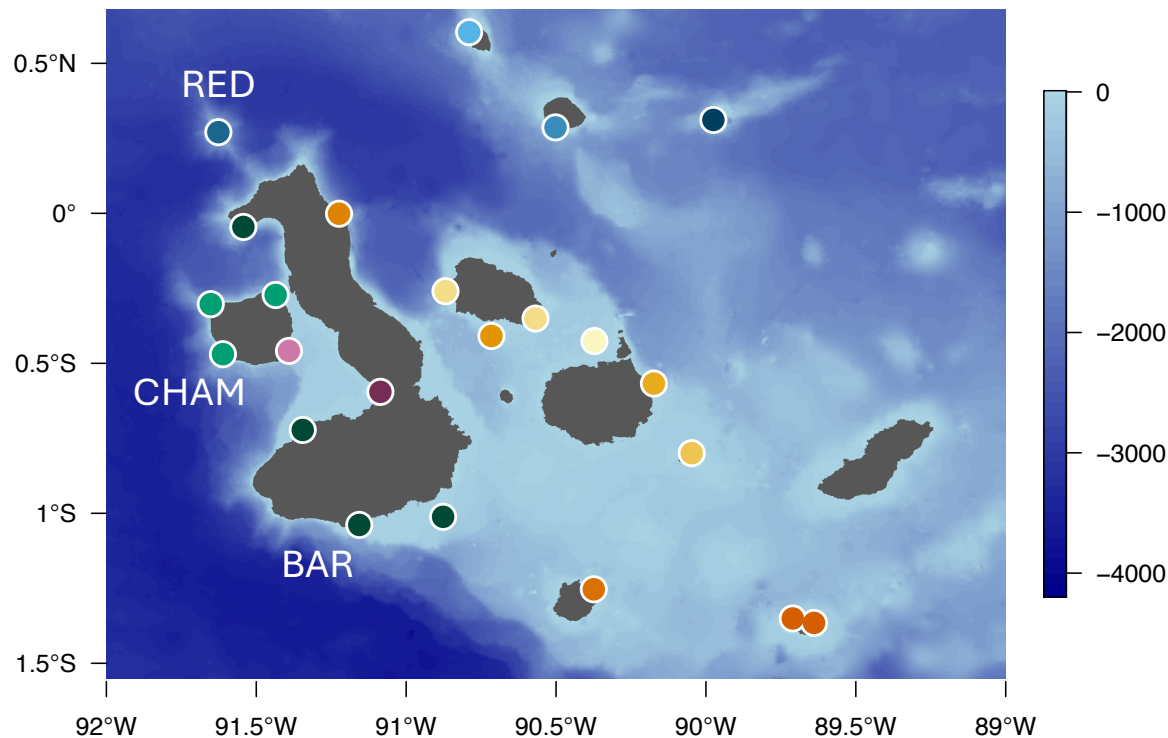
Metabarcoding Dataset



RED - Roca Redonda
CHAM - Cabo Hammond
BAR - Barahona Beach

C1 - Control 1 (extraction control)
C2 - Control 2 (PCR control)

Metabarcoding Dataset



What DNA region is this data from?

What is the most common taxa?

Do we see differences between our samples?

bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

HOME | SUBMIT | FAQ | BLOG | ALERTS / RSS | RESOURCES | ABOUT | CHANNELS

Search
Advanced Search

New Results

[Follow this preprint](#) [Previous](#)

[Next](#)

Fine-scale oceanographic processes shape marine biodiversity patterns in the Galápagos Islands

Posted November 13, 2024.

[Luke E Holman](#), [Diana A. Pazmiño](#), [Shyam Gopalakrishnan](#), [Alexander Forryan](#), [Alex R. Hearn](#),
[Alberto C. Naveira-Garabato](#), [Marc Rius](#)
doi: <https://doi.org/10.1101/2024.03.06.583537>

This article is a preprint and has not been certified by peer review [what does this mean?]

[Download PDF](#)
[Print/Save Options](#)
[Supplementary Material](#)
[Revision Summary](#)
[Email](#)
[Share](#)
[Citation Tools](#)
[Get QR code](#)

[Abstract](#) [Full Text](#) [Info/History](#) [Metrics](#) [Supplementary material](#) [Preview PDF](#)

Subject Area

Ecology

Abstract

Uncovering the drivers that shape biodiversity patterns is critical to understand ecological and evolutionary dynamics. Despite evidence that biodiversity composition is influenced by processes at different spatial scales, little is known about the role of fine-scale oceanographic processes on the structure of marine communities. This is particularly important in biodiversity

Reviews and Context

[0](#) Comment

[0](#) TRIP Peer Reviews