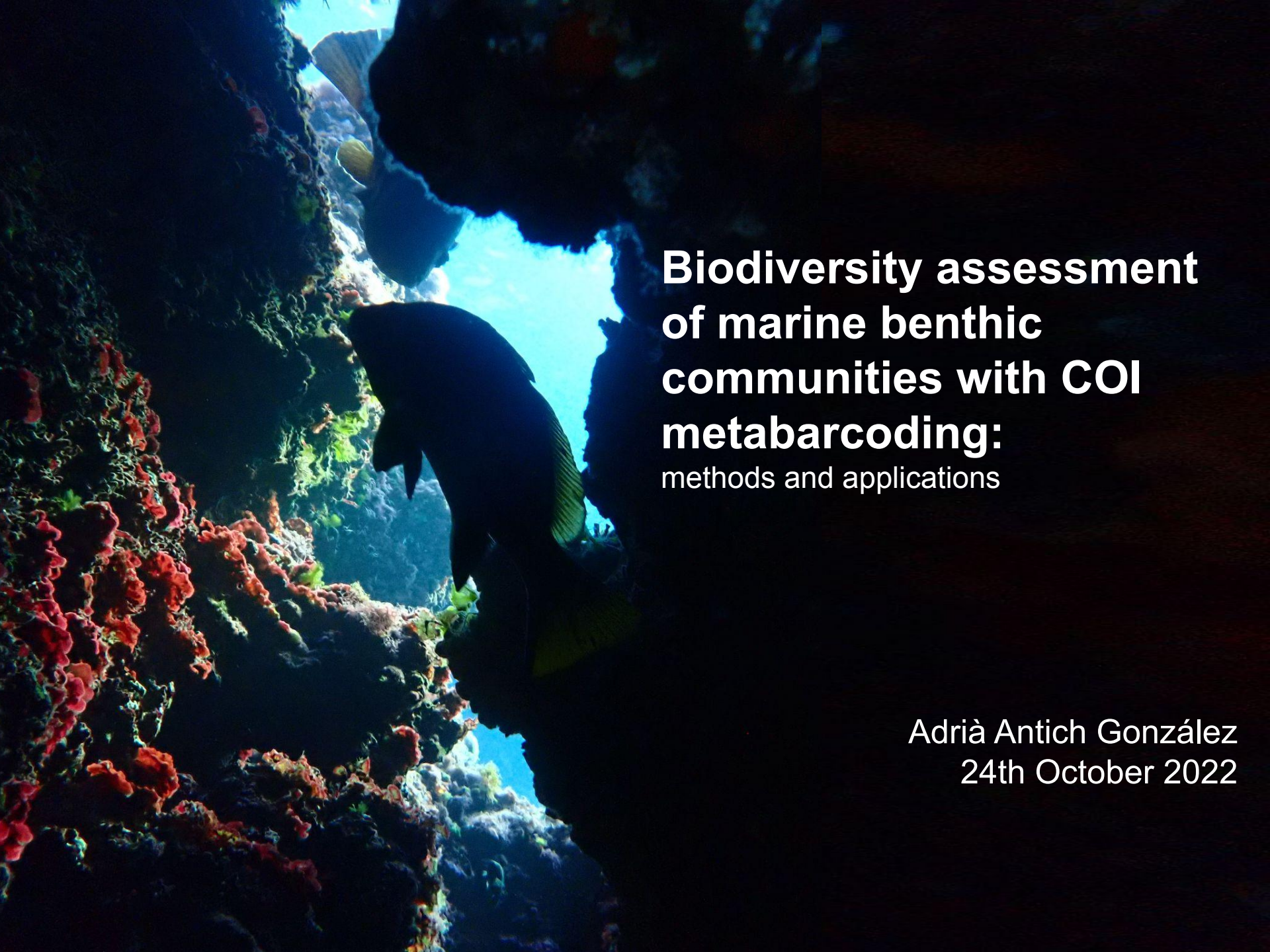


Biodiversity assessment of marine benthic communities with COI metabarcoding:

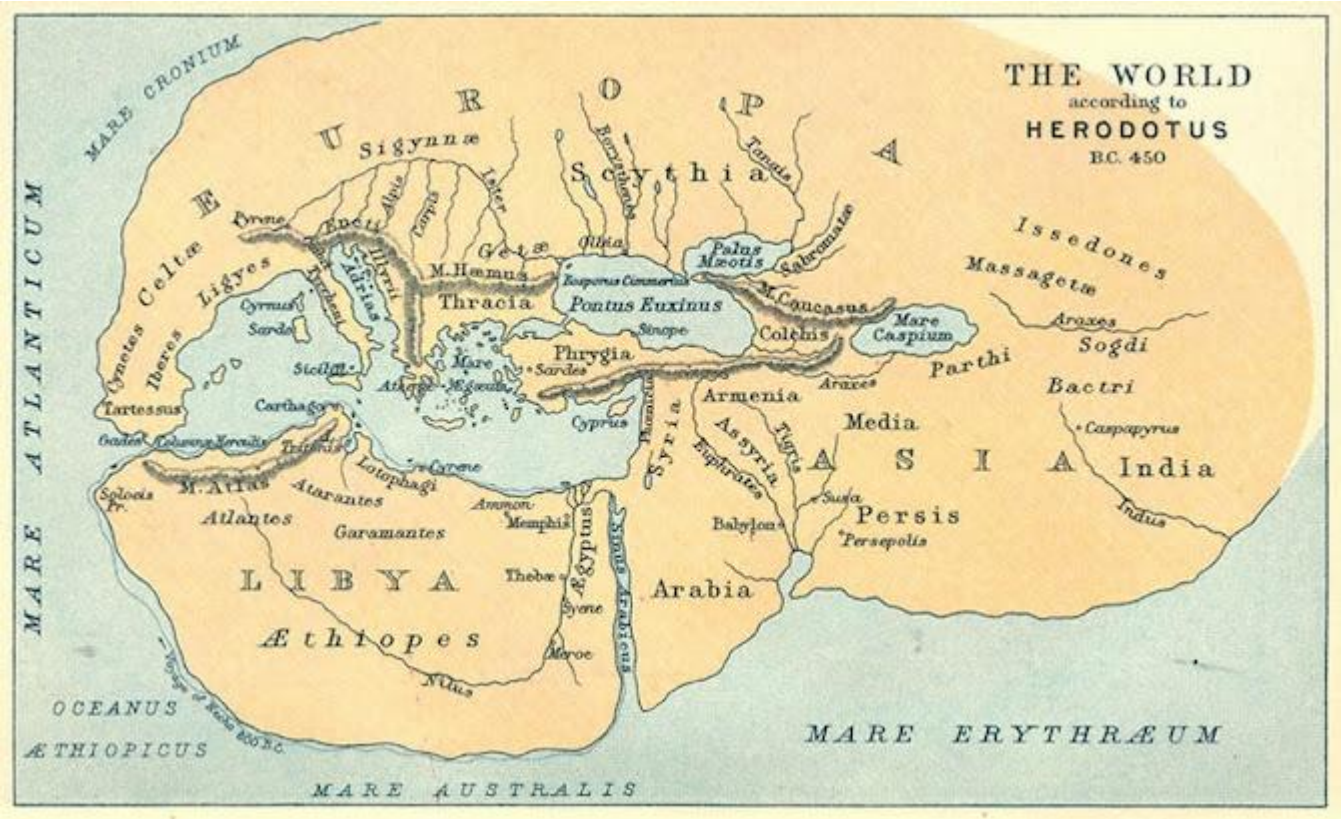
methods and applications

Adrià Antich González

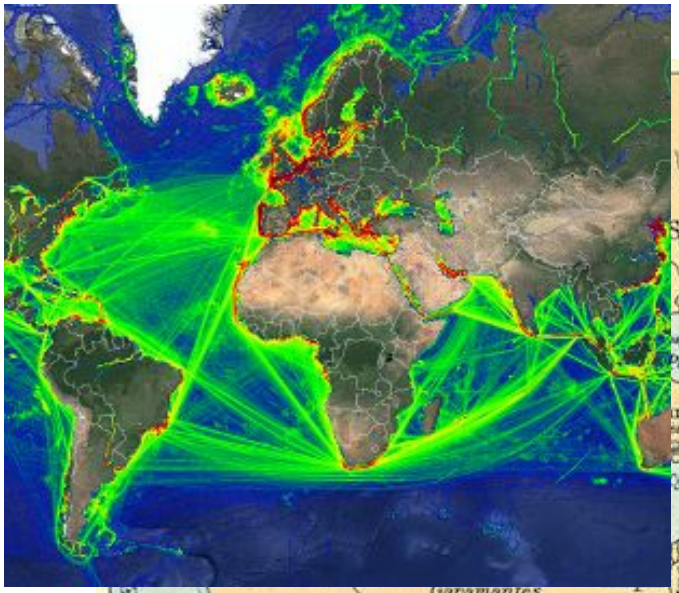
24th October 2022


INTRODUCTION

The importance of the sea and the benthic communities



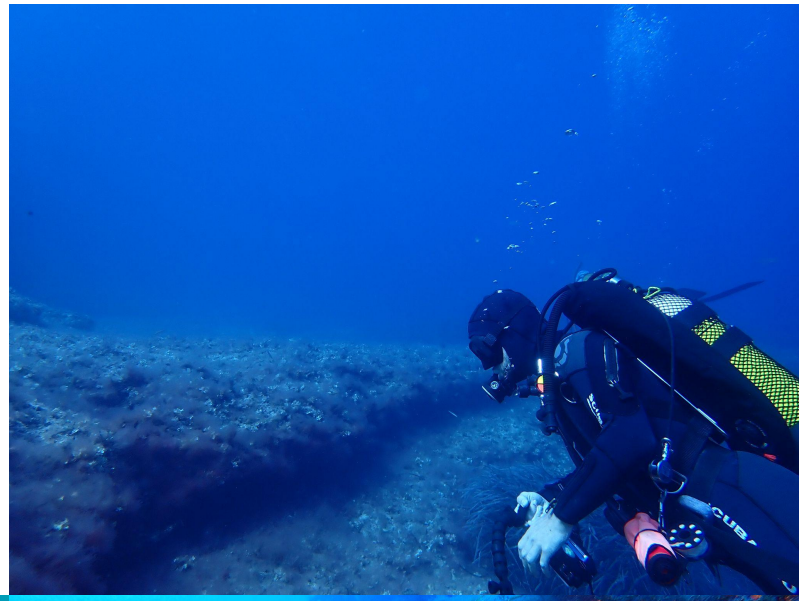
The importance of the sea and the benthic communities



The importance of the sea and the benthic communities

Benthic communities

biodiversity reservoirs
physical protection to waves
blue carbon
ludic activities
...



Major Threads

OUR WORK GET INVOLVED ABOUT US DONATE + ADOPT +

WWF

PRESS RELEASES

69% average decline in wildlife populations since 1970, says new WWF report

Wildlife populations in Latin America and the Caribbean plummeting at a staggering rate of 94%
Freshwater species populations have suffered an 83% fall
The report's Living Planet Index shows that there is no time to lose in securing a nature-positive society

DATE
October 13, 2022

MEDIA CONTACT
Brooke Hirsheimer
(202) 495-4759
brooke.hirsheimer@wwfus.org
Brendan Rohr
202-495-4621
brendan.rohr@wwfus.org

Share 306 Tweet

WASHINGTON, D.C. (October 12, 2022) - Monitored populations of vertebrates (mammals, birds, amphibians, reptiles and fish) **have seen a devastating 69% drop on average since 1970,** according to World Wildlife Fund's (WWF) *Living Planet Report 2022*. Populations in Latin America and the Caribbean have fared worst, with an average decline of 94%. Global freshwater species have also been disproportionately impacted, declining 83% on average.

Major Threads

PLOS ONE

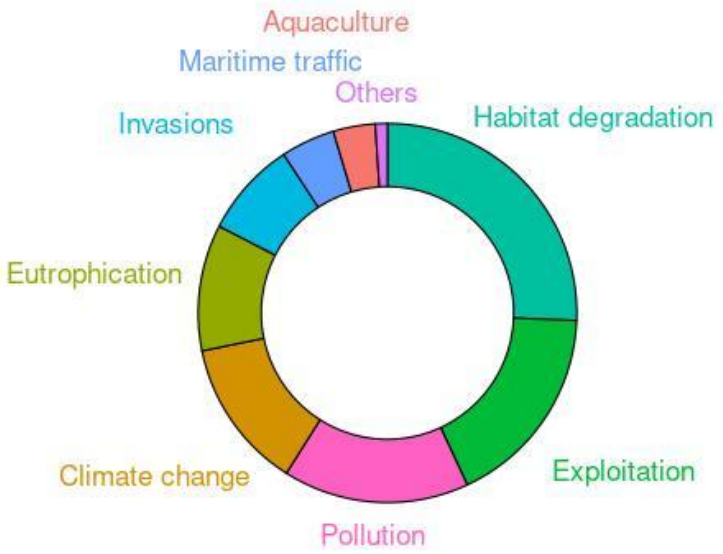
OPEN ACCESS

REVIEW

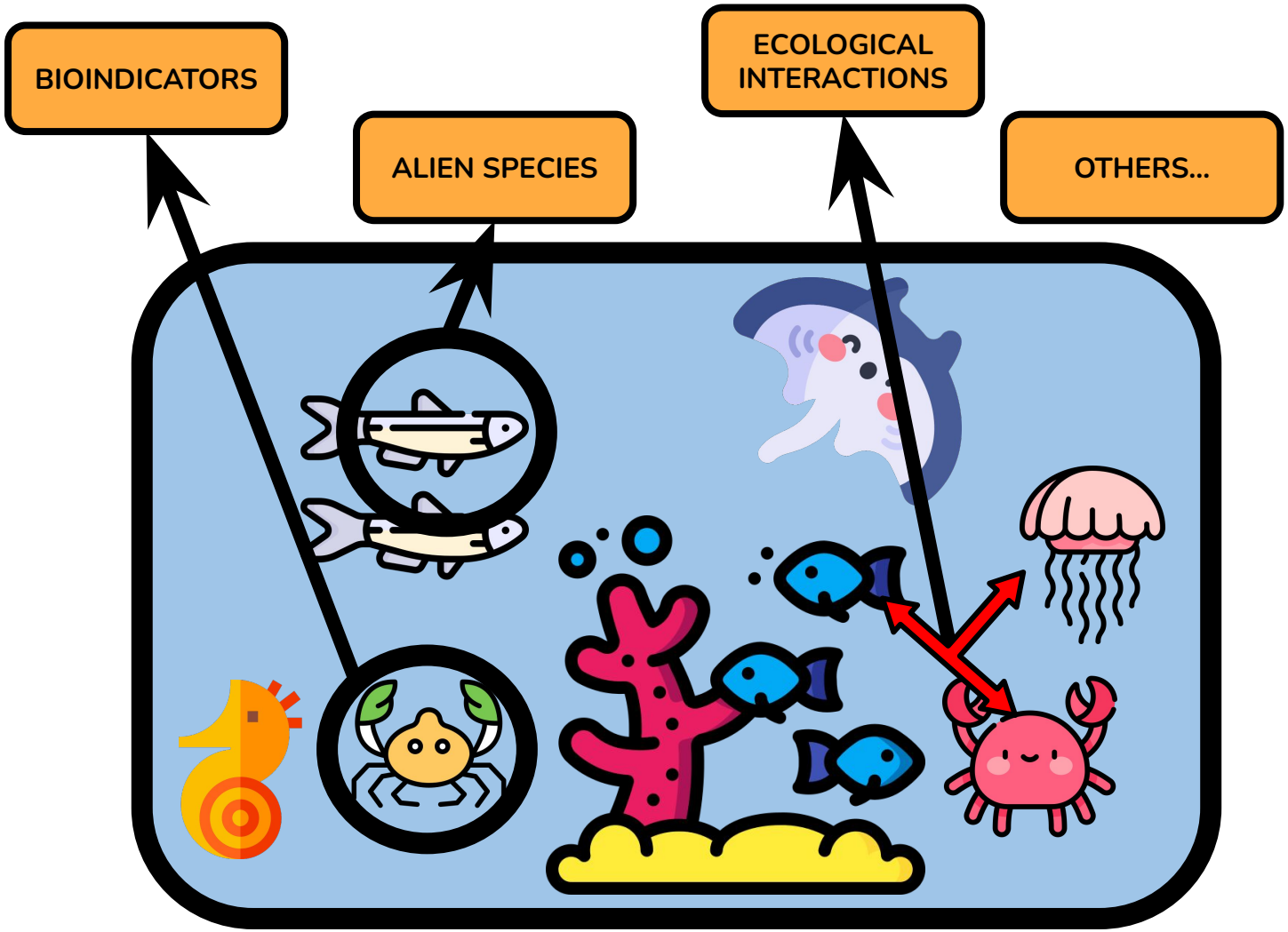
The Biodiversity of the Mediterranean Sea: Estimates, Patterns, and Threats

Marta Coll, Chiara Piroddi, Jeroen Steenbeek, Kristin Kaschner, Frida Ben Rais Lasram, Jacopo Aguzzi, Enric Ballesteros, Carlo Nike Bianchi, Jordi Corbera, Thanos Dailianis, Roberto Danovaro, Marta Estrada, Carlo Froglia, [...], Eleni Voultsiadou
[view all]

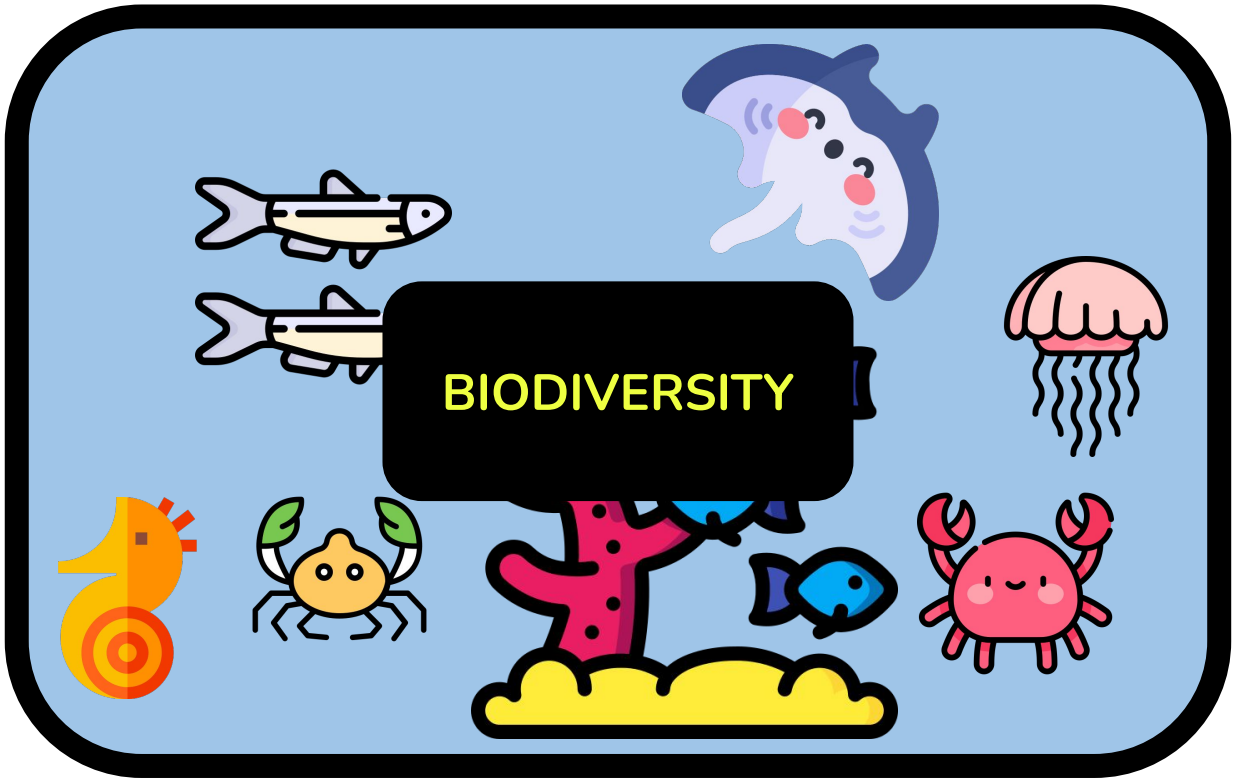
Published: August 2, 2010 • https://doi.org/10.1371/journal.pone.0011842



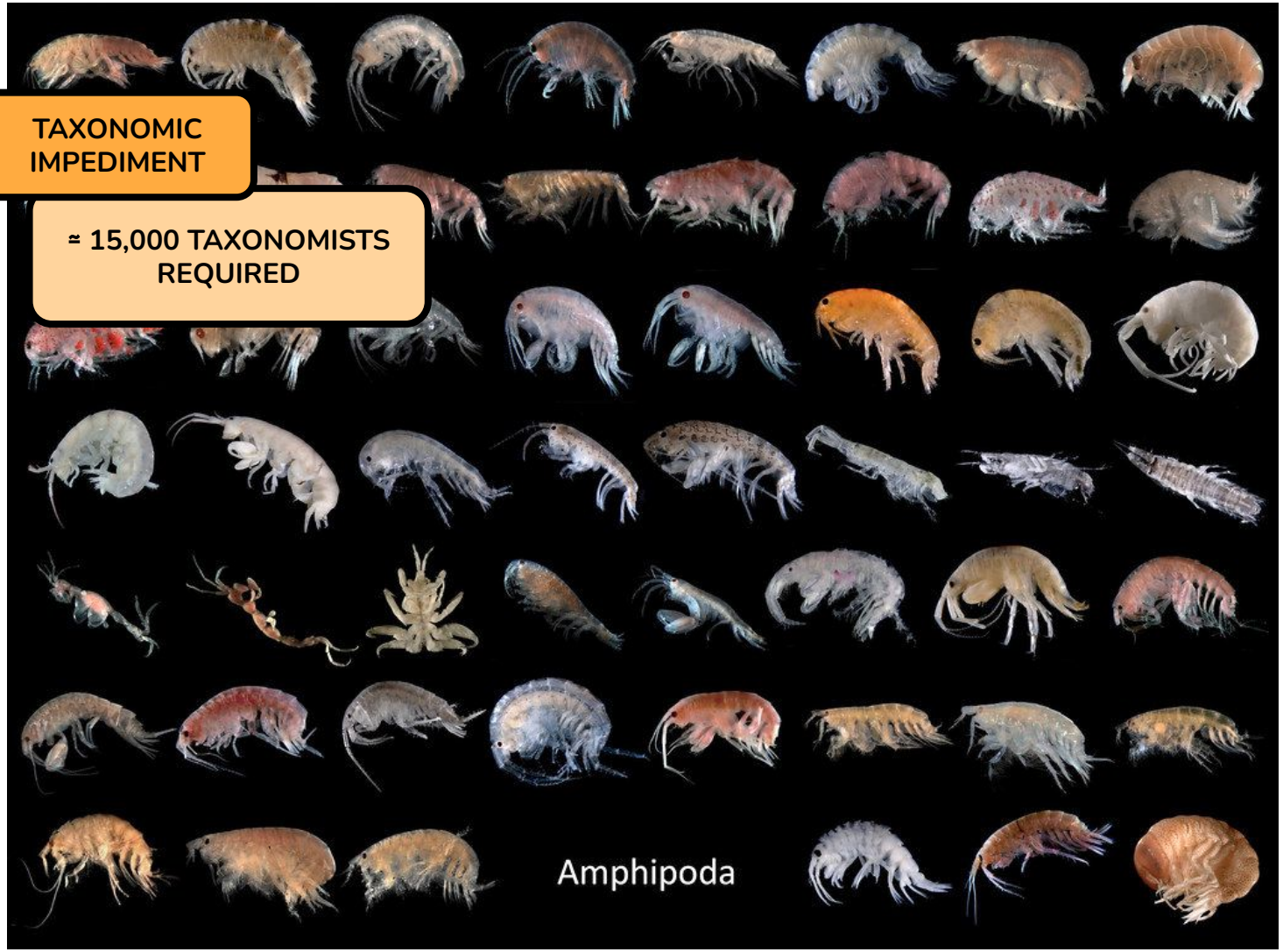
How and how fast?



How and how fast?



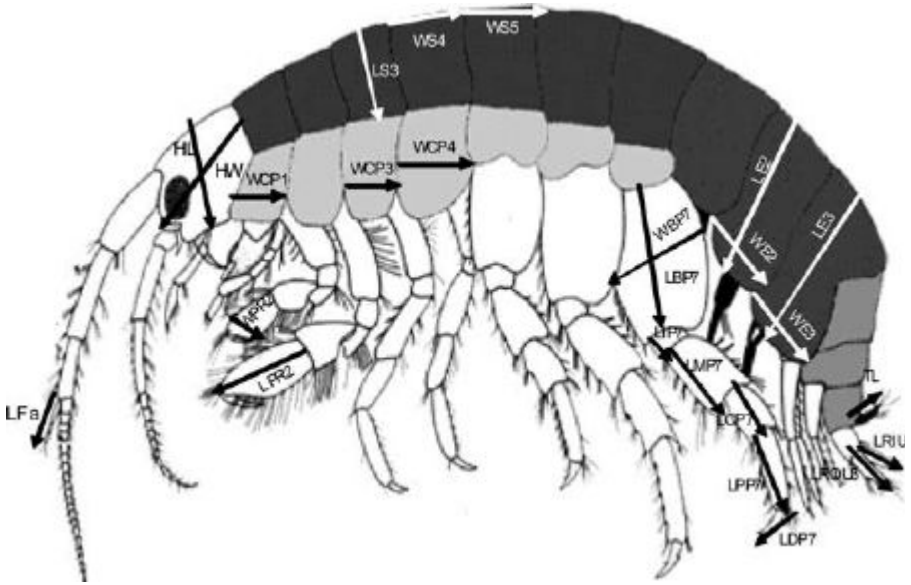
Biodiversity assessment

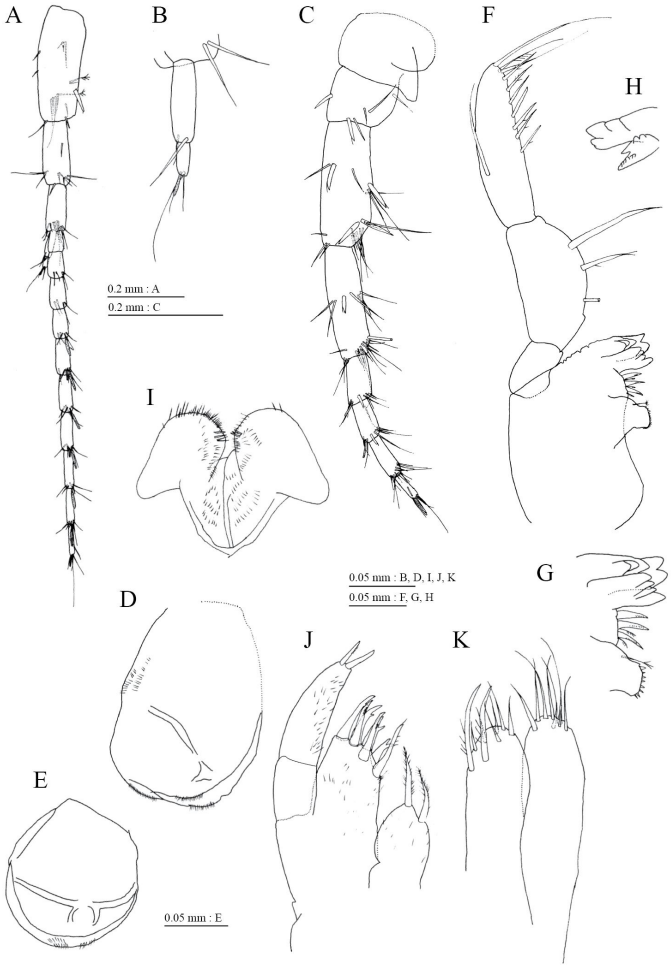


Biodiversity assessment

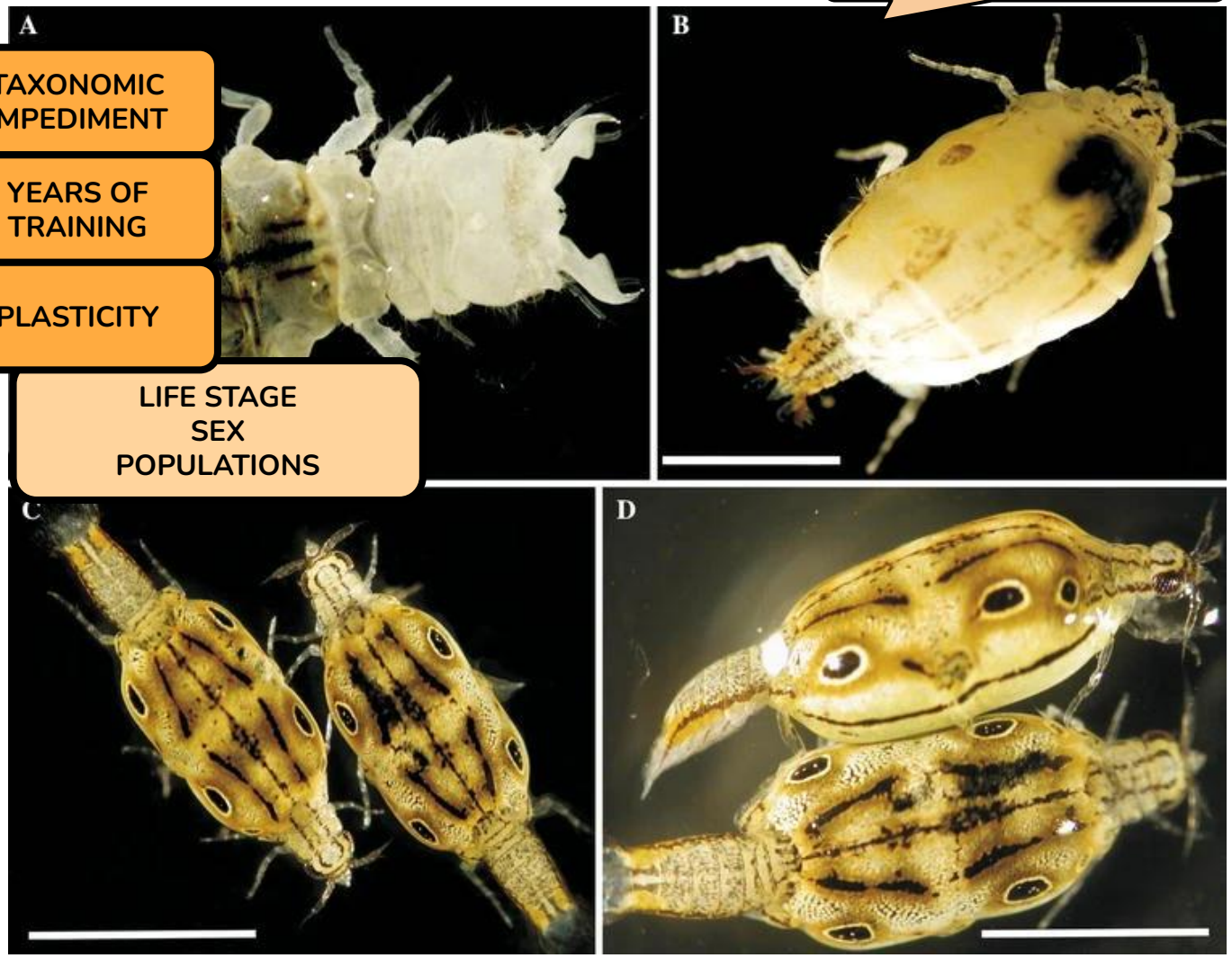
TAXONOMIC IMPEDIMENT

YEARS OF TRAINING





Biodiversity assessment



- TAXONOMIC IMPEDIMENT
- YEARS OF TRAINING
- PLASTICITY
- LIFE STAGE
- SEX
- POPULATIONS

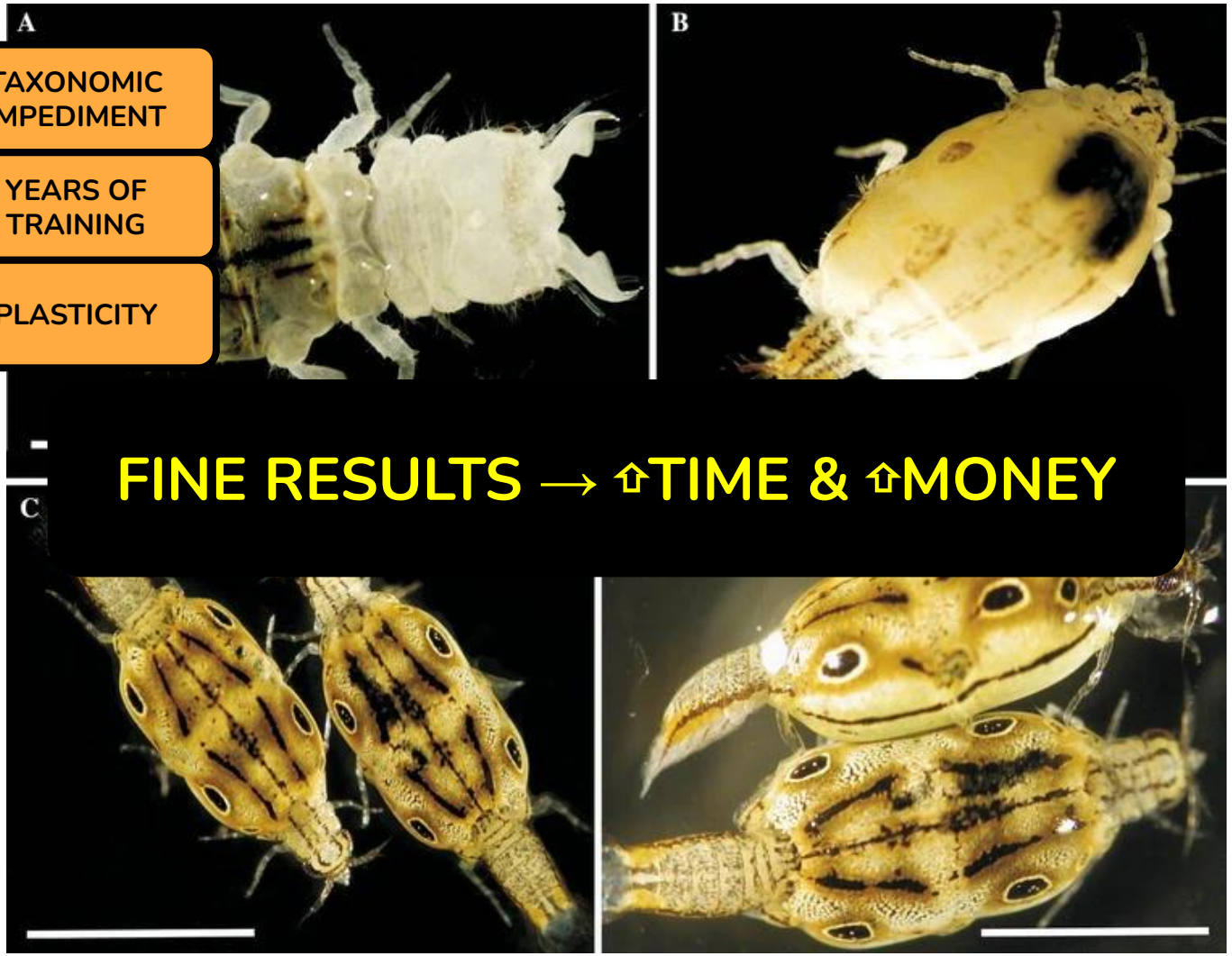
Coetzee et al. 2009

Biodiversity assessment

TAXONOMIC IMPEDIMENT

YEARS OF TRAINING

PLASTICITY

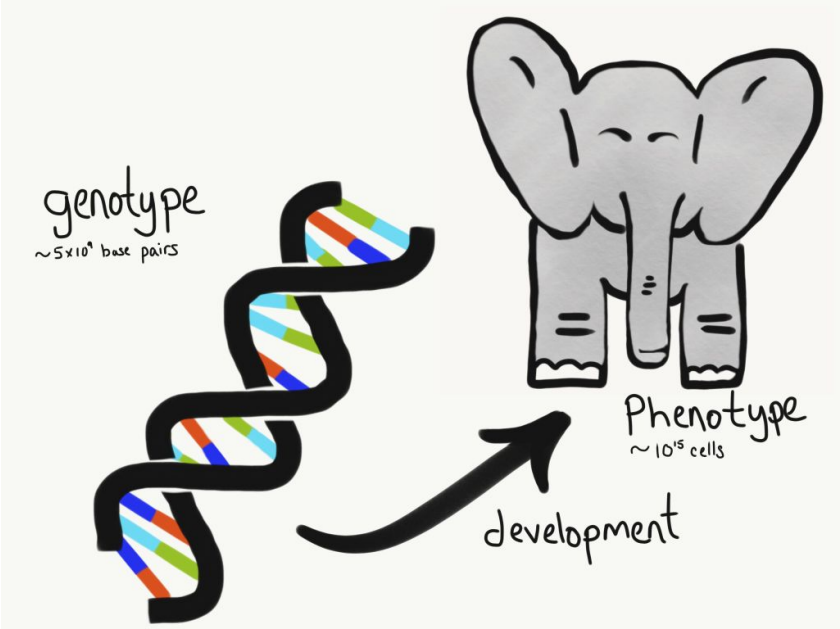


FINE RESULTS → ⇧TIME & ⇧MONEY

Coetzee et al. 2009

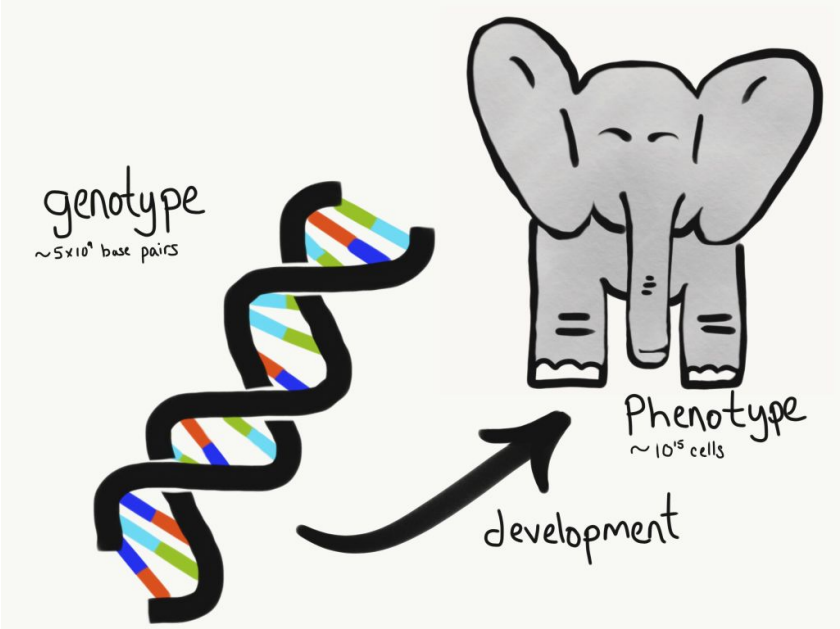
Metabarcoding

META - BARCODING

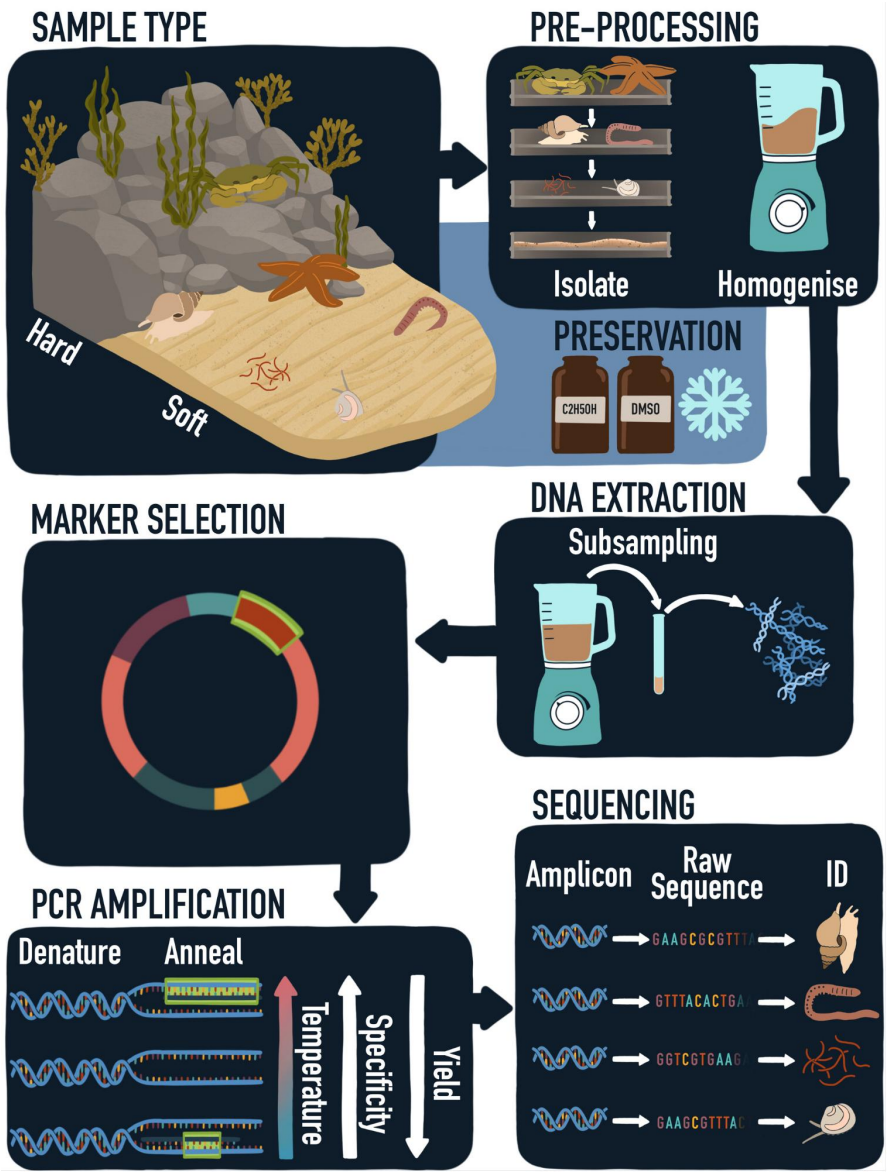


Metabarcoding

META + BARCODING

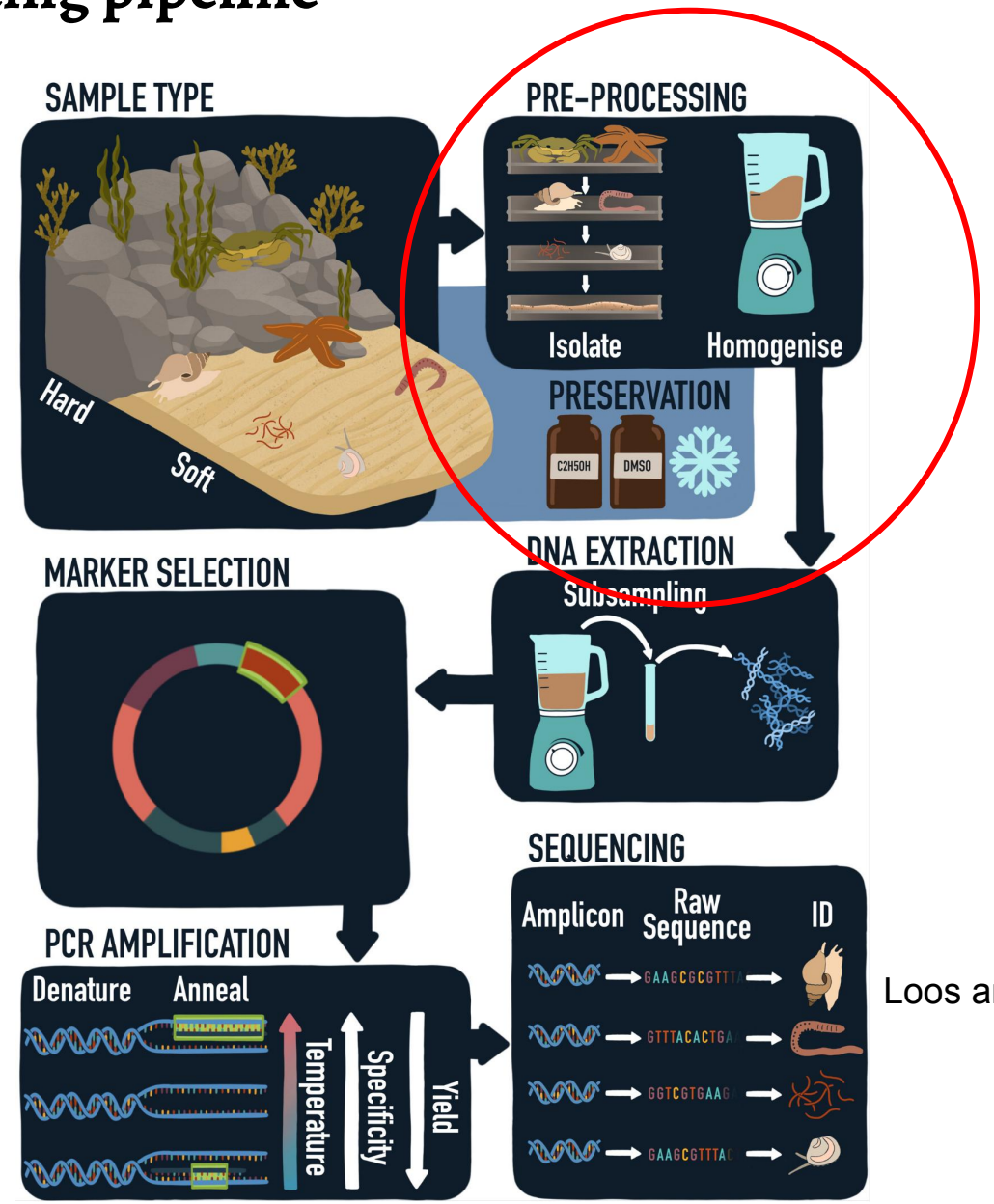


Metabarcoding pipeline



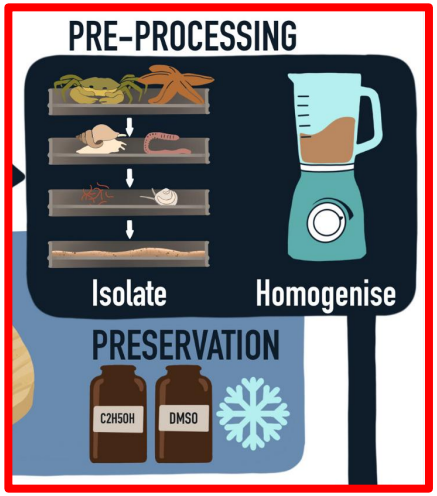
Loos and Nijland 2021

Metabarcoding pipeline



Loos and Nijland 2021

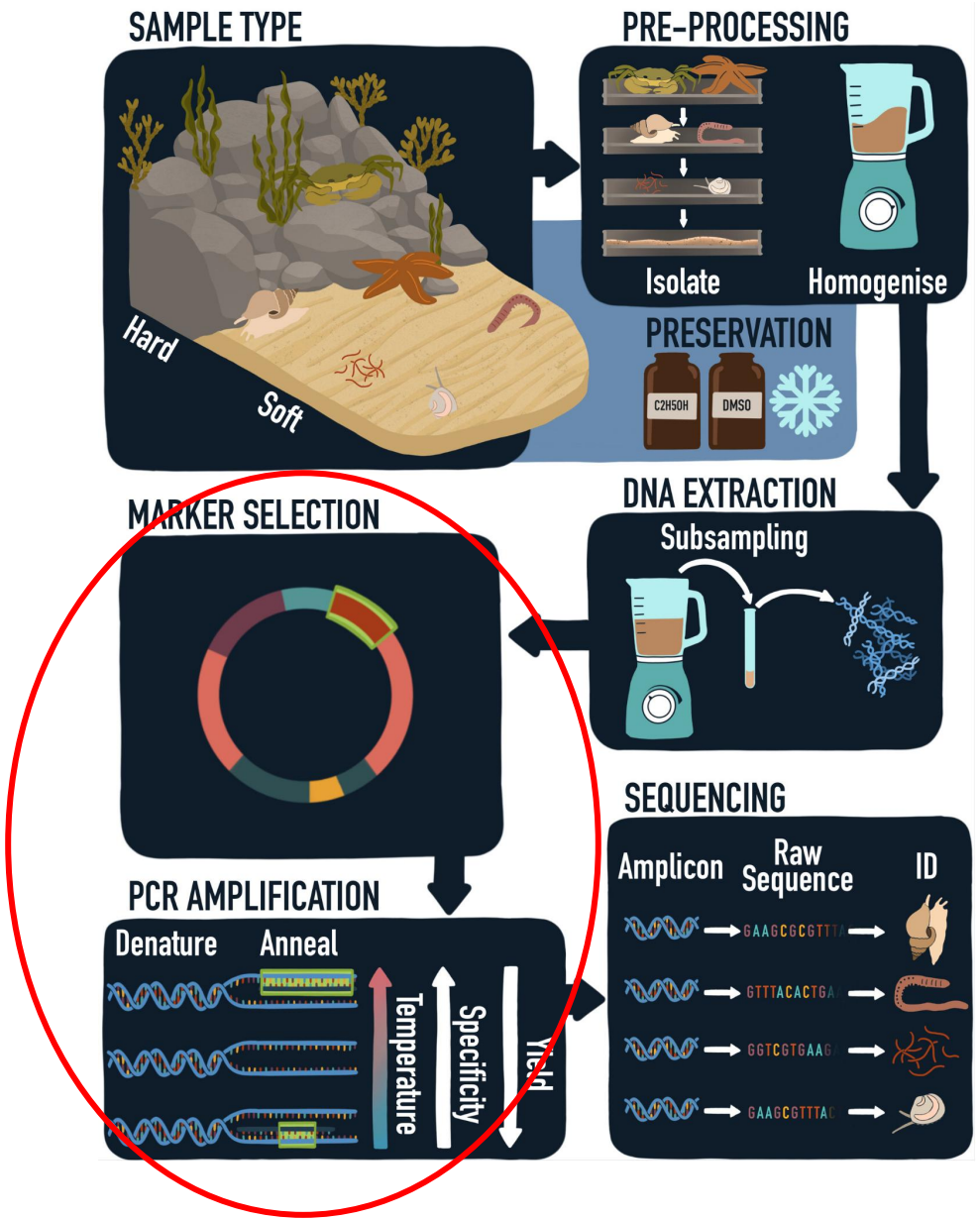
Metabarcoding pipeline



>10mm
>1mm
>63µm

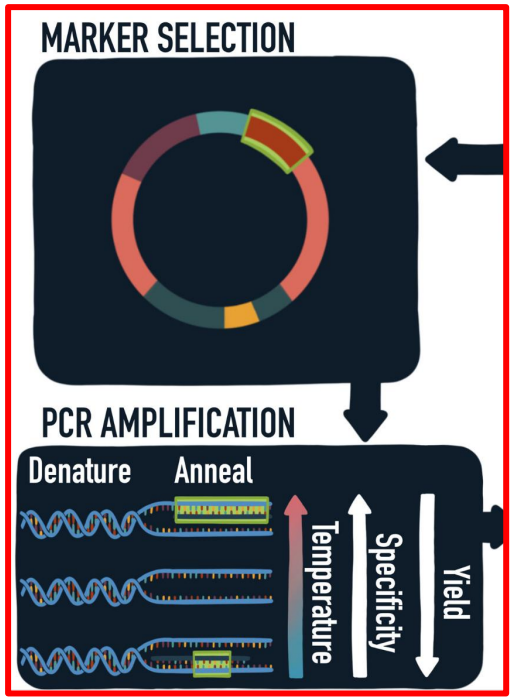
<200µm
&
>0.22µm

Metabarcoding pipeline



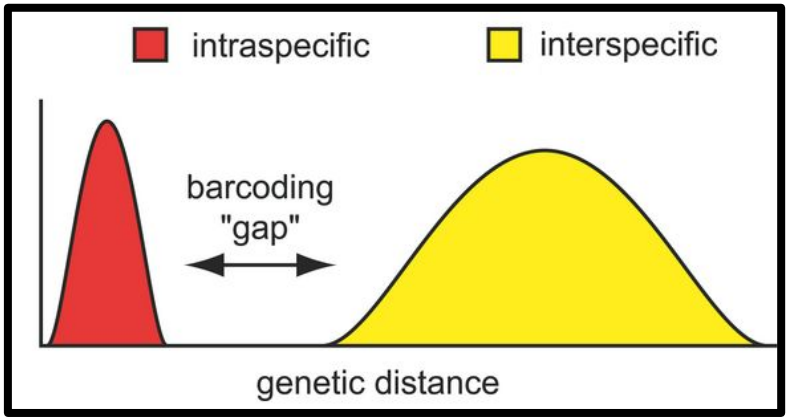
Loos and Nijland 2021

Metabarcoding pipeline

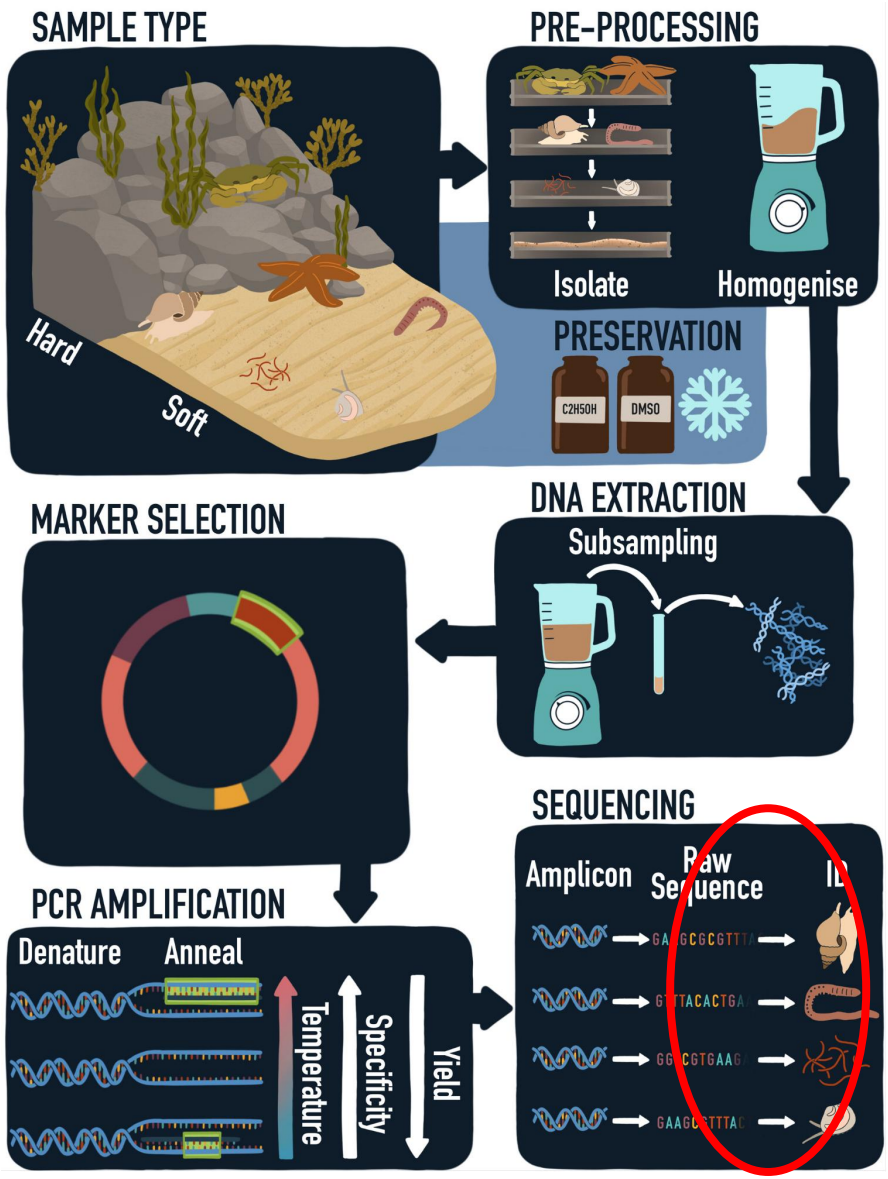


Cytochrome Oxidase I (COI)

- mitochondrial coding gene
- universal primers (Eukaryotes)
- used in phylogeography

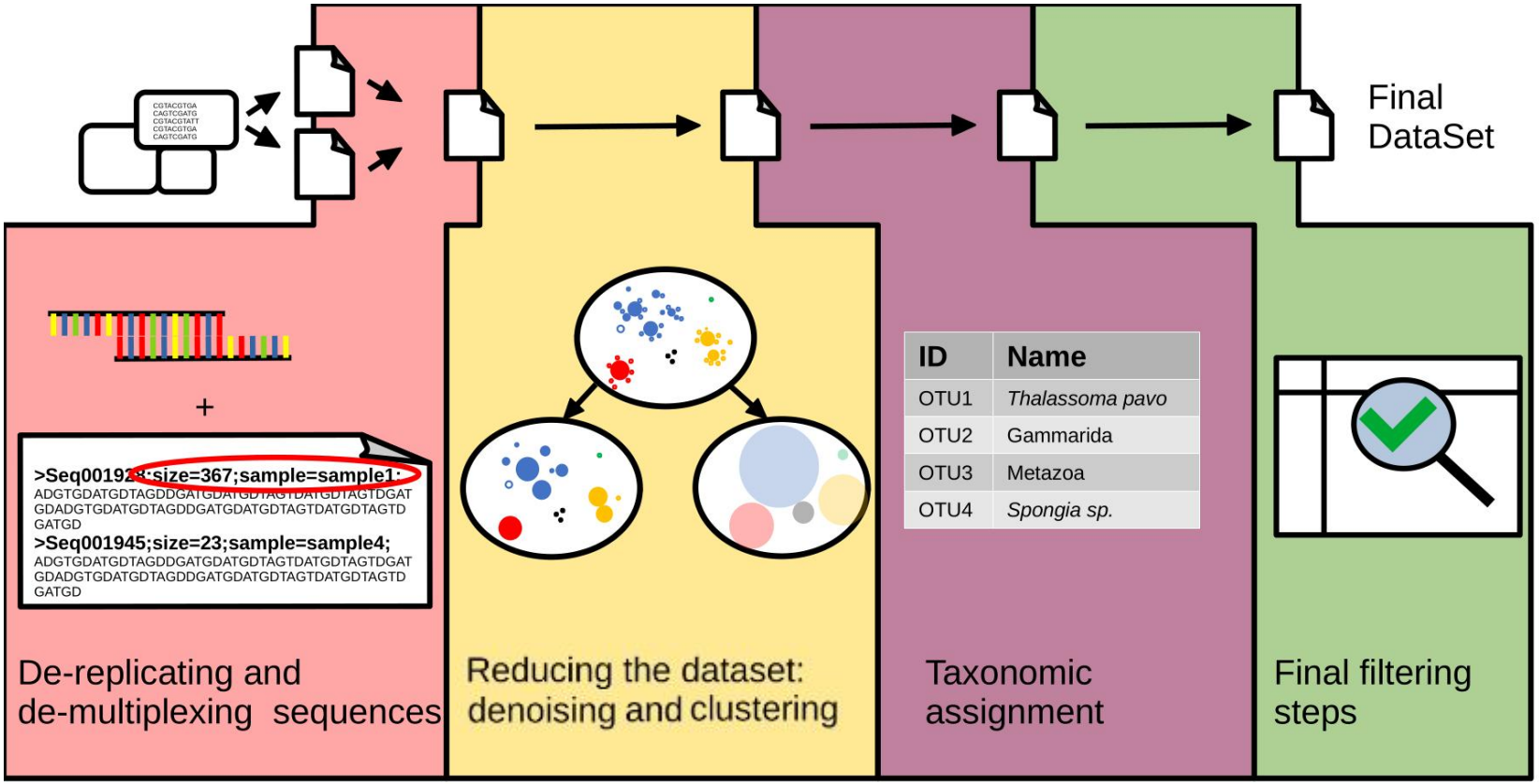


Metabarcoding pipeline



Loos and Nijland 2021

Metabarcoding pipeline



Read ⇨ each count of a sequence
Exact Sequence Variant (**ESV**) ⇨ Real sequence
Molecular Operational Taxonomic Unit (**MOTU**) ⇨ Cluster of sequences

Metabarcoding

TAXONOMIC IMPEDIMENT

YEARS OF TRAINING

PLASTICITY

Metabarcoding

TAXONOMIC IMPEDIMENT

YEARS OF TRAINING

PLASTICITY

OBJECTIVE

RELATIVELY LOW EXPERTISE

REPEATABLE & TRACEABLE

Metabarcoding

TAXONOMIC IMPEDIMENT

YEARS OF TRAINING

PLASTICITY

OBJECTIVE

RELATIVELY LOW EXPERTISE

REPEATABLE & TRACEABLE

FINE RESULTS → ⇩TIME & ⇩MONEY


OBJECTIVES

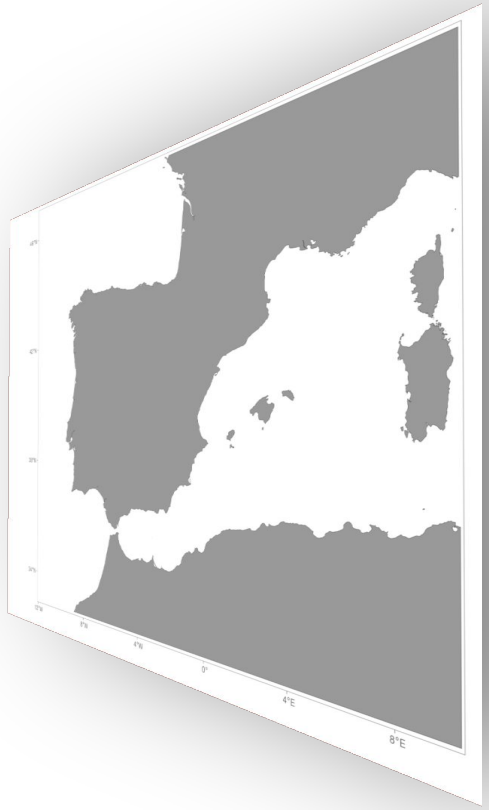
OBJECTIVES

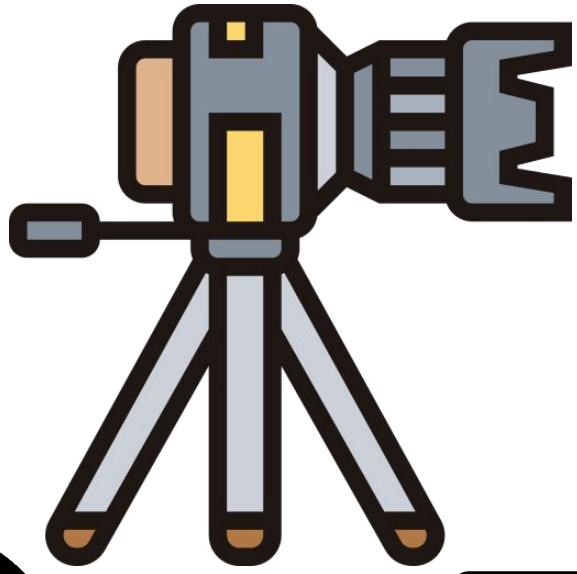
1. To explore different sampling methods to assess the marine benthic diversity.

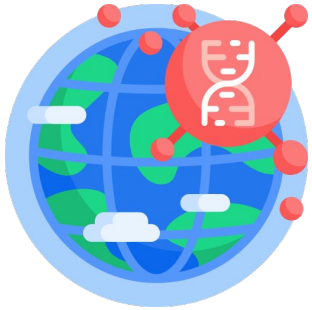
2. Develop the pipelines and software required to obtain both inter and intraspecific variability from marine benthic communities of eukaryotes using COI.

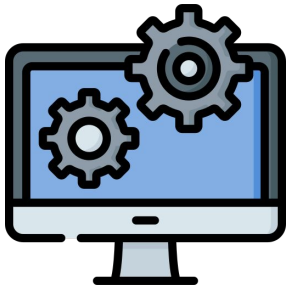
3. To apply these novel methodologies to analyse the biogeographic and metaphylogeographic patterns from benthic communities across two well described fronts in the eastern Iberian littoral.

OBJECTIVES









WHICH sample type?

WHAT do we assess?

HOW to treat data?

OBJECTIVES

4
Antich et al.2020

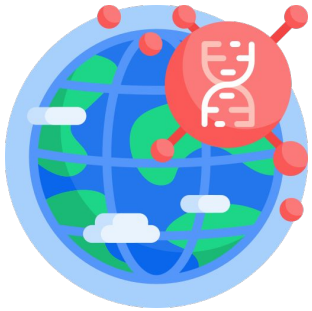
Received: 16 April 2020 | Revised: 6 August 2020 | Accepted: 7 September 2020
DOI: 10.1111/mec.15641

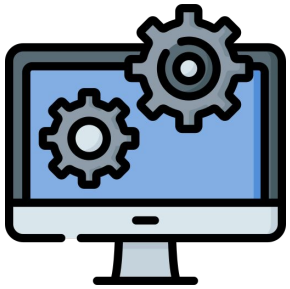
SPECIAL ISSUE

MOLECULAR ECOLOGY WILEY

Marine biomonitoring with eDNA: Can metabarcoding of water samples cut it as a tool for surveying benthic communities?

Adrià Antich[1] | Cruz Palacín[2] | Emma Cebrian[3] | Raül Golo[3] | Owen S. Wangensteen[4] | Xavier Turon[1]





WHICH sample type?

WHAT do we assess?

HOW to treat data?

OBJECTIVES

5

Turon et al.2020

Ecological Applications, 30(2), 2020, e02036

From metabarcoding to metaphylogeography: separating the wheat from the chaff

Xavier Turon,[1,4] Adrià Antich,[1] Creu Palacín,[2] Kim Præbel,[3] and Owen Simon Wangensteen[3]

OBJECTIVES

5
Turon et al.2020

6
Antich et al.2021

Antich et al. BMC Bioinformatics (2021) 22:177
https://doi.org/10.1186/s12859-021-04115-6

BMC Bioinforma

RESEARCH ARTICLE Open Access

To denoise or to cluster, that is not the question: optimizing pipelines for COI metabarcoding and metaphylogeography

Adrià Antich[1], Creu Palacin[2], Owen S. Wangensteen[3*] and Xavier Turon[1*]

7
Antich et al.2022

PeerJ

DnoisE: distance denoising by entropy. An open-source parallelizable alternative for denoising sequence datasets

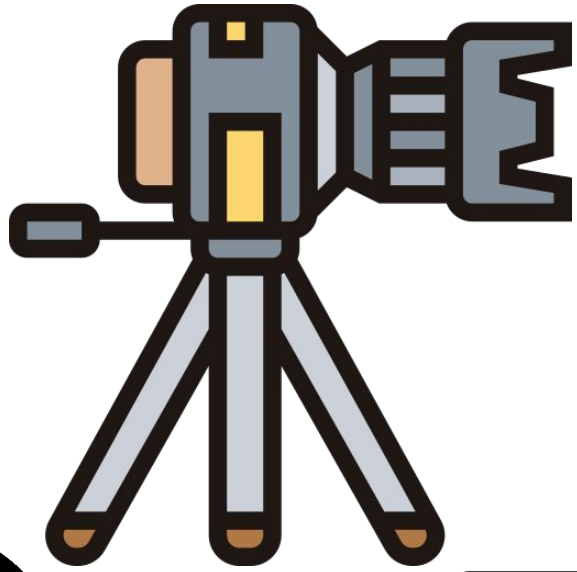
Adrià Antich[1], Creu Palacín[2], Xavier Turon[1] and Owen S. Wangensteen[3]

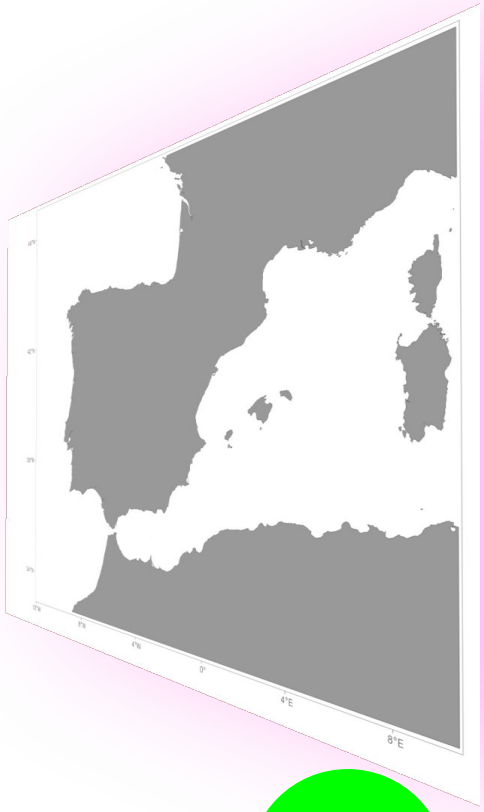
WHICH sample type?

WHAT do we assess?

HOW to treat data?

OBJECTIVES





8

Antich et al.(press)

1 **Metabarcoding reveals high-resolution biogeographic and**

2 **metaphylogeographic patterns through marine barriers**

3 Antich A[1], Palacin C[2], Zarcero J[1], Turon X[1]*, Wangensteen OS[3]*

WHAT do we assess?

WHICH SAMPLE TYPE?

Environmental DNA

4

Antich et al.2020

WHICH SAMPLE TYPE?

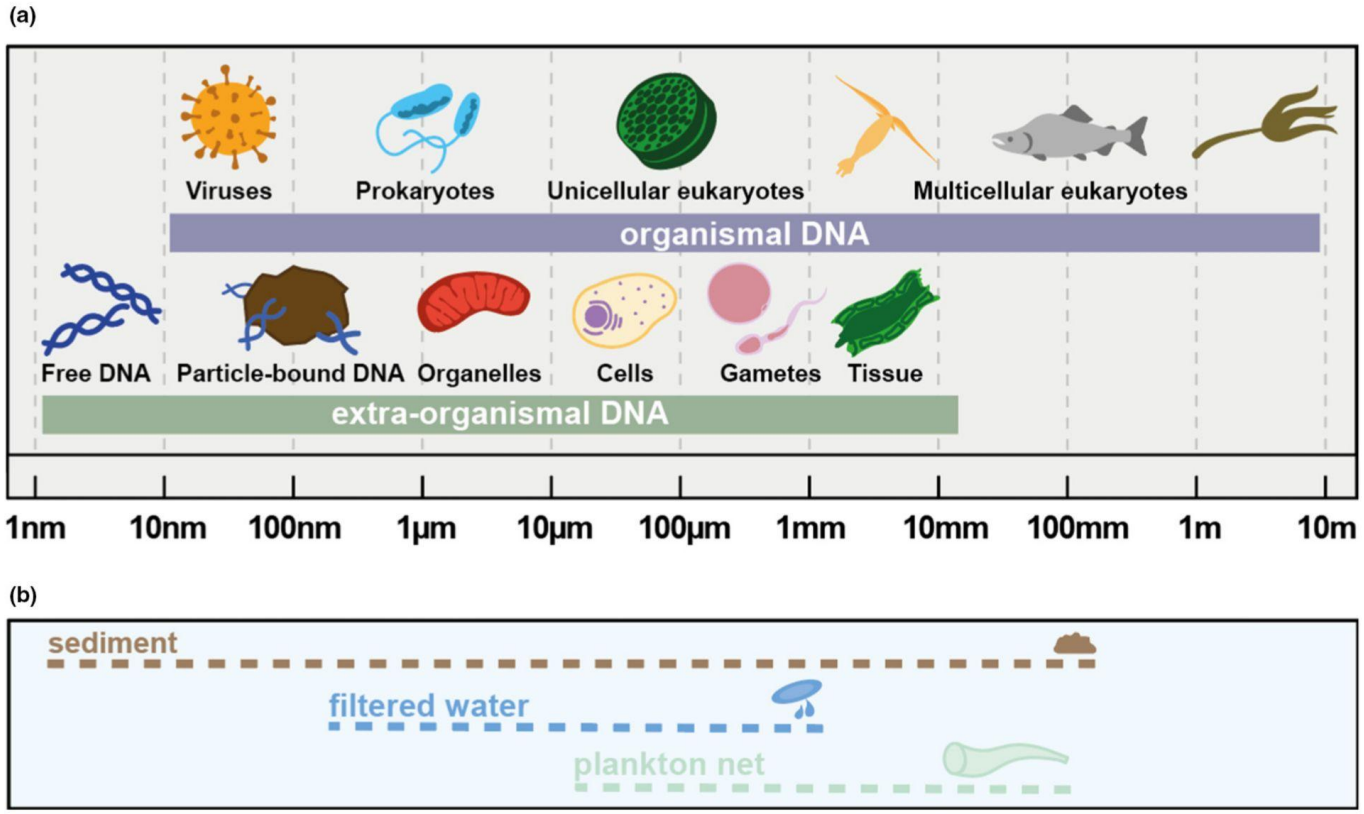
The DNA that can be obtained **from environmental samples** (i.e. soil, water, faeces, etc.) **without isolating any organism**, as **opposed to community DNA** (comDNA) that is obtained from bulk samples of previously **isolated organisms**.

(Deiner et al., 2017; Taberlet et al., 2012)

Environmental DNA

4
Antich et al.2020

WHICH SAMPLE TYPE?

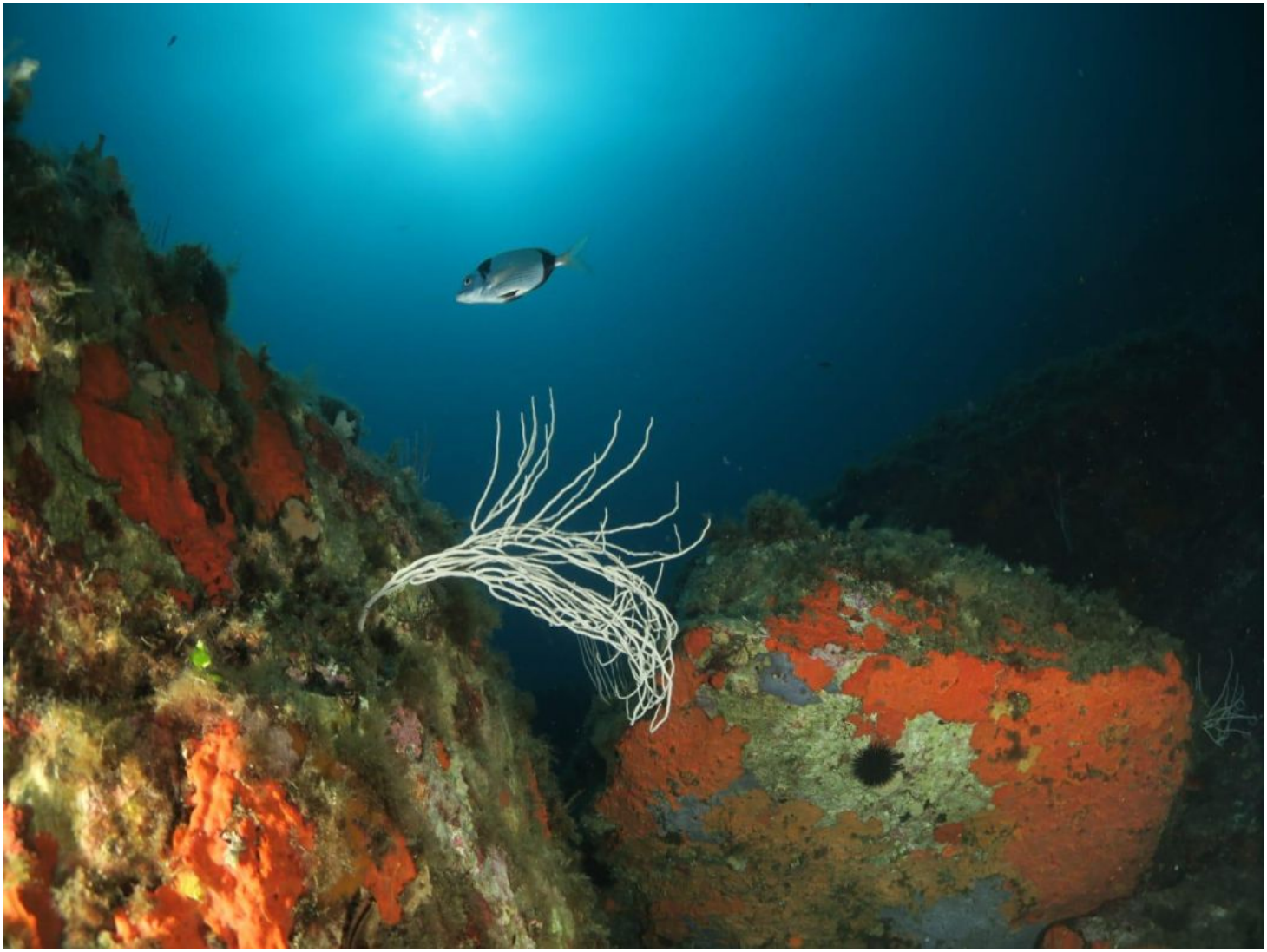


Rodriguez-Ezpeleta et al. 2021

Can we use Environmental DNA from water samples?

4

Antich et al.2020

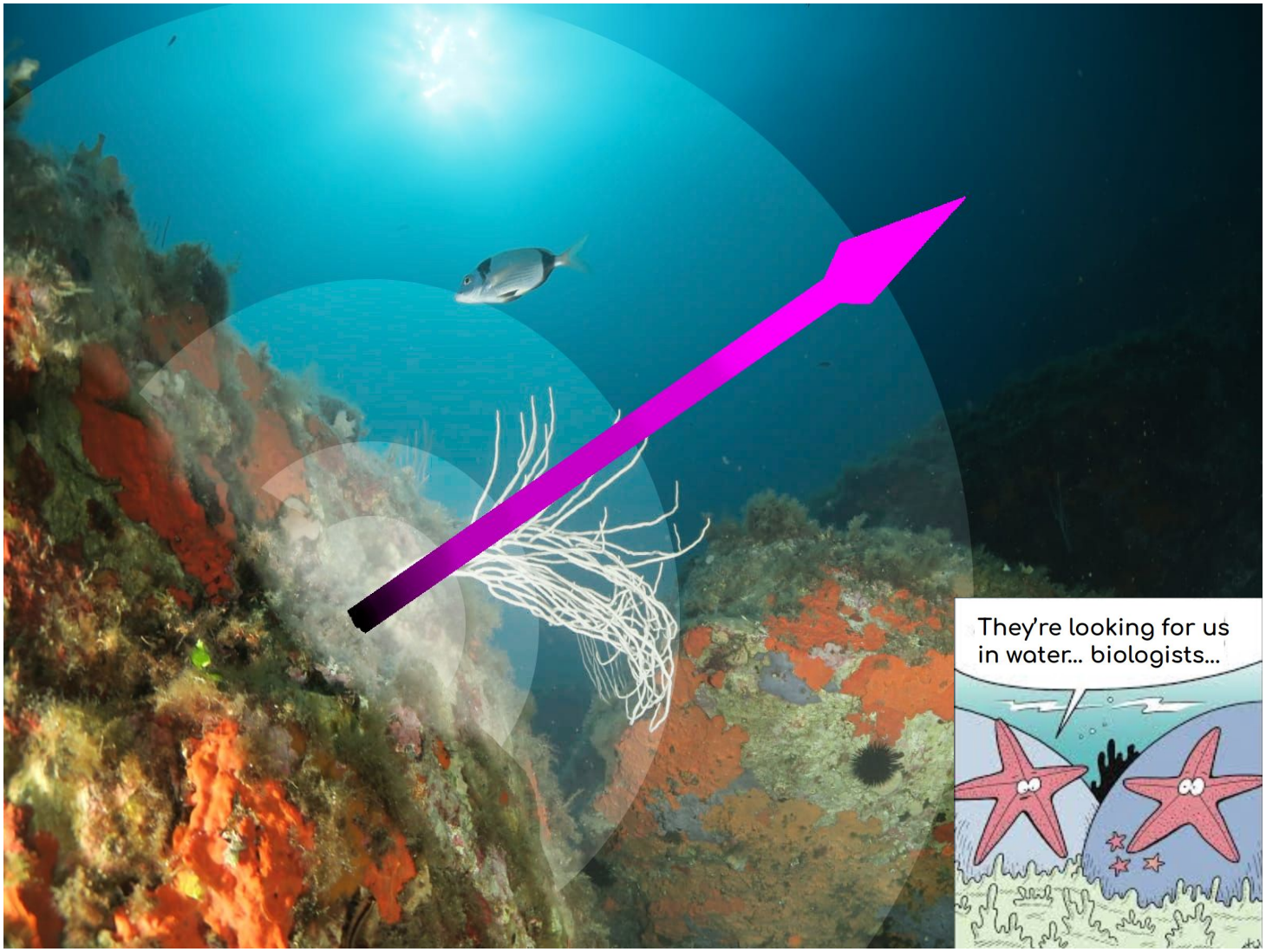


WHICH SAMPLE TYPE?

Can we use Environmental DNA from water samples?

4

Antich et al.2020

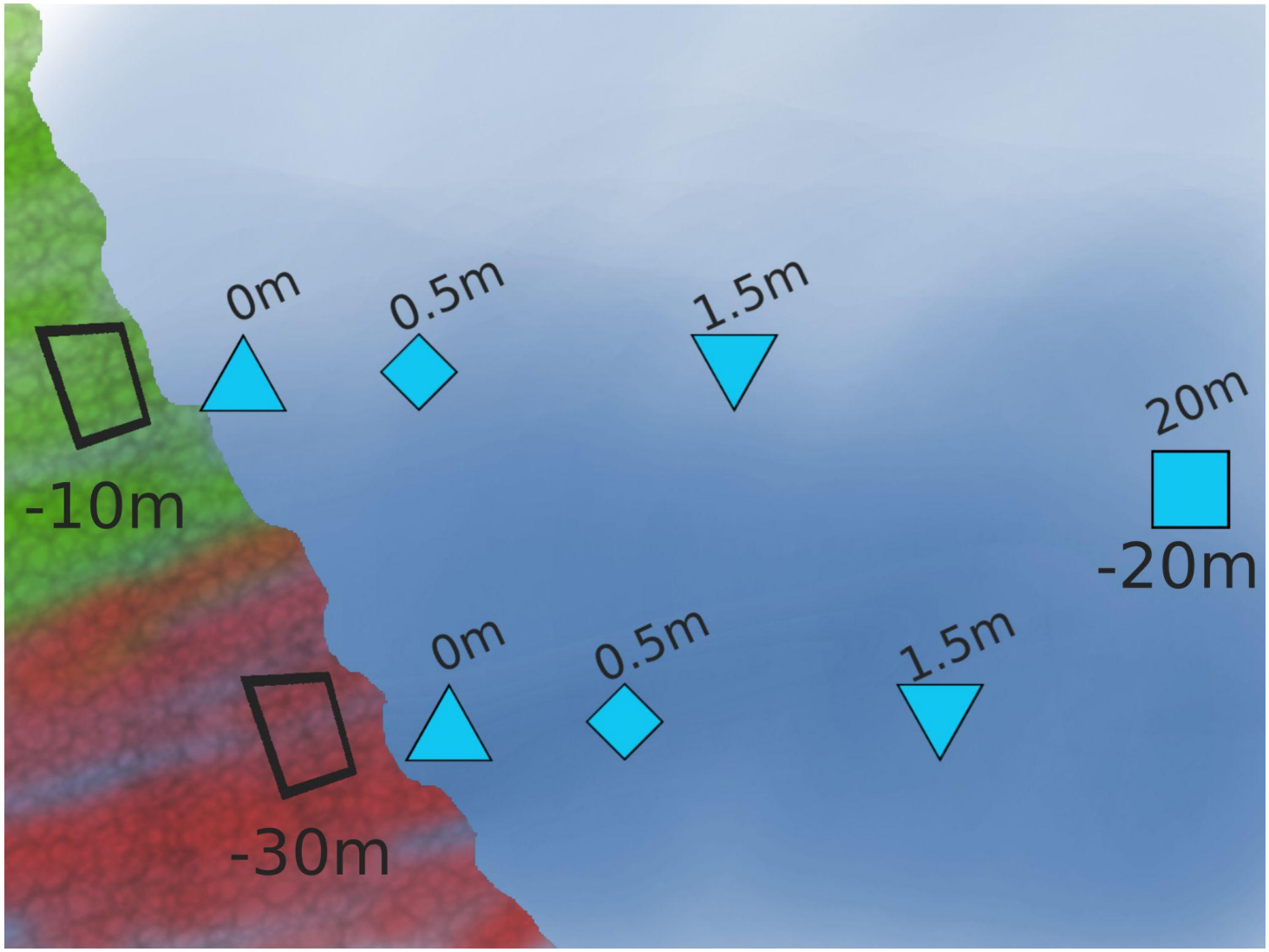


WHICH SAMPLE TYPE?

Can we use Environmental DNA from water samples?

4

Antich et al.2020

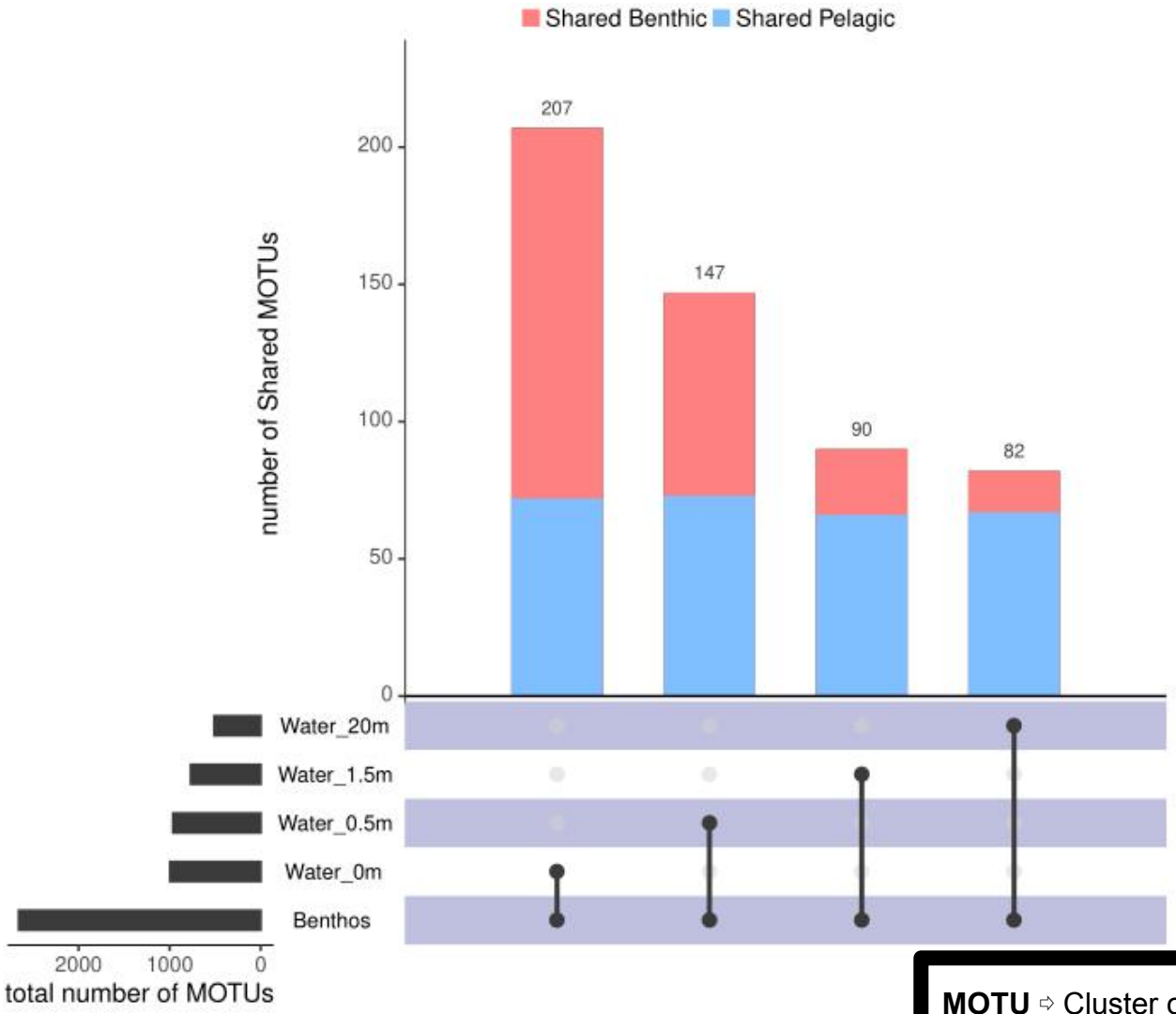


WHICH SAMPLE TYPE?

Can we use Environmental DNA from water samples?

4

Antich et al.2020



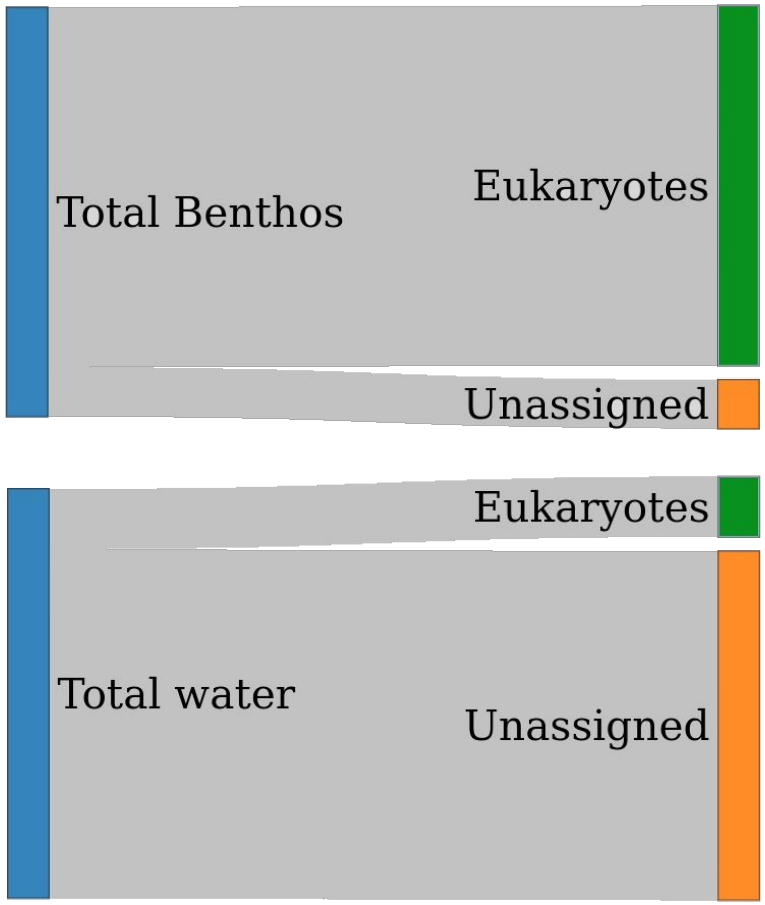
MOTU ⇨ Cluster of sequences

Can we use Environmental DNA from water samples?

4

Antich et al.2020

PROPORTION OF READS



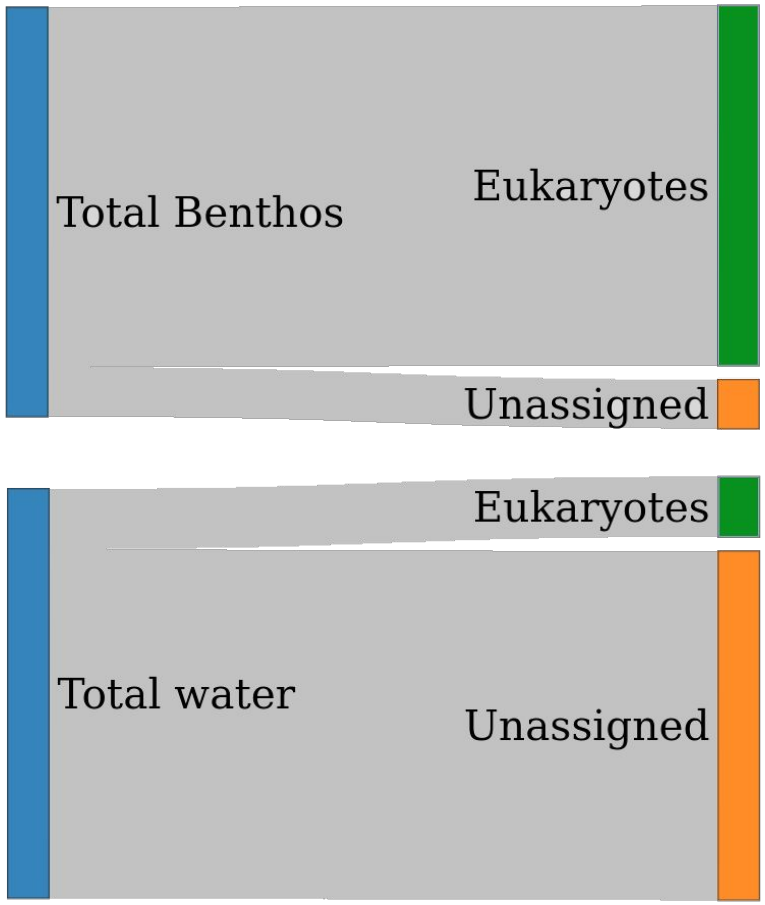
Read ⇨ each count of a seq.

Can we use Environmental DNA from water samples? ⇨ NO

4

Antich et al.2020

PROPORTION OF READS



Read ⇨ each count of a seq.

WHAT DO WE ASSESS?

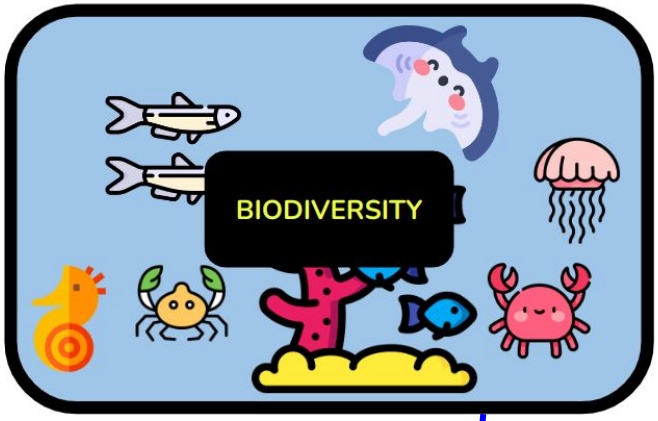
Biodiversity with metabarcoding

Biodiversity with metabarcoding

WHAT DO WE ASSESS?

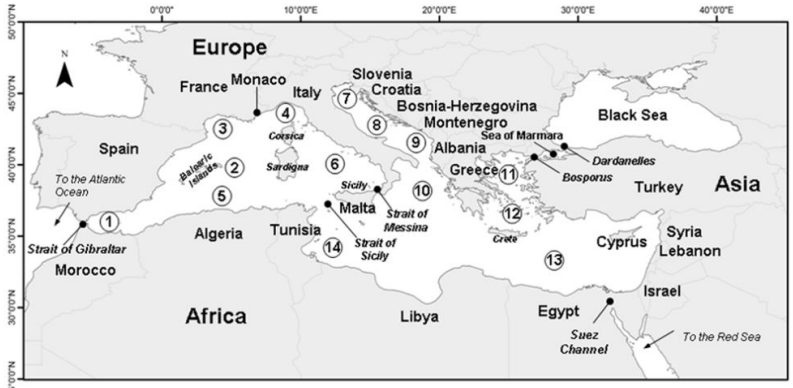
How and how fast?

... and WHERE?



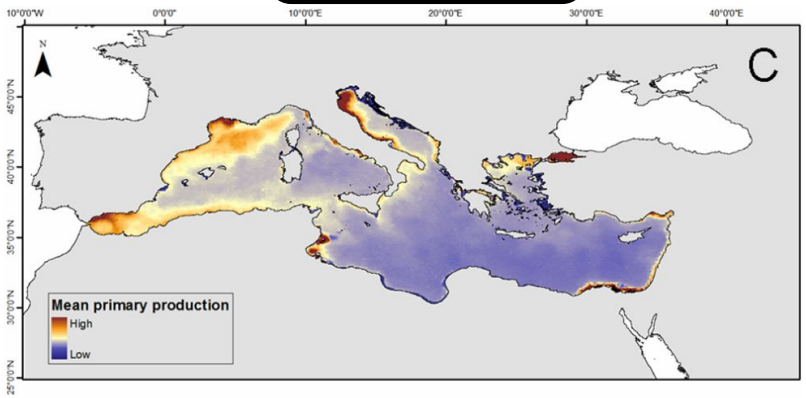
Marine barriers

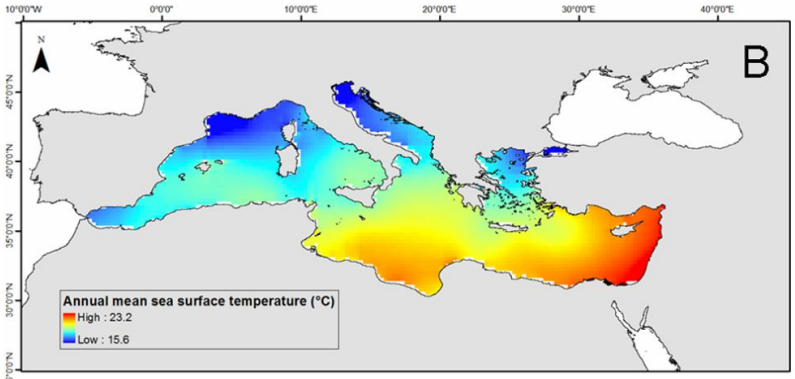
WHAT DO WE ASSESS?



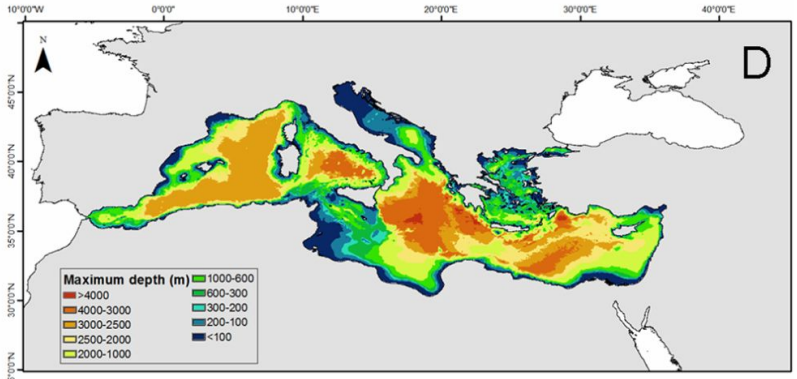
1. Alboran Sea, 2. Balearic Sea, 3. Gulf of Lions, 4. Ligurian Sea, 5. Algeria and Tunisian waters, 6. Tyrrhenian Sea, 7. North Adriatic Sea, 8. Central Adriatic Sea, 9. South Adriatic Sea, 10. Ionian Sea, 11. North Aegean Sea, 12. South Aegean Sea, 13. Levant Sea, 14. Gulf of Gabès.

PRIMARY PRODUCTION





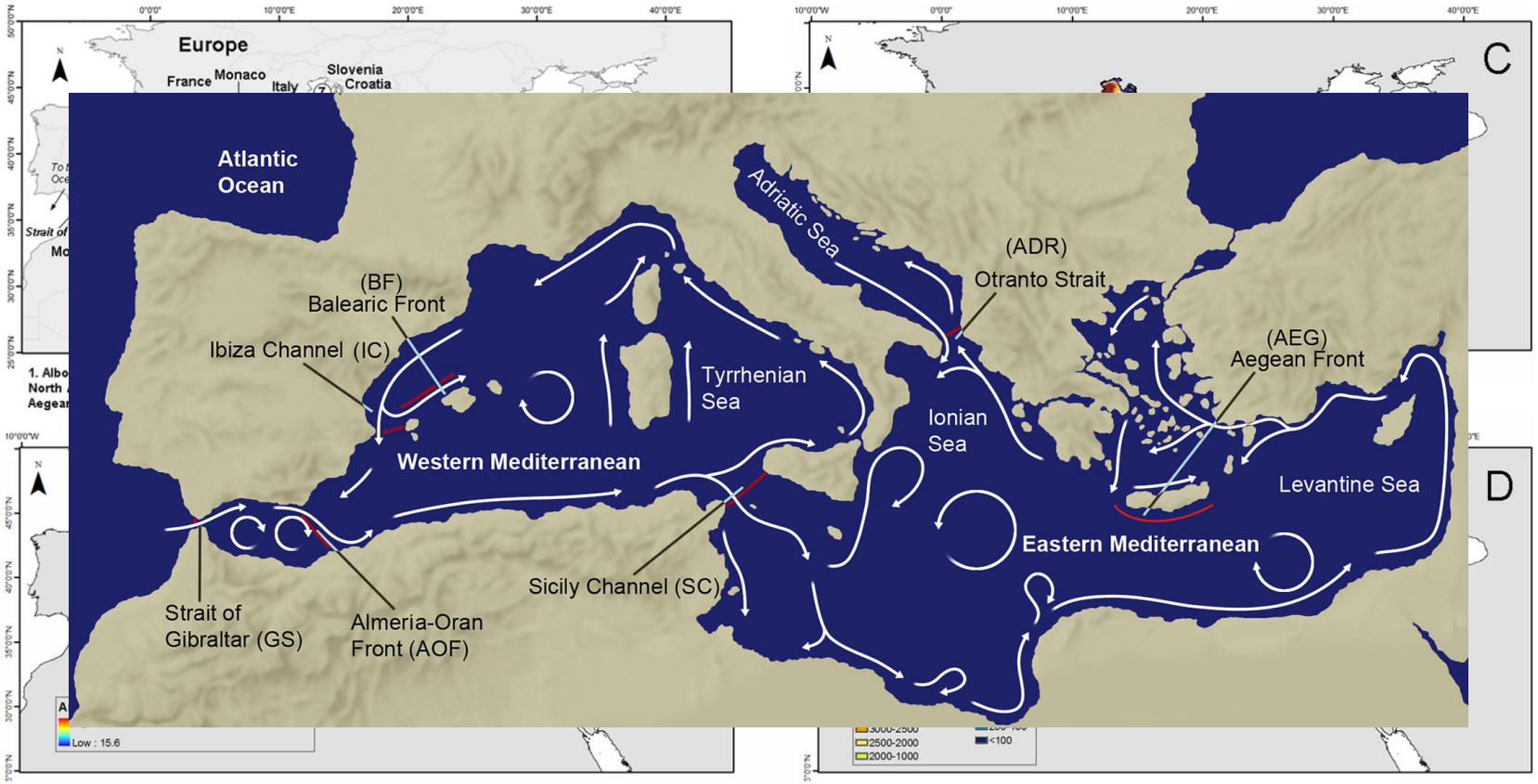
TEMPERATURE



MAX DEPTH

Coll et al. 2010

Marine barriers



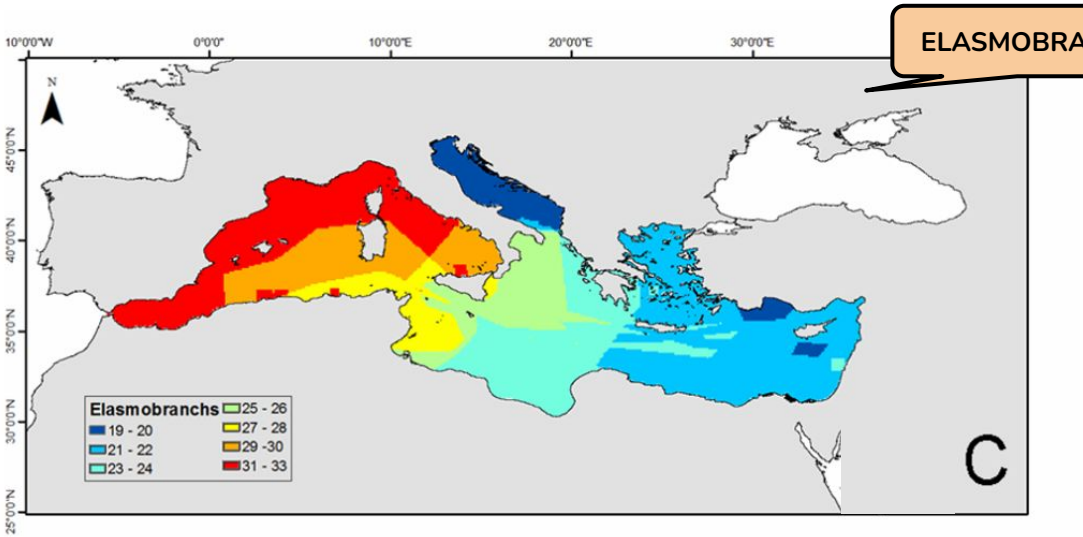
Pascual et al. 2017

Coll et al. 2010

Marine barriers

WHAT DO WE ASSESS?

INTER-SPECIES



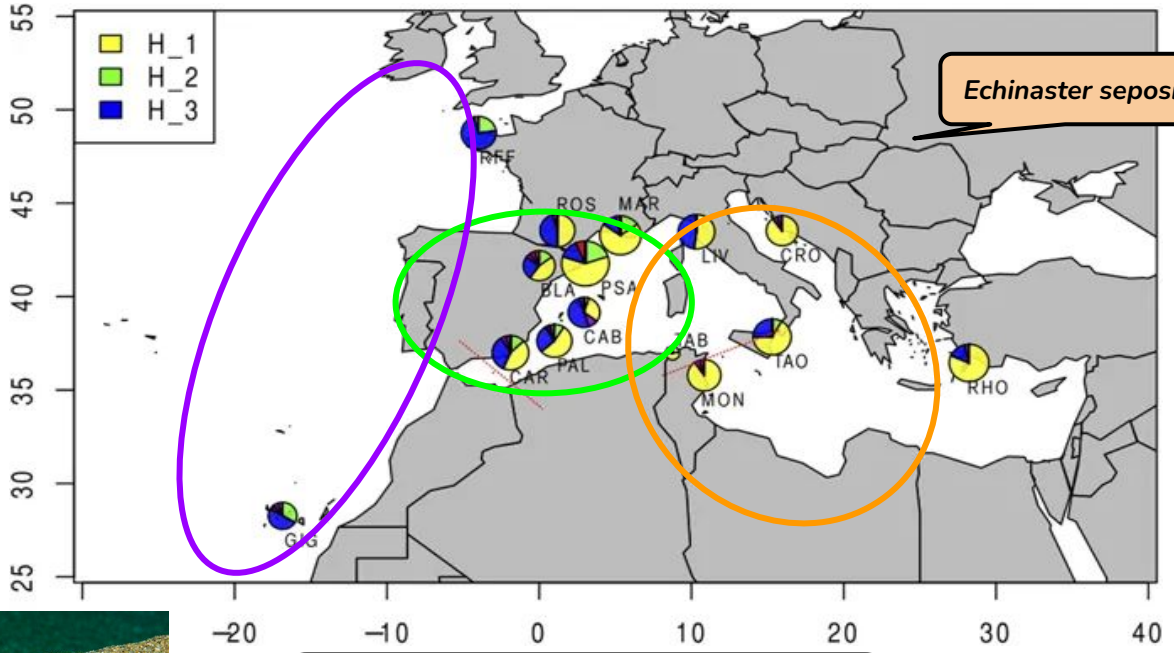
BIOGEOGRAPHY

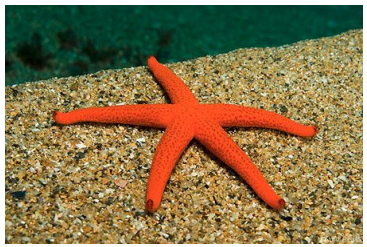
Coll et al. 2010

Marine barriers

WHAT DO WE ASSESS?

INTRA-SPECIES





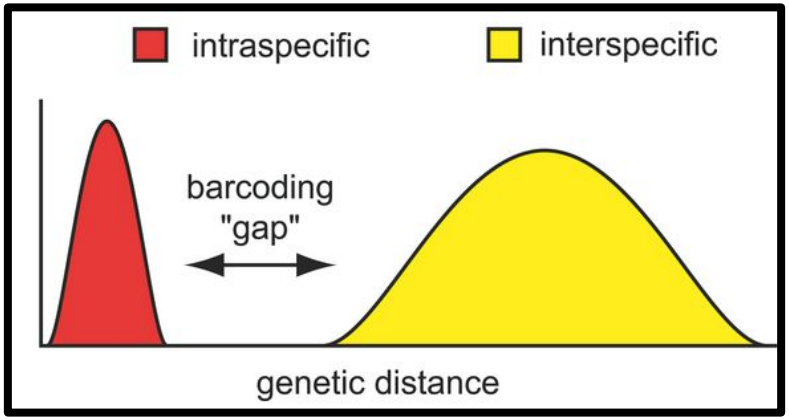
PHYLOGEOGRAPHY

Garcia-Cisneros et al. 2016

5

Turon et al.2020

META - BARCODING



MOTU

ESV

proxy for

Species

Haplotype

MOTU ⇨ Cluster of sequences
ESV ⇨ Real Sequence

5

Turon et al.2020

META - BARCODING + PHYLOGEOGRAPHY

5

Turon et al.2020

META - BARCODING + PHYLOGEOGRAPHY

↓

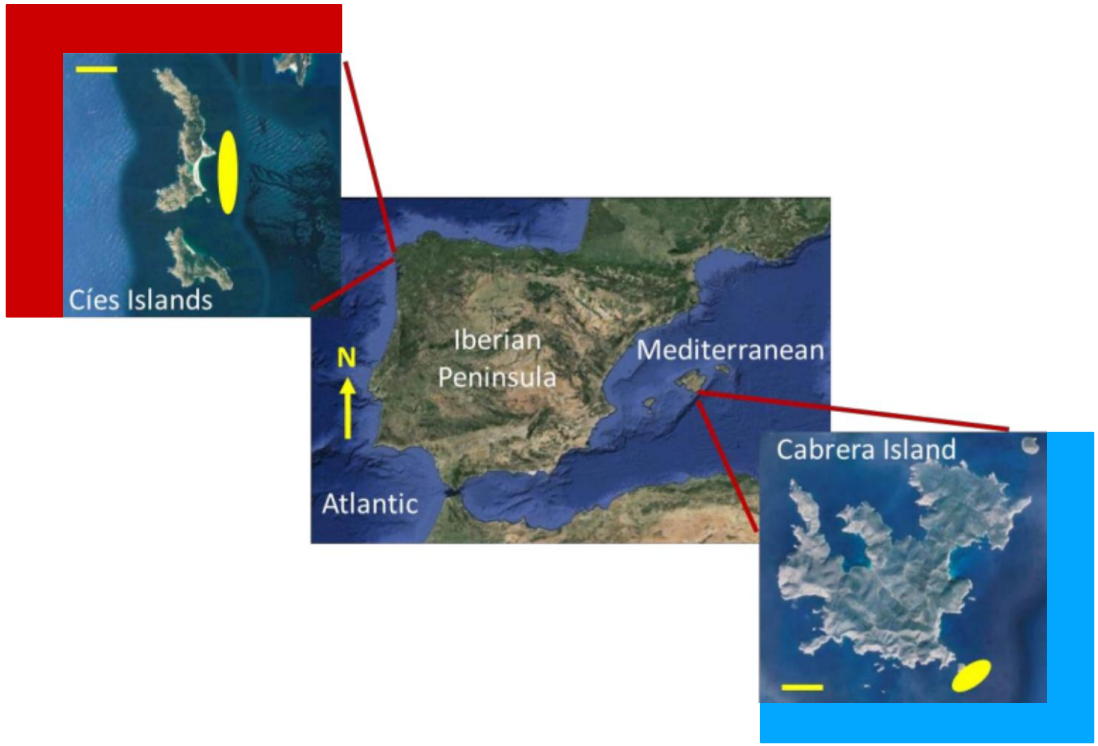
METAPHYLOGEOGRAPHY

Turon et al. 2020

Metaphylogeography

5

Turon et al.2020



WHAT DO WE ASSESS?

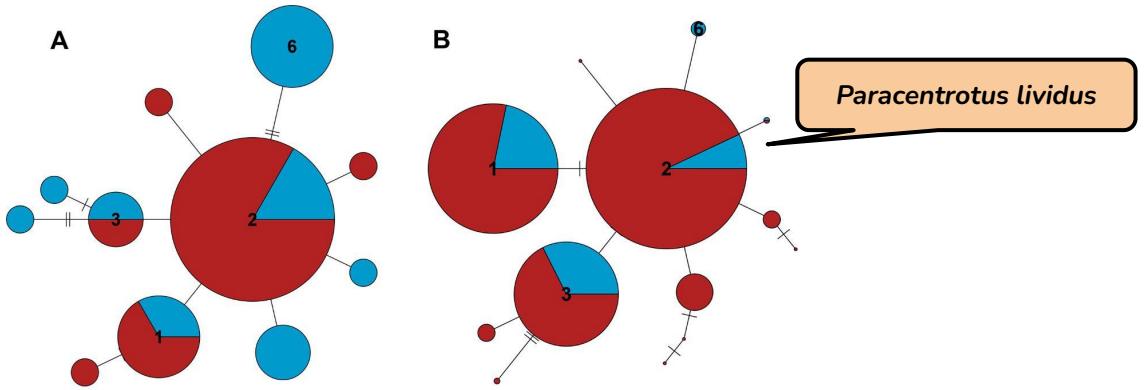
Metaphylogeography

5

Turon et al.2020

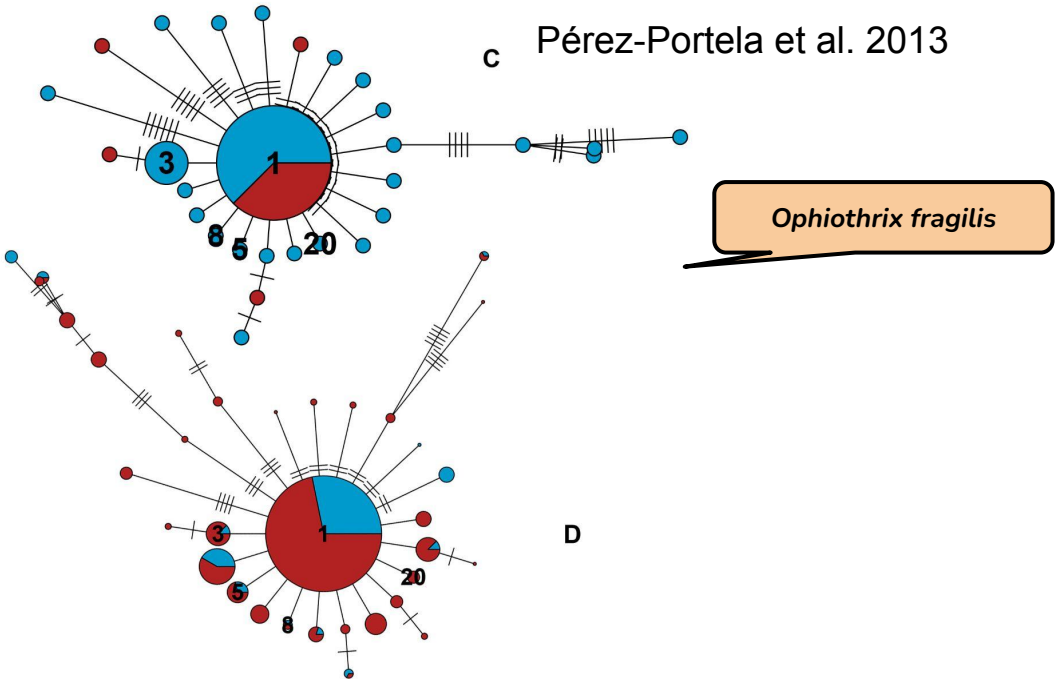
WHAT DO WE ASSESS?

Duran et al. 2004



Paracentrotus lividus

Pérez-Portela et al. 2013



Ophiothrix fragilis

ESV ⇨ Real Sequence
MOTU ⇨ Cluster of sequences

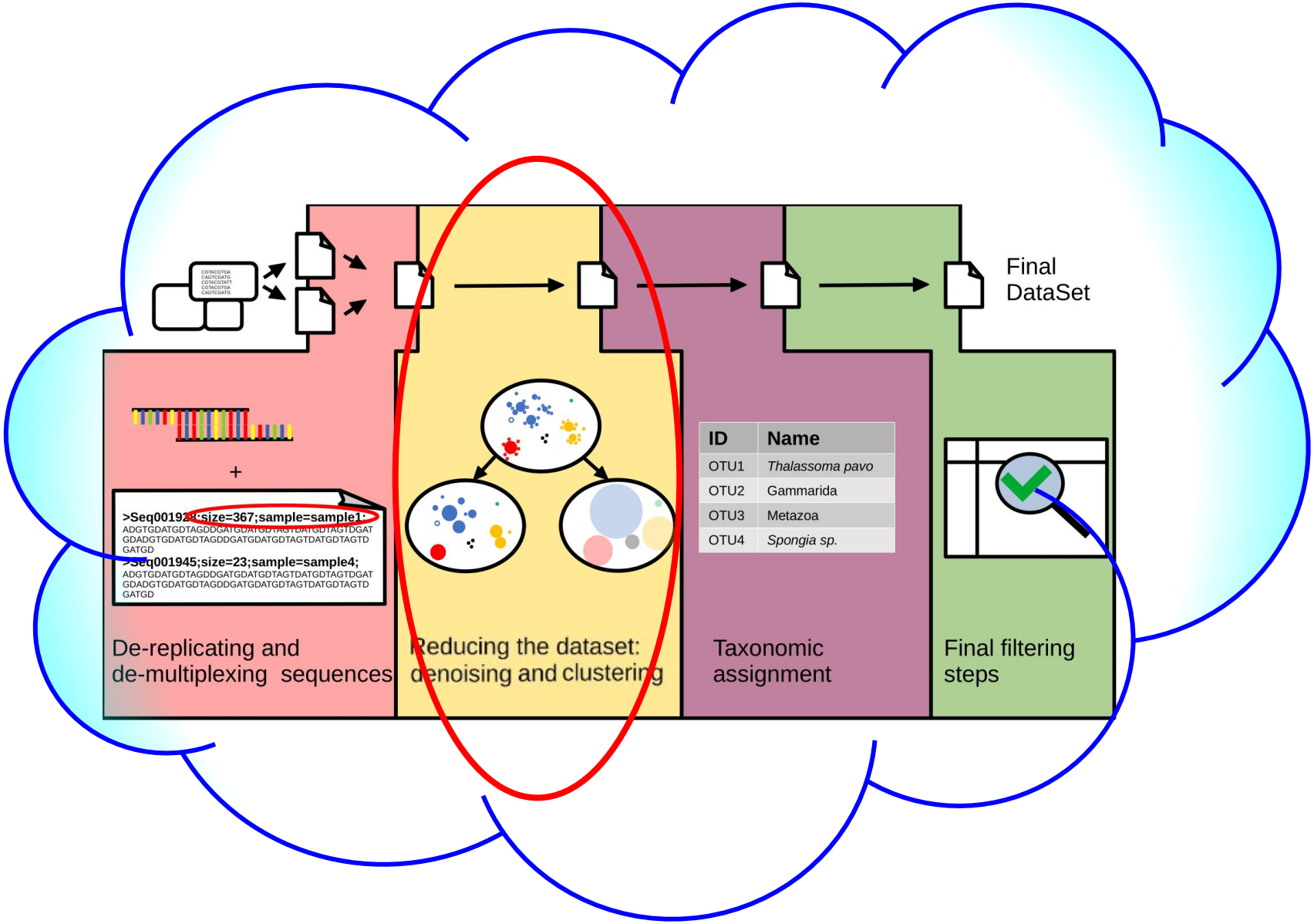
HOW TO TREAT DATA?

Remember metabarcoding pipeline...

5 Turon et al.2020

6 Antich et al.2021

7 Antich et al.2022

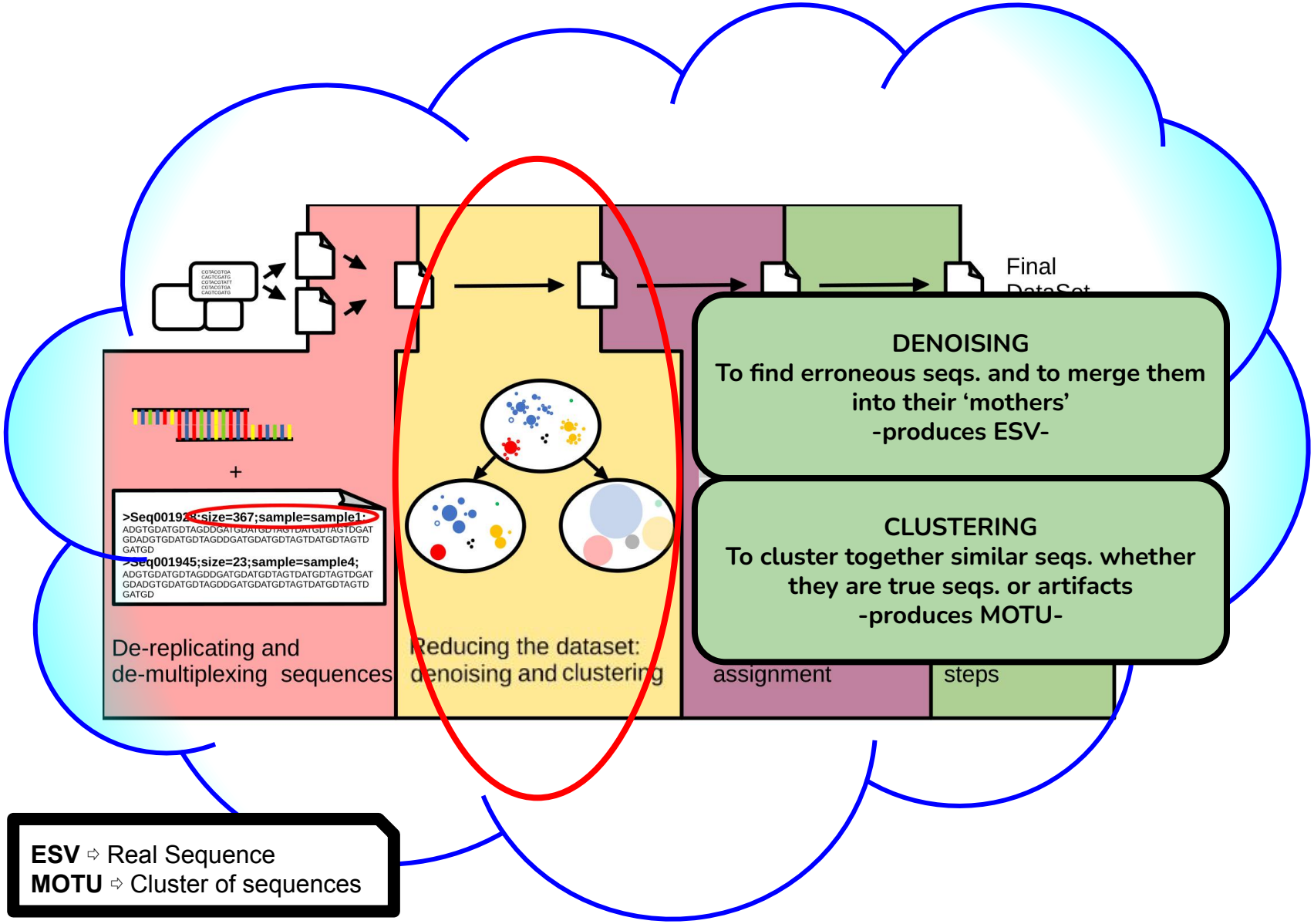


Remember metabarcoding pipeline...

5 Turon et al.2020

6 Antich et al.2021

7 Antich et al.2022



ESV ⇨ Real Sequence
MOTU ⇨ Cluster of sequences

5 Turon et al.2020

6 Antich et al.2021

7 Antich et al.2022

HOW DO WE TREAT DATA?

OPEN

The ISME Journal (2017) 11, 2639–2643
www.nature.com/ismej

PERSPECTIVE

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

Benjamin J Callahan[1], Paul J McMurdie[2] and Susan P Holmes[3]
[1]*Department of Population Health and Pathobiology, NC State University, Raleigh NC, USA;* [2]*Whole Biome Inc, San Francisco CA, USA and* [3]*Department of Statistics, Stanford University, Stanford CA, USA*

5 Turon et al.2020

6 Antich et al.2021

7 Antich et al.2022

HOW DO WE TREAT DATA?

OPEN

[2017] 11, 2639–2643

PERSPECTIVE

Exact sequence variants should [illegible]place operational taxonomic units in marker-gen[illegible]ta analysis

Benjamin J Callahan[1], Paul J McMurdie[2] and [illegible] Holmes[3]
[1]*Department of Population Health and Pathobiology, [illegible]ate University, Raleigh NC, USA;* [2]*Whole Biome Inc, San Francisco CA, USA and* [3]*Department of Sta[illegible]s, Stanford University, Stanford CA, USA*

HOW DO WE TREAT DATA?

OPEN

[2017] 11, 2639–2643

PERSPECTIVE

Exact sequence variants should [illegible]place operational taxonomic units in marker-gen[illegible]ta analysis

Benjamin J Callahan[1], Paul J McMurdie[2] and [illegible] Holmes[3]
[1]*Department of Population Health and Pathobiology, [illegible]ate University, Raleigh NC, USA;* [2]*Whole Biome Inc, San Francisco CA, USA and* [3]*Department of Sta[illegible]s, Stanford University, Stanford CA, USA*

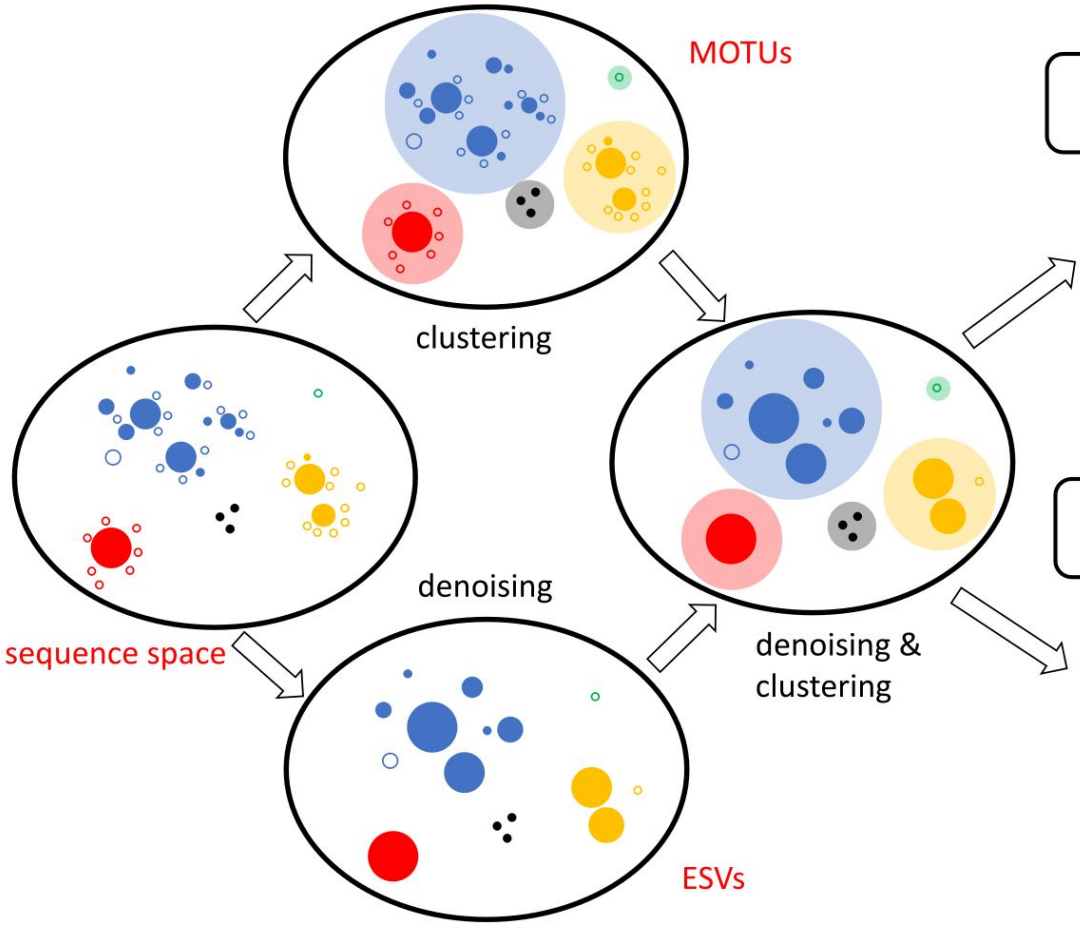
Maybe for <u>ribosomal</u> markers in prokaryotes, but not for <u>coding</u> genes with high variability in eukaryotes

Clustering + Denoising

5 Turon et al.2020

6 Antich et al.2021

7 Antich et al.2022



MOTU-level (~species) analyses

- α-diversity
- β-diversity
- Biogeography
- Functional analyses
- Ecological indices
- Indicator species

ESV-level (~haplotypes) analyses

- Metaphylogeography
- Haplotype networks
- Population genetics
- Connectivity
- Selection
- Indicator haplotypes

ESV ⇨ Real Sequence
MOTU ⇨ Cluster of sequences

HOW DO WE TREAT DATA?

Clustering

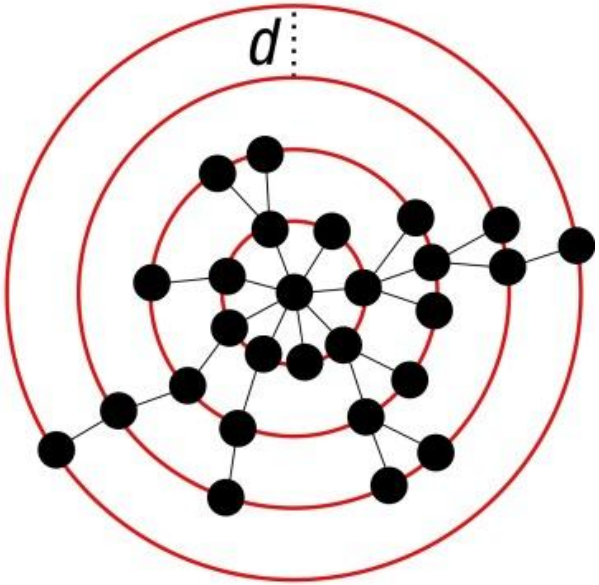
6

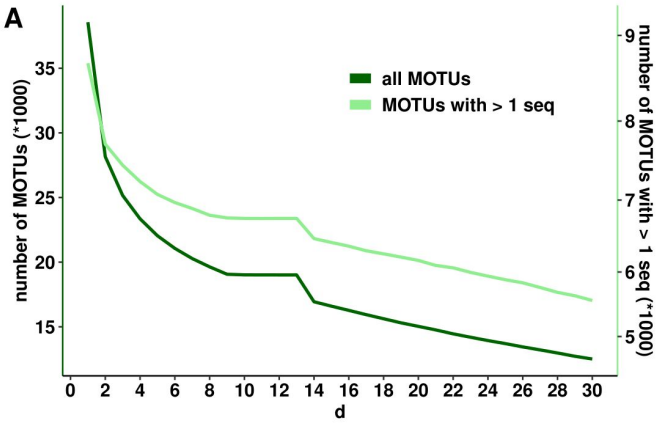
Antich et al.2021

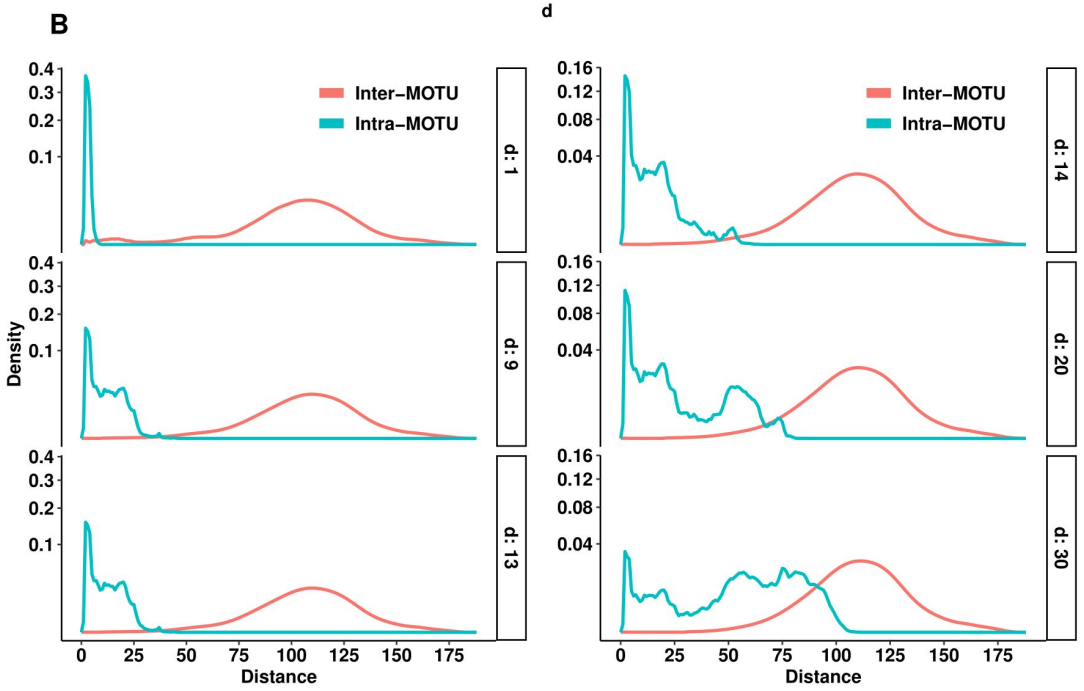
HOW DO WE TREAT DATA?

SWARM V3

(Mahé et al. 2021)







MOTU ⇨ Cluster of sequences

Clustering

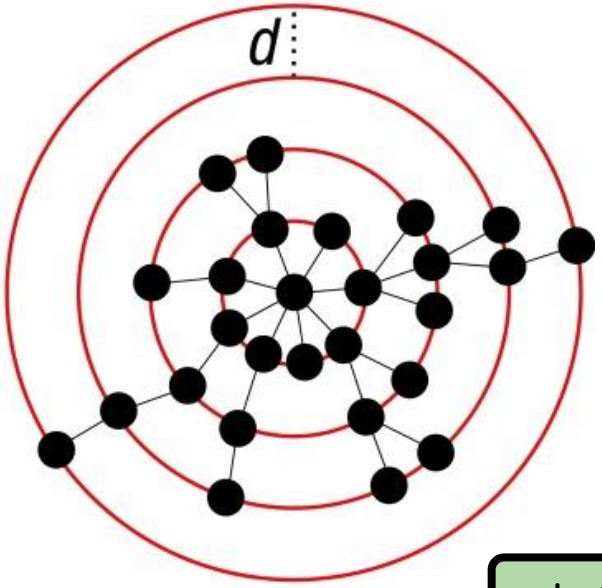
6

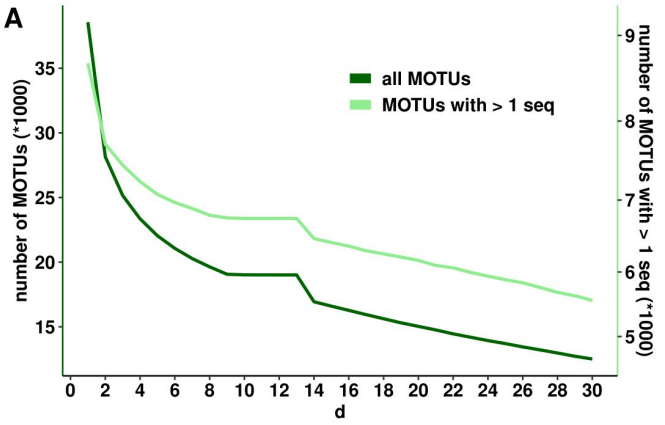
Antich et al.2021

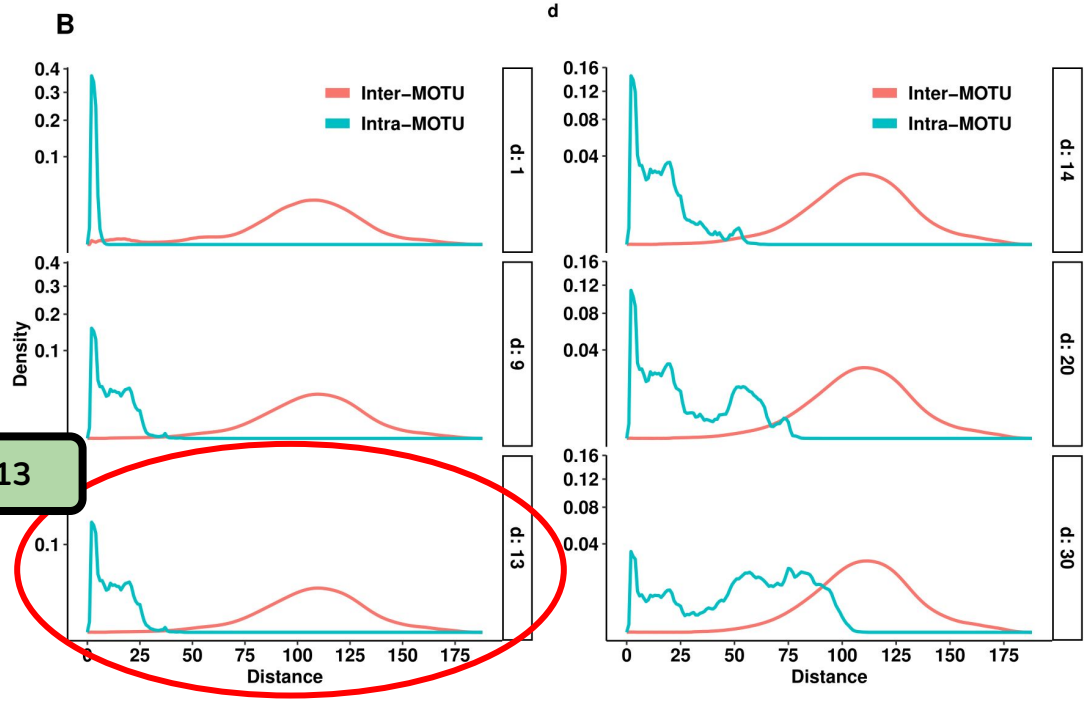
HOW DO WE TREAT DATA?

SWARM V3

(Mahé et al. 2021)







d = 13

MOTU ⇨ Cluster of sequences

Denoising

5 Turon et al.2020

6 Antich et al.2021

7 Antich et al.2022

HOW DO WE TREAT DATA?

UNOISE

(Edgar 2016)

$$\beta(d) = 1/2^{\alpha d+1}$$

Denoising

5 Turon et al.2020

6 Antich et al.2021

7 Antich et al.2022

HOW DO WE TREAT DATA?

UNOISE

(Edgar 2016)

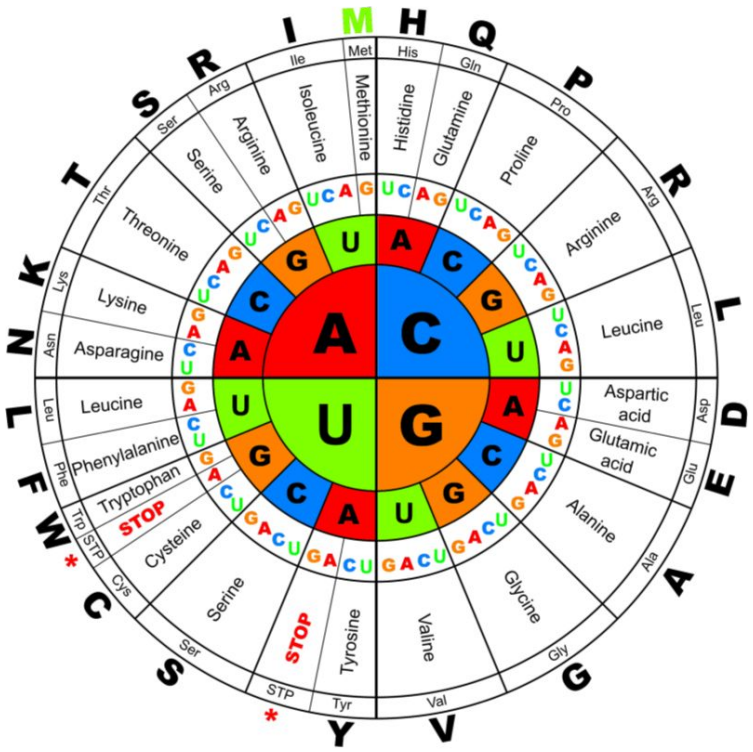
$$\beta(d) = 1/2^{\alpha d+1}$$

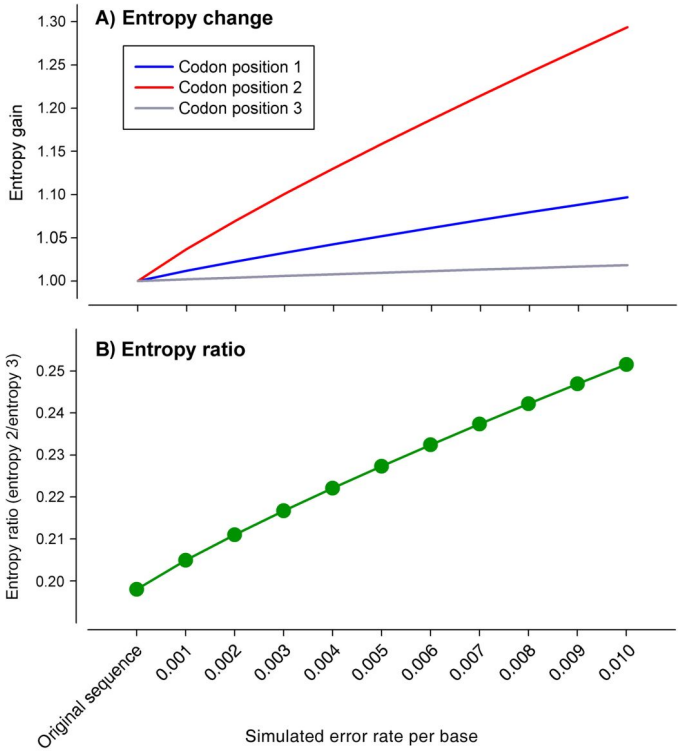
Denoising

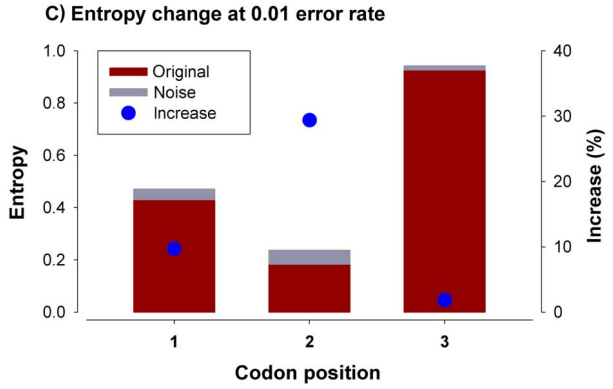
5 Turon et al.2020 | 6 Antich et al.2021 | 7 Antich et al.2022

UNOISE

(Edgar 2016)







Entropy → measure of variability

In Nature → ⇩ Entropy ratio (E2/E3)

Denoising

5 Turon et al.2020

6 Antich et al.2021

7 Antich et al.2022

UNOISE
(Edgar 2016)

DnoisE
(Antich et al. 2022)

OPEN SOURCE (PYTHON3)
PARALLELIZABLE
ENTROPY RATIO
MULTIPLE MERGING CRITERIA

$$\beta(d) = 1/2^{\alpha d+1}$$

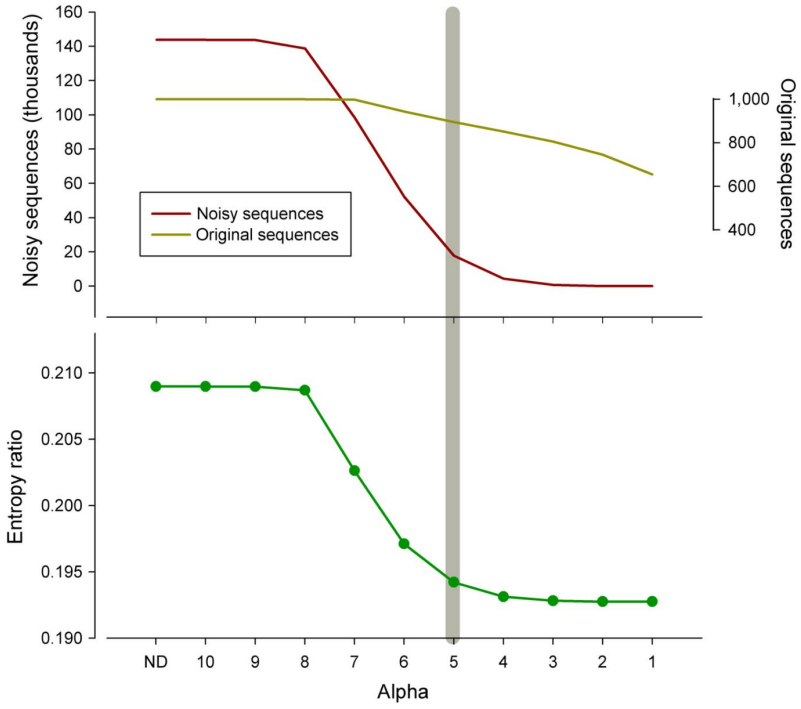
$$d_{corr} = \sum_{i=1}^{3} d(i) \times entropy(i) \times 3/(entropy(1) + entropy(2) + entropy(3))$$

Denoising

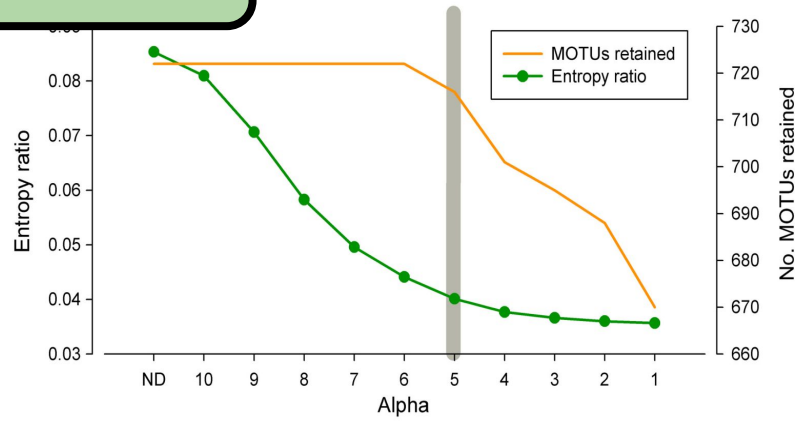
5 Turon et al.2020

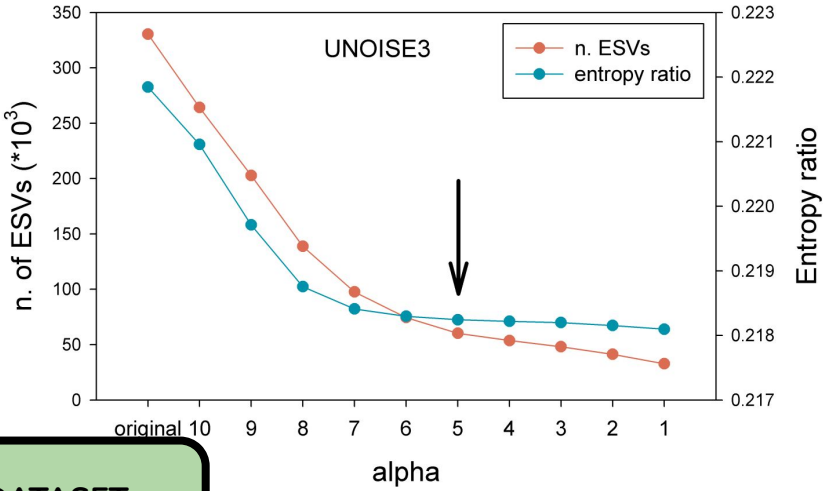
6 Antich et al.2021

SIMULATED DATASET
d NOT CORR



REAL DATASET
d NOT CORR





REAL DATASET
dcorr

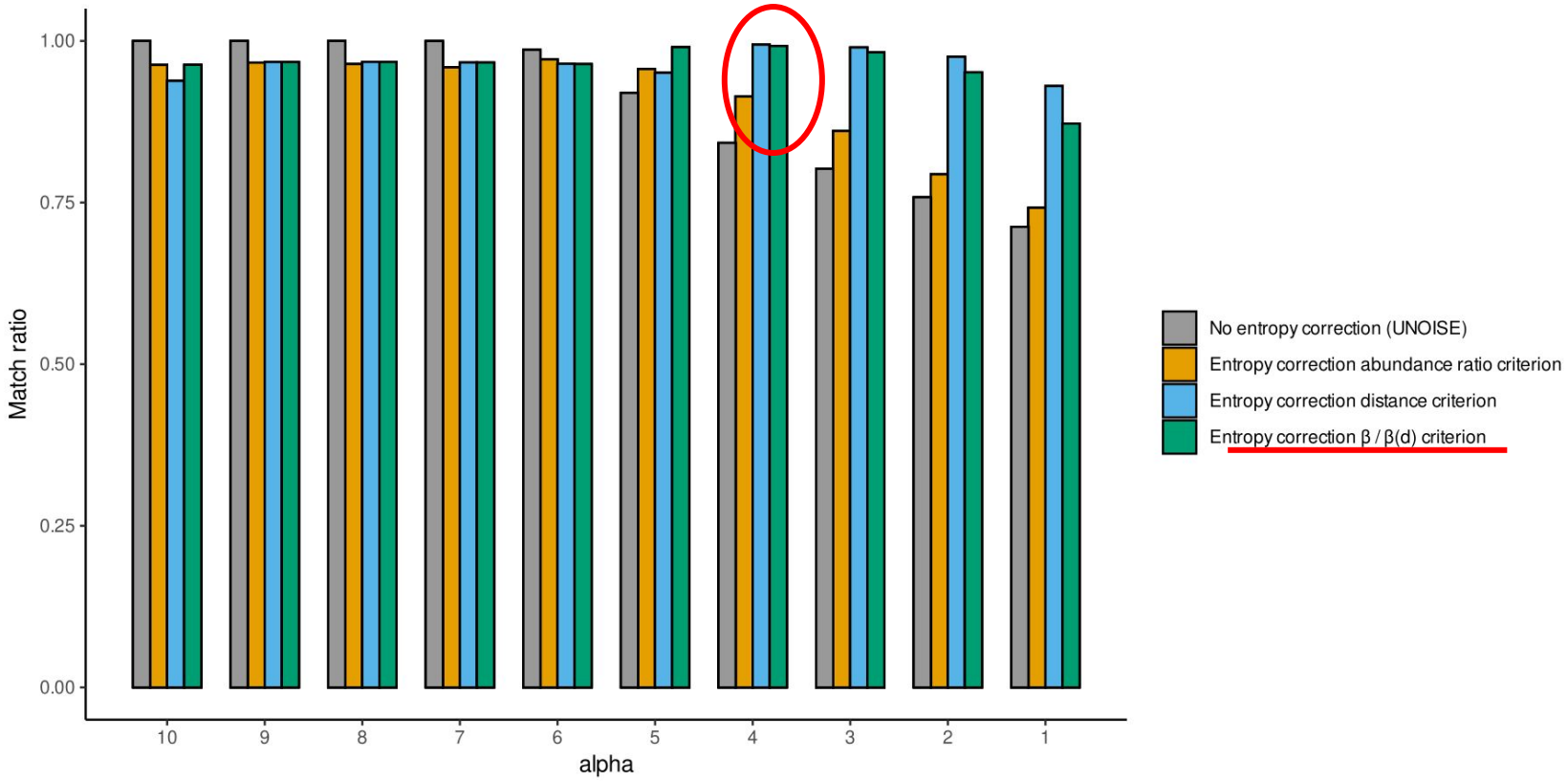
ESV ⇨ Real Sequence
MOTU ⇨ Cluster of sequences

Denoising and merging criteria

7

Antich et al.2022

HOW DO WE TREAT DATA?



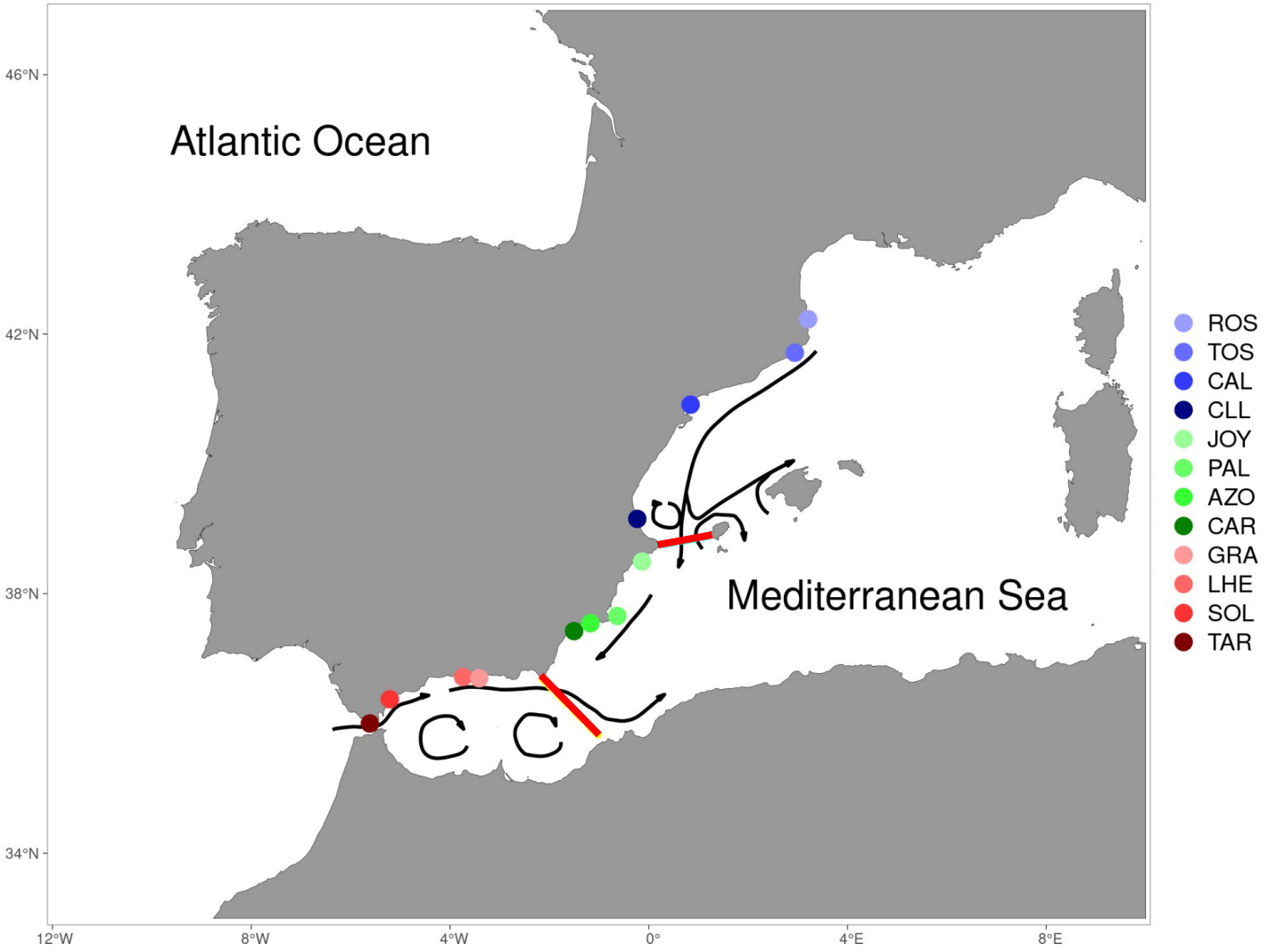
SIMULATED DATASET
dcorr

DESCRIBING OUR SHORES

Where?

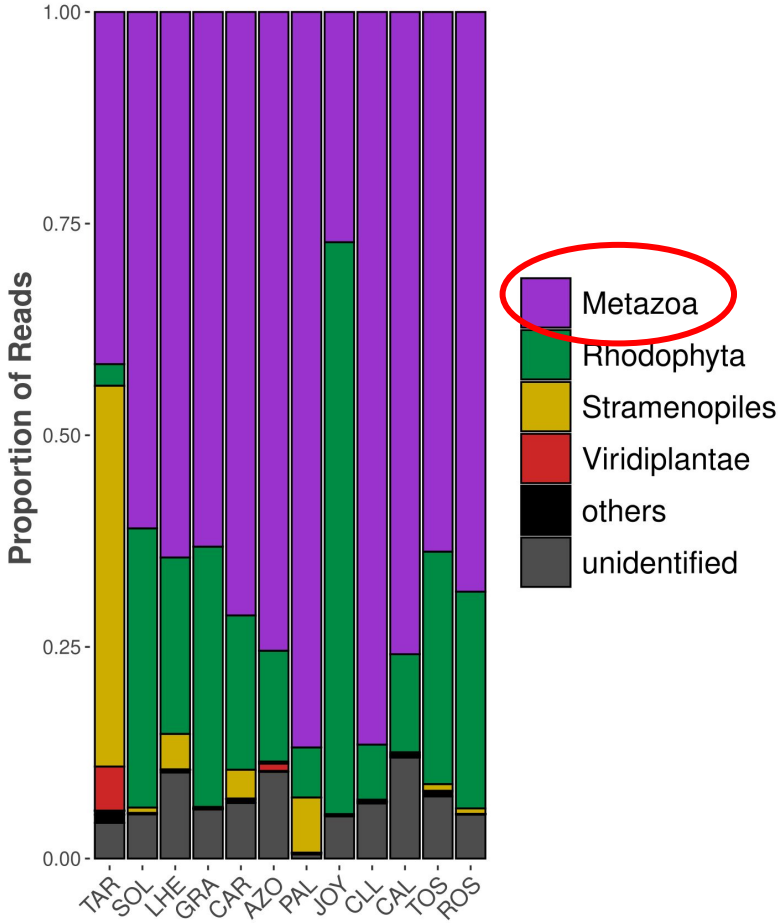
8

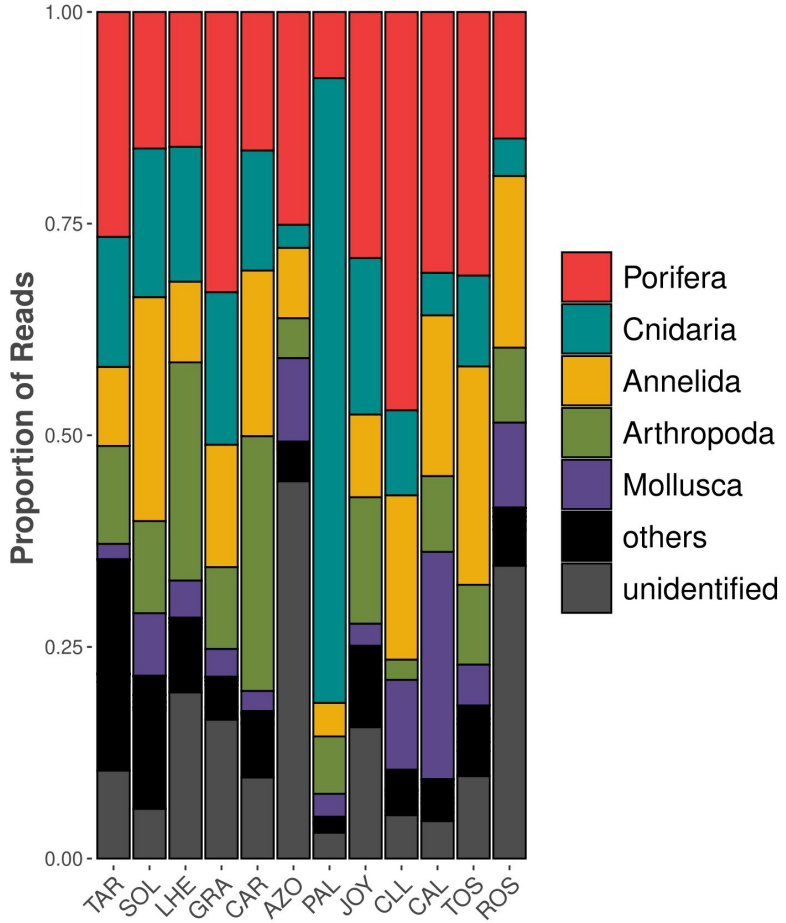
Antich et al.(press)



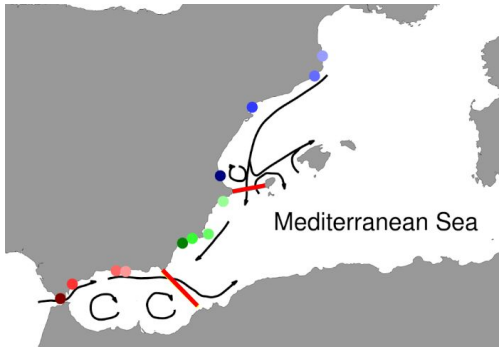
8

Antich et al.(press)

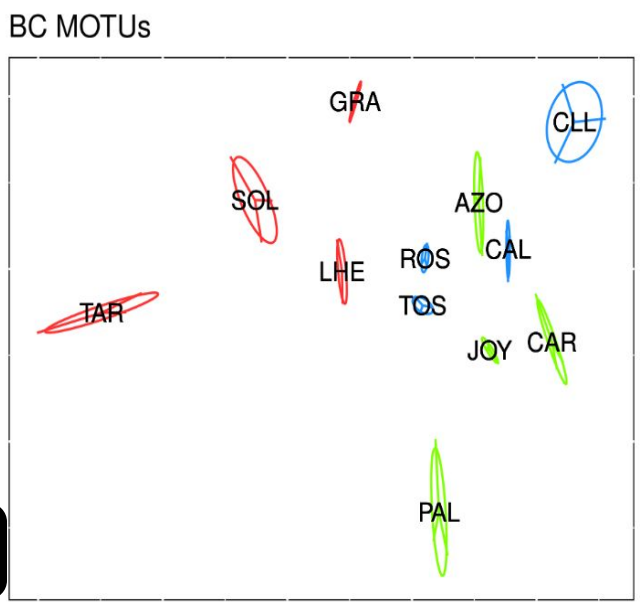


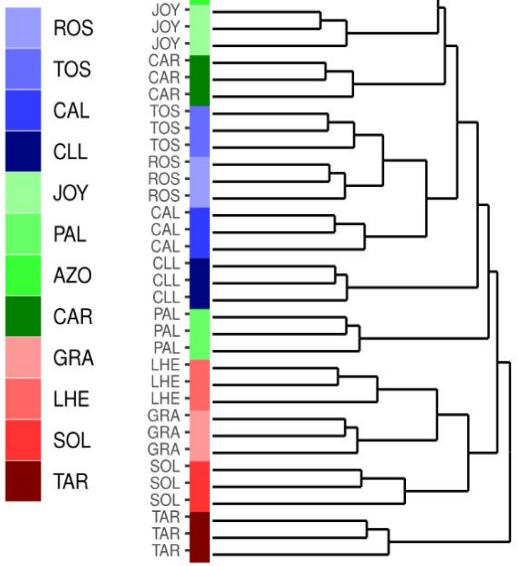


Read ⇨ each count of a seq.



8
Antich et al.(press)

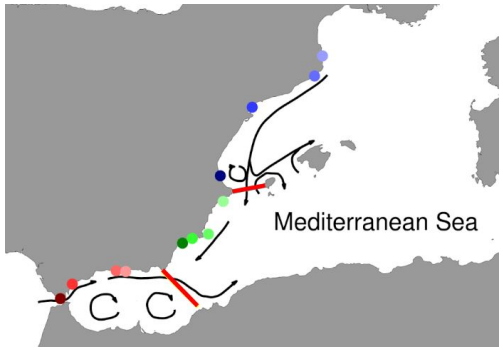




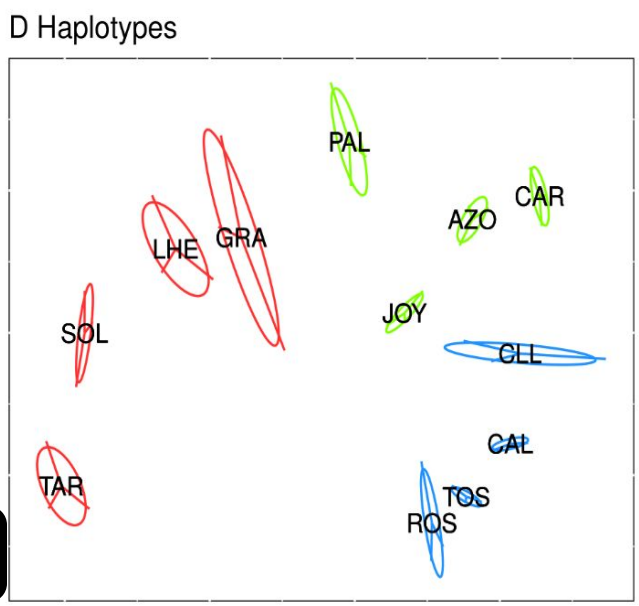
3,392 MOTUs

stress 0.143

MOTU ⇨ Cluster of sequences

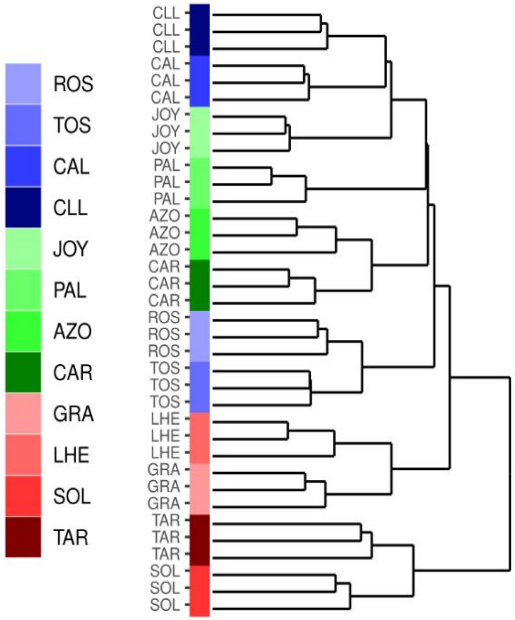


8
Antich et al.(press)



339 MOTUs

stress 0.202

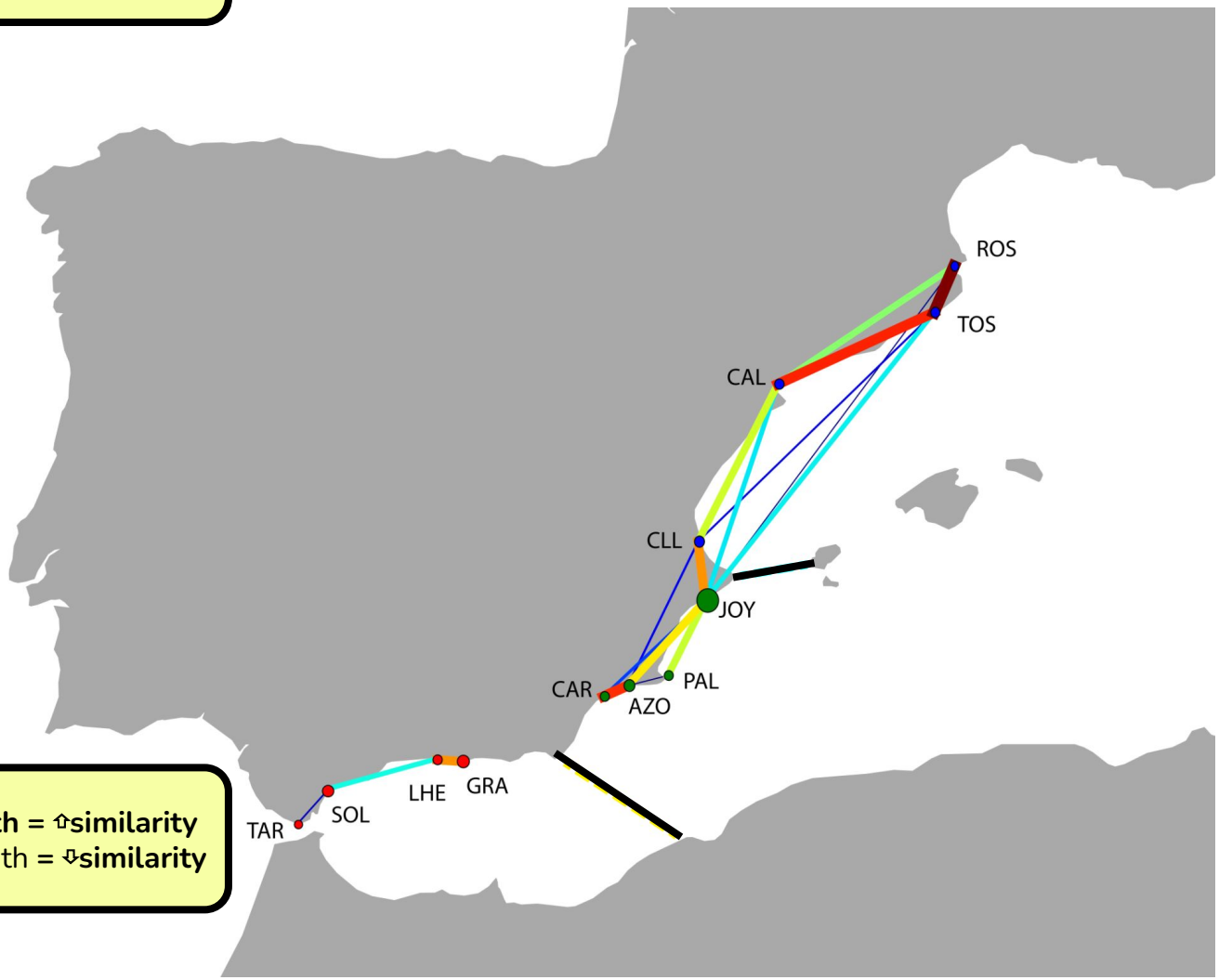


Haplotype ⇨ gene variant

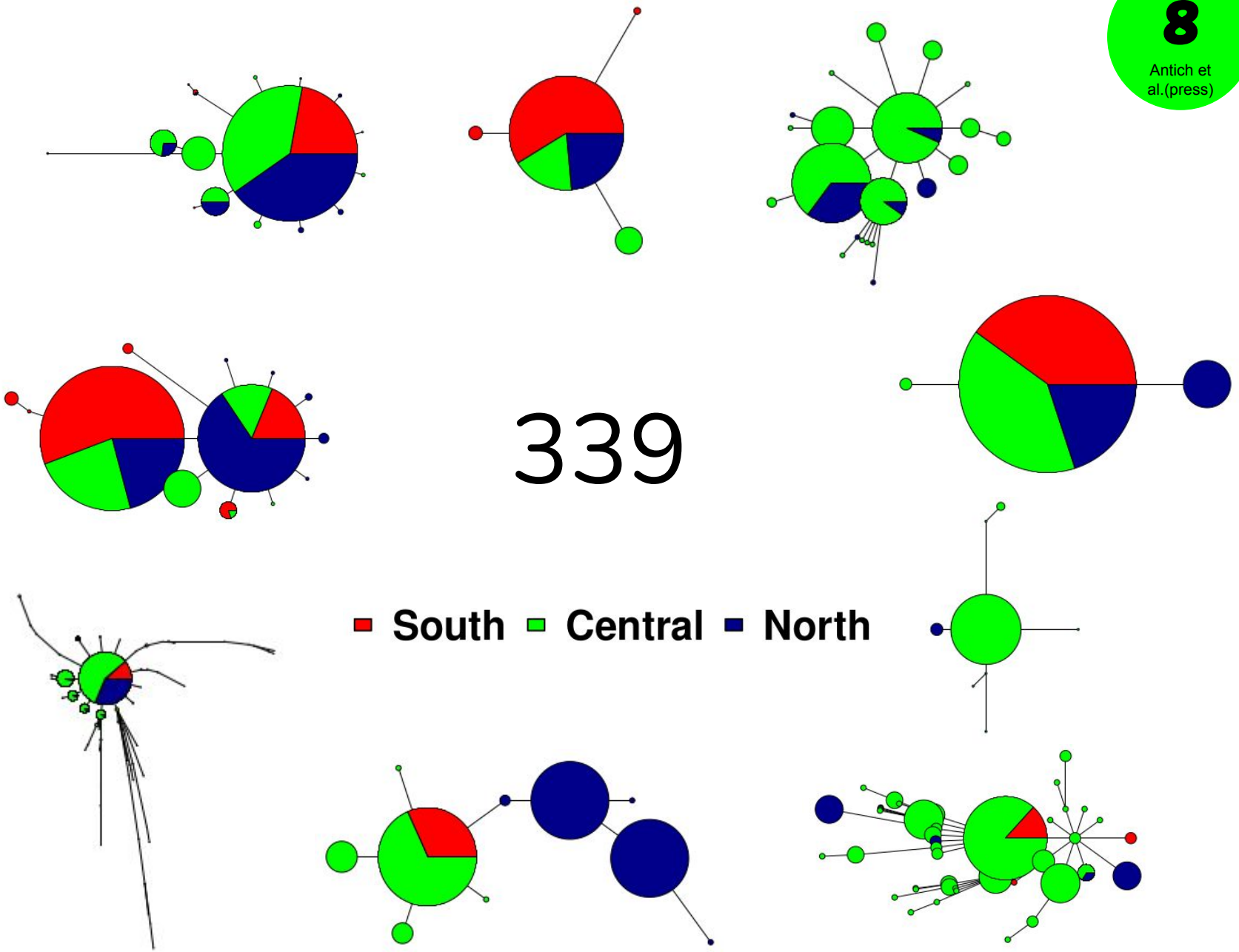
8

Antich et al.(press)

EDENetwork
D dissimilarities



red + ⇧**width** = ⇧**similarity**
blues + ⇩width = ⇩**similarity**


8
Antich et al.(press)
339
South
Central
North


CONCLUSIONS

CONCLUSIONS

Obj 1. To explore different sampling methods to assess the marine benthic diversity.

1.1 Direct sampling is necessary to properly assess the biodiversity of eukaryotic marine benthic communities using metabarcoding approaches. With universal primers targeting the highly variable COI marker, the environmental DNA captured from the water at the boundary layer failed to retrieve the benthic diversity and thus the traditional method of ’quadrat sampling’ is recommended.

Obj 2. Develop the pipelines and software required to obtain both inter and intraspecific variability from marine benthic communities of eukaryotes using COI.

2.1 We have developed a pipeline designed and tested for a highly variable metabarcode, the Leray fragment of COI, to obtain both the inter- and intra-MOTU variability. This pipeline combines the two common strategies to process sequences in metabarcoding studies, clustering and denoising.

2.2 We introduced the field of Metaphylogeography as the study of phylogeographic patterns of hundreds of species simultaneously using metabarcoding data.

Obj 2. Develop the pipelines and software required to obtain both inter and intraspecific variability from marine benthic communities of eukaryotes using COI.

2.3 We have calibrated the parameters of our preferred clustering program, SWARM, for COI metabarcoding of marine benthic communities.

2.4 We have designed, programmed, calibrated and tested the DnoisE program. This parallelizable open-source software is a new formulation of the UNOISE algorithm and includes a correction factor based on the different entropy values of each position in the codon to better assess the probability that a change in a given codon position is a natural variation or a sequencing artifact. DnoisE has been designed and tested for metabarcoding data of coding genes but can also be used (disabling entropy correction) with noncoding markers.

CONCLUSIONS

Obj 3. To apply these novel methodologies to analyse the biogeographic and metaphylogeographic patterns from benthic communities across two well described fronts in the eastern Iberian littoral.

3.1 We have successfully applied the metabarcoding methods developed to assess the biogeographic and metaphylogeographic patterns of the Eastern Iberian Coast to analyse the Atlanto-Mediterranean transition along two well studied barriers, the Almeria-Oran Front (AOF) and the Ibiza Channel (IC).

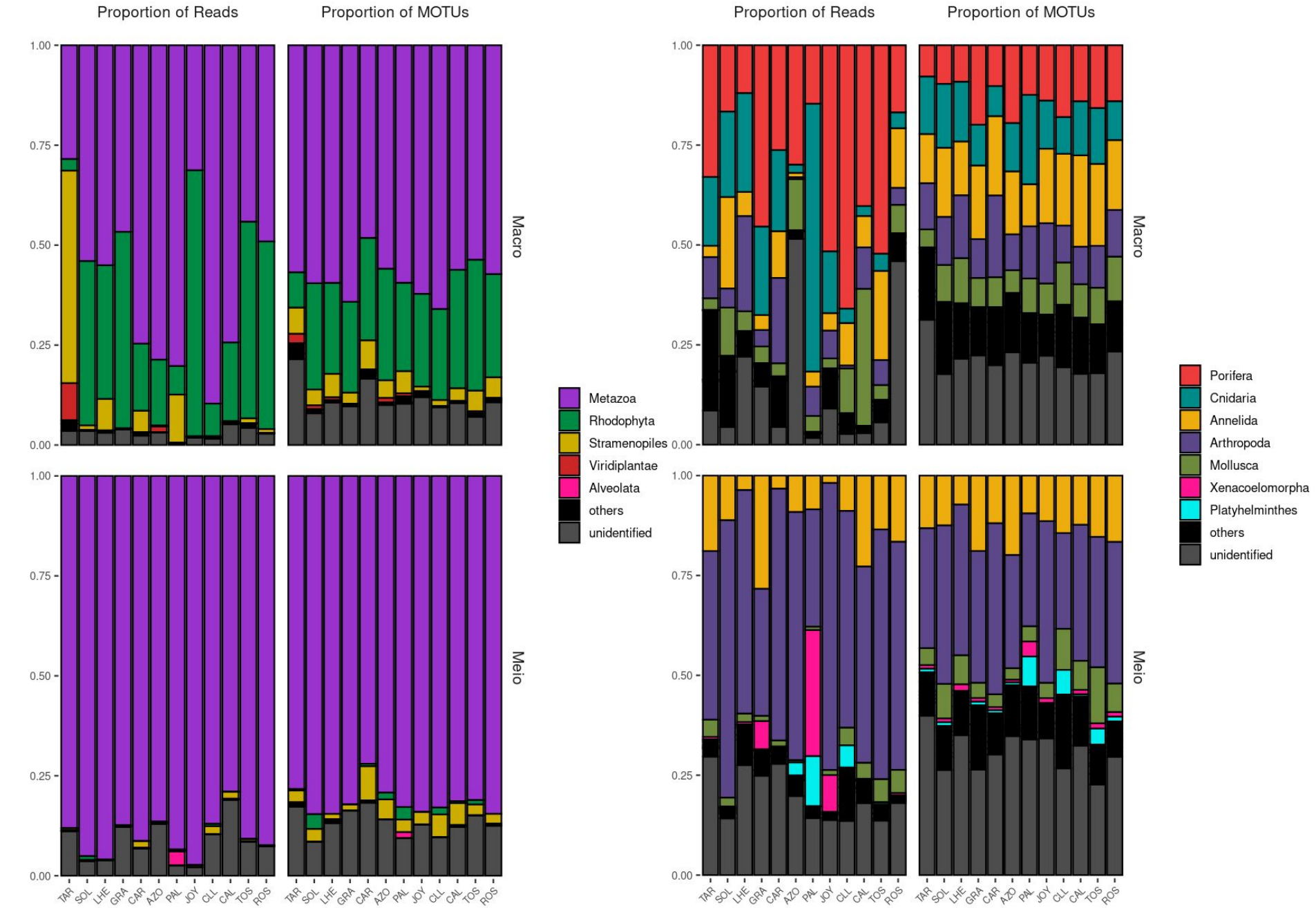
3.2 In all the analyses, the AOF had a strong effect separating regions, confirming previous reports. For the IC, a clear effect was detected only with the metaphylogeographic approach. Network analysis confirmed the important role of the AOF and showed that the localities close to the IC are highly connected with central and northern regions, indicating a potential transition area near the IC, rather than a well-delimited break.

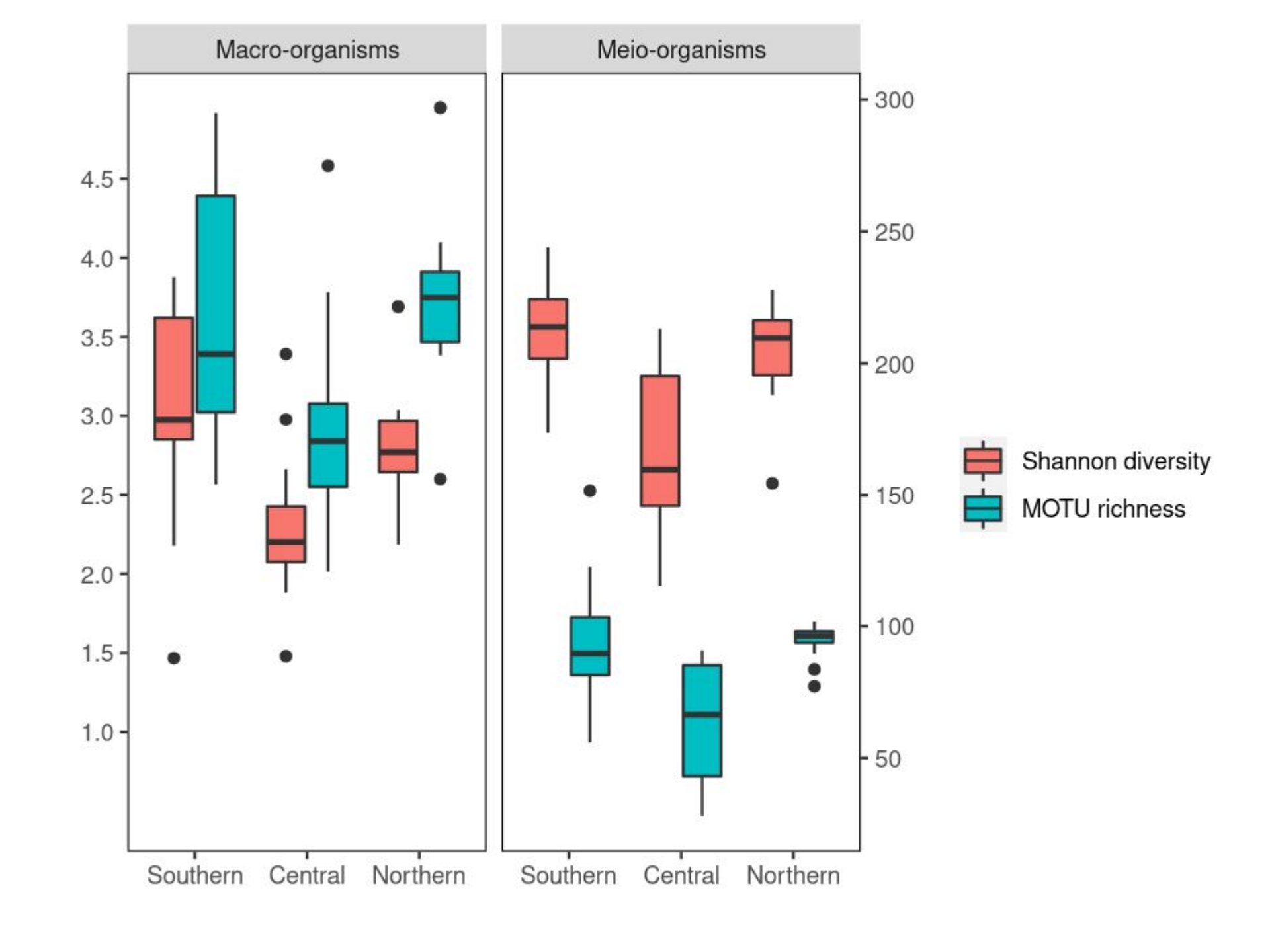
Obj 3. To apply these novel methodologies to analyse the biogeographic and metaphylogeographic patterns from benthic communities across two well described fronts in the eastern Iberian littoral.

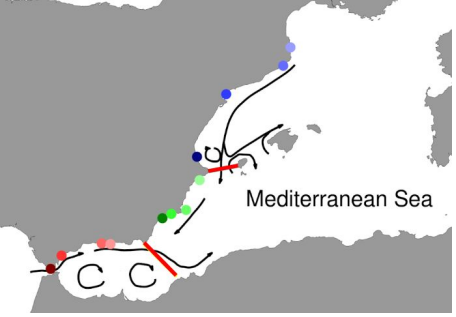
3.3 We favour the use of MOTUs as a proxy of species and ESVs within MOTUs as a proxy of haplotypes for metabarcoding with highly variable markers.

3.4 Metabarcoding offers the opportunity to fast and efficiently assess biogeographic and metaphylogeographic patterns and has the potential to become a cornerstone in biodiversity assessment of complex littoral benthic communities.

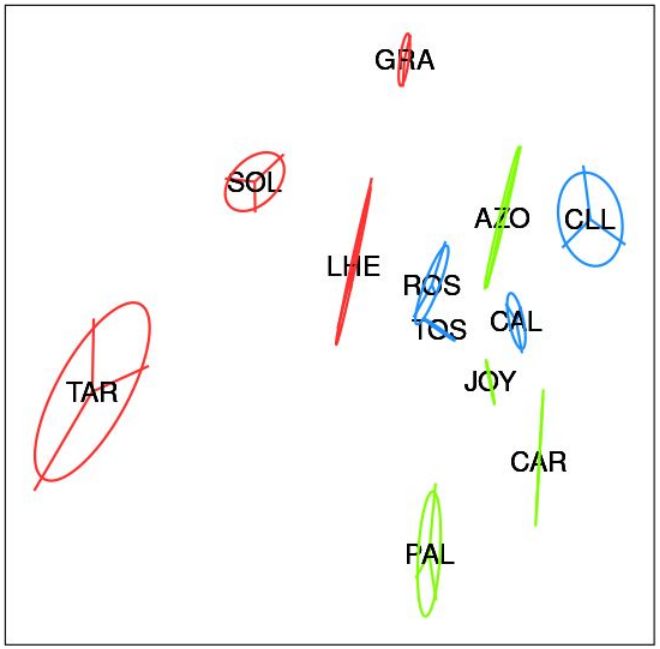
Moltes Gràcies!!!


Proportion of Reads
Proportion of MOTUs
Macro
Meio
1.00
0.75
0.50
0.25
0.00
TAR
SOL
LHE
GRA
CAR
AZO
PAL
JOY
CLL
CAL
TOS
ROS
Metazoa
Rhodophyta
Stramenopiles
Viridiplantae
Alveolata
others
unidentified
Porifera
Cnidaria
Annelida
Arthropoda
Mollusca
Xenacoelomorpha
Platyhelminthes

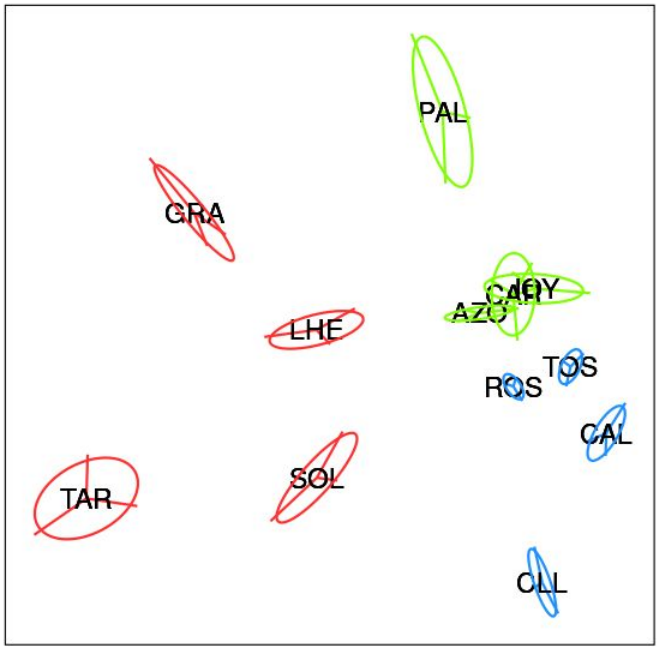

Macro-organisms
Meio-organisms
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
300
250
200
150
100
50
Southern
Central
Northern
Southern
Central
Northern
Shannon diversity
MOTU richness



BC Macro-organisms



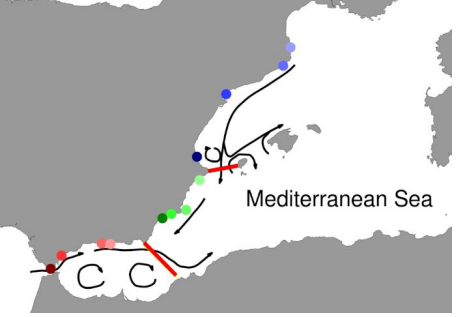
BC Meio-organisms



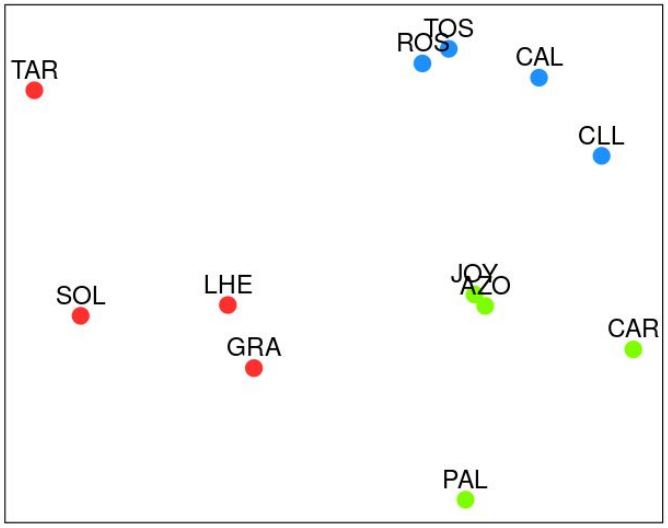
Northern

Central

Southern



D Macro-organisms



D Meio-organisms