azti tecnalia
Transforming
Science into
Business

Benchmarking DNA metabarcoding for measuring biodiversity and assessing environmental status

ICES/PICES
6TH ZOOPLANKTON PRODUCTION SYMPOSIUM
"NEW CHALLENGES IN A CHANGING OCEAN"
SCANDIC BERGEN CITY, BERGEN, NORWAY
9–13 MAY 2016

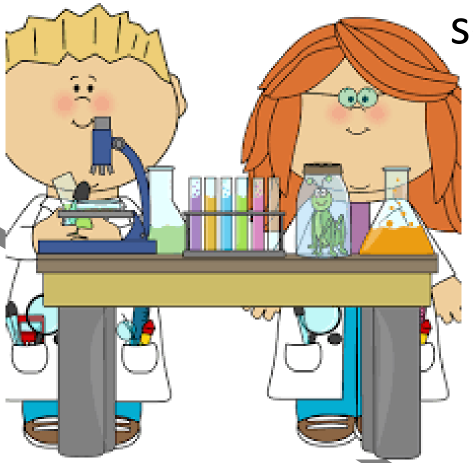
NAIARA RODRIGUEZ-EZPELETA

EVA AYLAGAS **JON CORELL**

Types of samples
- Sediment vs water, filtering 2L vs filtering 20 L, extracting extracellular or intracellular DNA

Transforming Science into Business

DNA extraction method
- Different protocols/kits are more efficient for some taxonomic groups than for others



Barcode choice
- COI, 18S

Primer choice
- Leray, Folmer

Barcode amplification conditions
- PCR Annealing temperature

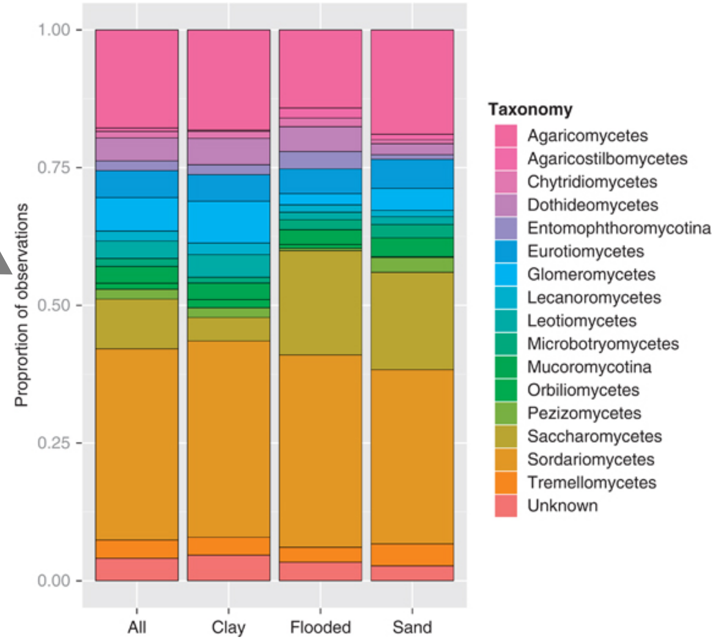
Sequencing platform
- Illumina, Ion Torrent

Analysis pipeline
- Qiime, mothur

Quality filtering
- Strict, loose
- Chimera removal
- Singleton removal

Classification strategies
- OTU based
- Taxonomy based

Database
- More or less complete



SAMPLE TYPE

Same barcode, primer, amplification conditions, analysis pipeline	Intracellular	Extracellular
Annelida;Clitellata;Haplotaxida;Tubificidae;Tubificoides;Tubificoides_amplivasatus;	0	1
Annelida;Polychaeta;Phyllodocida;Glyceridae;Glycera;Glycera_unicornis;	1	0
Annelida;Polychaeta;Phyllodocida;Goniadidae;Glycinde;Glycinde_nordmanni;	1	0
Annelida;Polychaeta;Phyllodocida;Nephtyidae;Nephtys;Nephtys_hombergii;	1	0
Annelida;Polychaeta;Phyllodocida;Nephtyidae;Nephtys;Nephtys_kersivalensis;	1	0
Annelida;Polychaeta;Phyllodocida;Nephtyidae;Nephtys;_;	1	0
Annelida;Polychaeta;Phyllodocida;Phyllodocidae;Phyllodoce;Phyllodoce_rosea;	1	0
Annelida;Polychaeta;Phyllodocida;Pilargidae;Pilargis;Pilargis_verrucosa;	1	0
Annelida;Polychaeta;Terebellida;Cirratulidae;Chaetozone;Chaetozone_gibber;	1	0
Annelida;Polychaeta;_;Capitellidae;Notomastus;_;	1	0
Annelida;Polychaeta;_;Paraonidae;Paradoneis;Paradoneis_ilvana;	1	1
Arthropoda;Malacostraca;Amphipoda;Ampeliscidae;Ampelisca;Ampelisca_pectenata;	1	0
Mollusca;Bivalvia;Lucinoida;Thyasiridae;Axinulus;Axinulus_croulinensis;	1	0
Echinodermata;Ophiuroidea;Ophiurida;Ophiothrichidae;Macrophiothrix;_;	0	1
Nemertea;Anopla;Heteronemertea;_;Cerebratulus;_;	1	0
Porifera;Demospongiae;Hadromerida;Hemiasterellidae;_;_;	0	1

PRIMER

Same sample type, barcode, amplification conditions, analysis pipeline	Leray	Folmer
Annelida;Polychaeta;Phyllodocida;Pilargidae;Litocorsa;Litocorsa_stremma;	1	1
Annelida;Polychaeta;Eunicida;Lumbrineridae;Gallardoneris;Gallardoneris_iberica;	1	0
Annelida;Polychaeta;Sabellida;Oweniidae;Galathowenia;Galathowenia_oculata;	1	1
Annelida;Polychaeta;Terebellida;Ampharetidae;Auchenoplax;Auchenoplax_crinita;	1	1
Annelida;Polychaeta;Terebellida;Ampharetidae;Eclysippe;Eclysippe_vanelli;	1	1
Annelida;Polychaeta;Terebellida;Terebellidae;Polycirrus;_;	1	0
Annelida;Polychaeta;_;Maldanidae;Praxillella;_;	1	0
Annelida;Polychaeta;Spionida;Magelonidae;Magelona;Magelona_lusitanica;	1	1
Annelida;Polychaeta;_;Paraonidae;Levinsenia;Levinsenia_gracilis;	1	1
Mollusca;Bivalvia;Veneroida;Semelidae;Abra;Abra_alba;	1	0
Mollusca;Bivalvia;Lucinoida;Thyasiridae;Thyasira;Thyasira_flexuosa;	1	0
Mollusca;Bivalvia;Lucinoida;Thyasiridae;Thyasira;Thyasira_ferruginea;	1	1

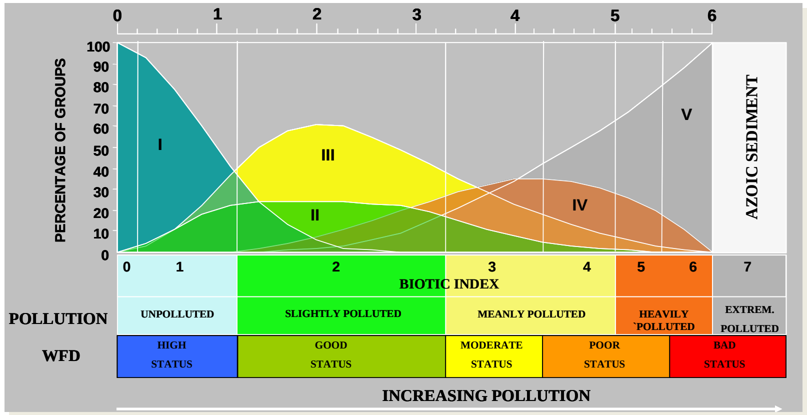
ANNEALING TEMPERATURE

Same sample type, barcode, primer, analysis pipeline	46	TD
Annelida;Polychaeta;Phyllodocida;Nephtyidae;Nephtys;Nephtys_kersivalensis;	1	0
Annelida;Polychaeta;Phyllodocida;Nephtyidae;Nephtys;_;	1	1
Annelida;Polychaeta;Phyllodocida;Pilargidae;Ancistrosyllis;Ancistrosyllis_groenlandica;	1	1
Annelida;Polychaeta;Phyllodocida;Pilargidae;Pilargis;Pilargis_verrucosa;	0	1
Annelida;Polychaeta;Eunicida;Lumbrineridae;Gallardoneris;Gallardoneris_iberica;	1	0
Annelida;Polychaeta;Terebellida;Cirratulidae;Chaetozone;Chaetozone_gibber;	1	1
Annelida;Polychaeta;Terebellida;Terebellidae;Pista;Pista_cristata;	1	0
Annelida;Polychaeta;Spionida;Magelonidae;Magelona;Magelona_minuta;	1	0
Annelida;Polychaeta;Spionida;Spionidae;Aonides;_;	1	1
Arthropoda;Malacostraca;Amphipoda;Ampeliscidae;Ampelisca;Ampelisca_pectenata;	1	1
Arthropoda;Malacostraca;Amphipoda;Ampeliscidae;Ampelisca;Ampelisca_typica;	0	1
Arthropoda;Malacostraca;Isopoda;Cirolanidae;Natatolana;Natatolana_borealis;	1	1
Echinodermata;Echinoidea;Echinoida;Echinidae;Paracentrotus;Paracentrotus_lividus;	1	0
Echinodermata;Ophiuroidea;Ophiurida;Ophiuridae;Ophiura;_;	1	1
Echinodermata;Ophiuroidea;Ophiurida;Ophiuridae;Ophiura;Ophiura_texturata;	1	1
Echinodermata;Ophiuroidea;Ophiurida;Ophiuridae;Ophiura;Ophiura_albida;	1	1
Sipuncula;Sipunculidea;Golfingiida;Phascolionidae;Onchnesoma;Onchnesoma_steenstrupii;	1	0

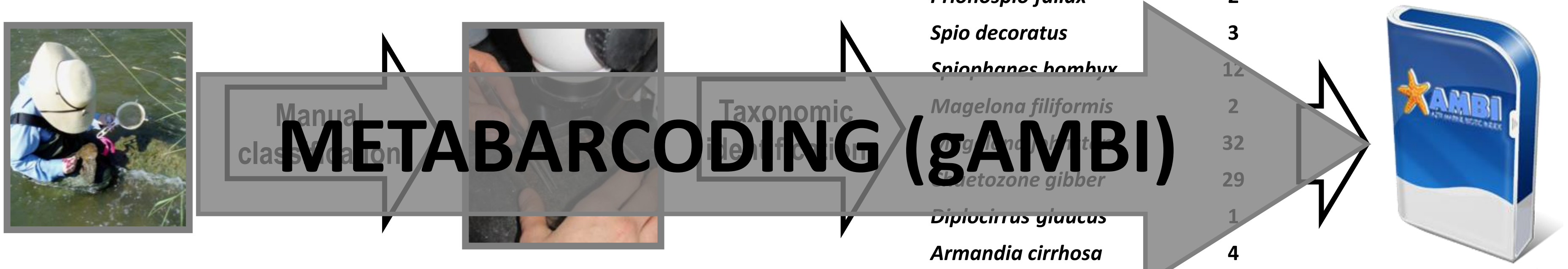
AZTI´s Marine Biotic Index - AMBI

Based on abundance-weighted pollution tolerances of the species present in a sample, where tolerance is expressed categorically as one of five ecological groups:

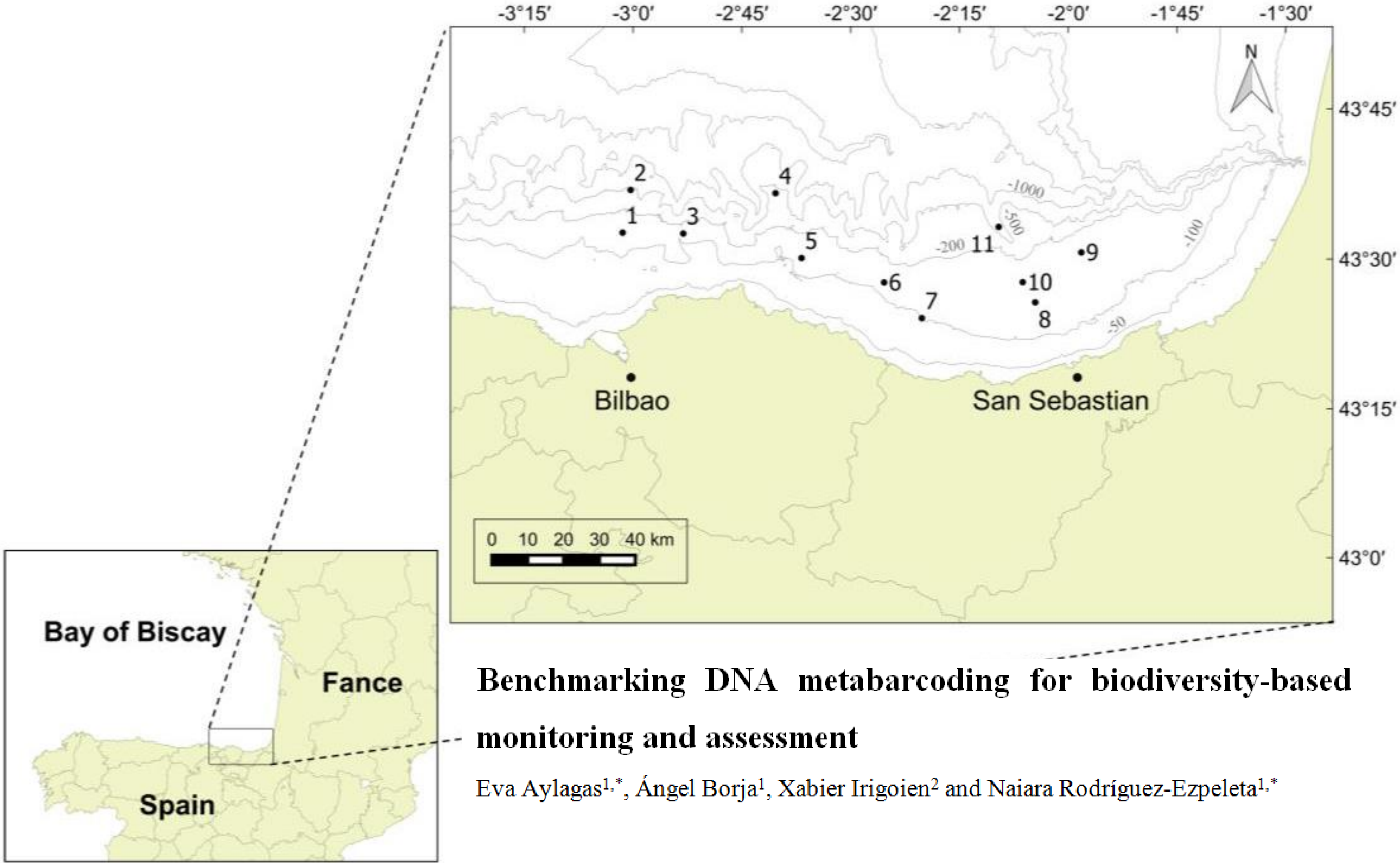
I	sensitive to pressure
II	indifferent
III	tolerant
IV	opportunist of second order
V	opportunist of first order



$$\boxed{AMBI = ((0 * \%GI) + (1.5 * \%GII) + (3 * \%GIII) + (4.5 * \%GIV) + (6 * \%GV))/100}$$



SPECIES	ABUNDANCE
Prionospio fallax	2
Spio decoratus	3
Spiophanes bombyx	12
Magelona filiformis	2
Magelona johnstoni	32
Chaetozone gibber	29
Diplocirrus glaucus	1
Armandia cirrhosa	4
Capitella capitata	4
Mediomastus fragilis	3



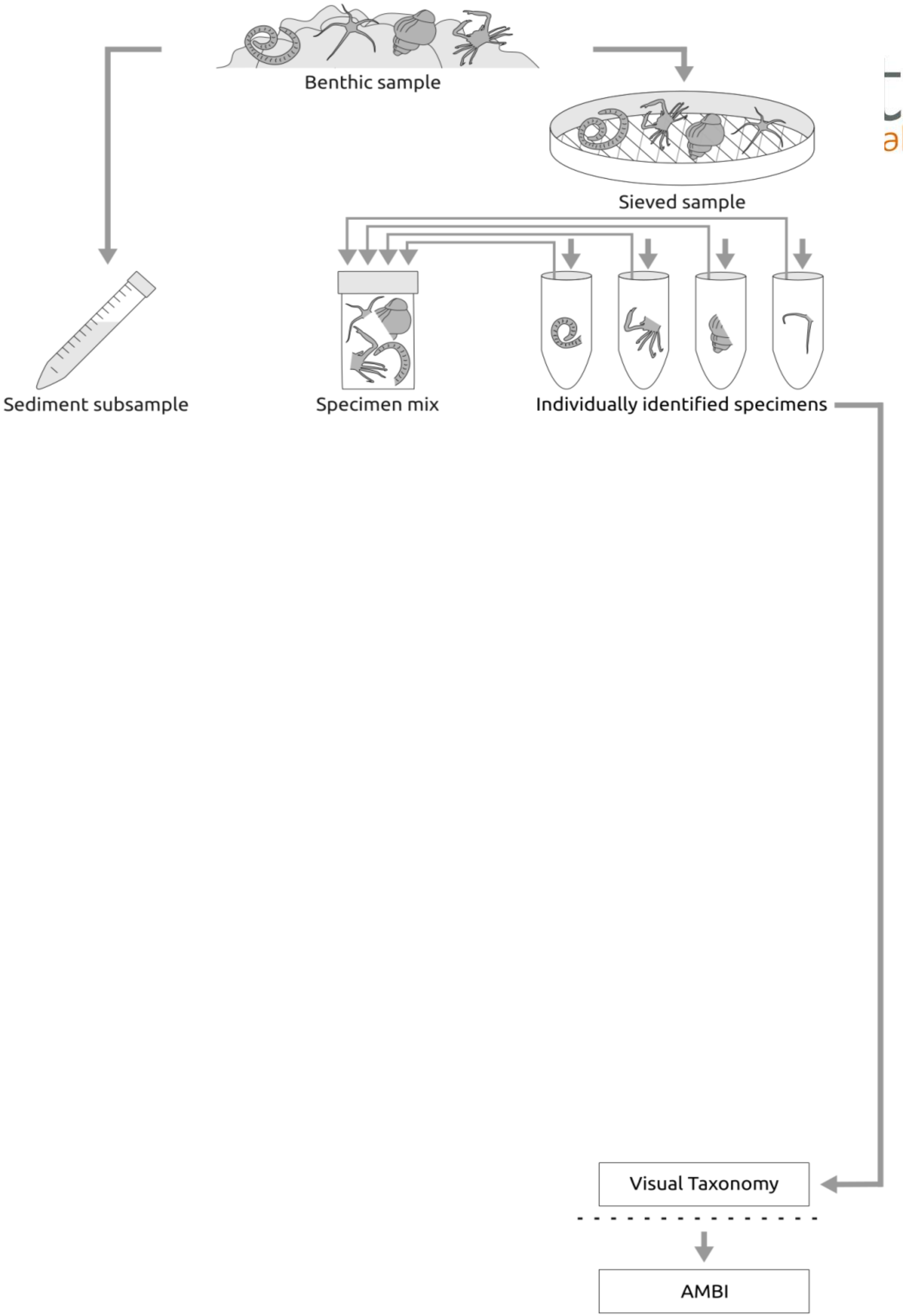
Benchmarking DNA metabarcoding for biodiversity-based monitoring and assessment

Eva Aylagas[1,*], Ángel Borja[1], Xabier Irigoien[2] and Naiara Rodríguez-Ezpeleta[1,*]

[1]AZTI, Marine Research Division, Sukarrieta, Bizkaia, Spain

[2]Red Sea Research Center, King Abdullah University of Science and Technology (KAUST), Saudi Arabia

Sample collection
& processing



Benthic sample

Sieved sample

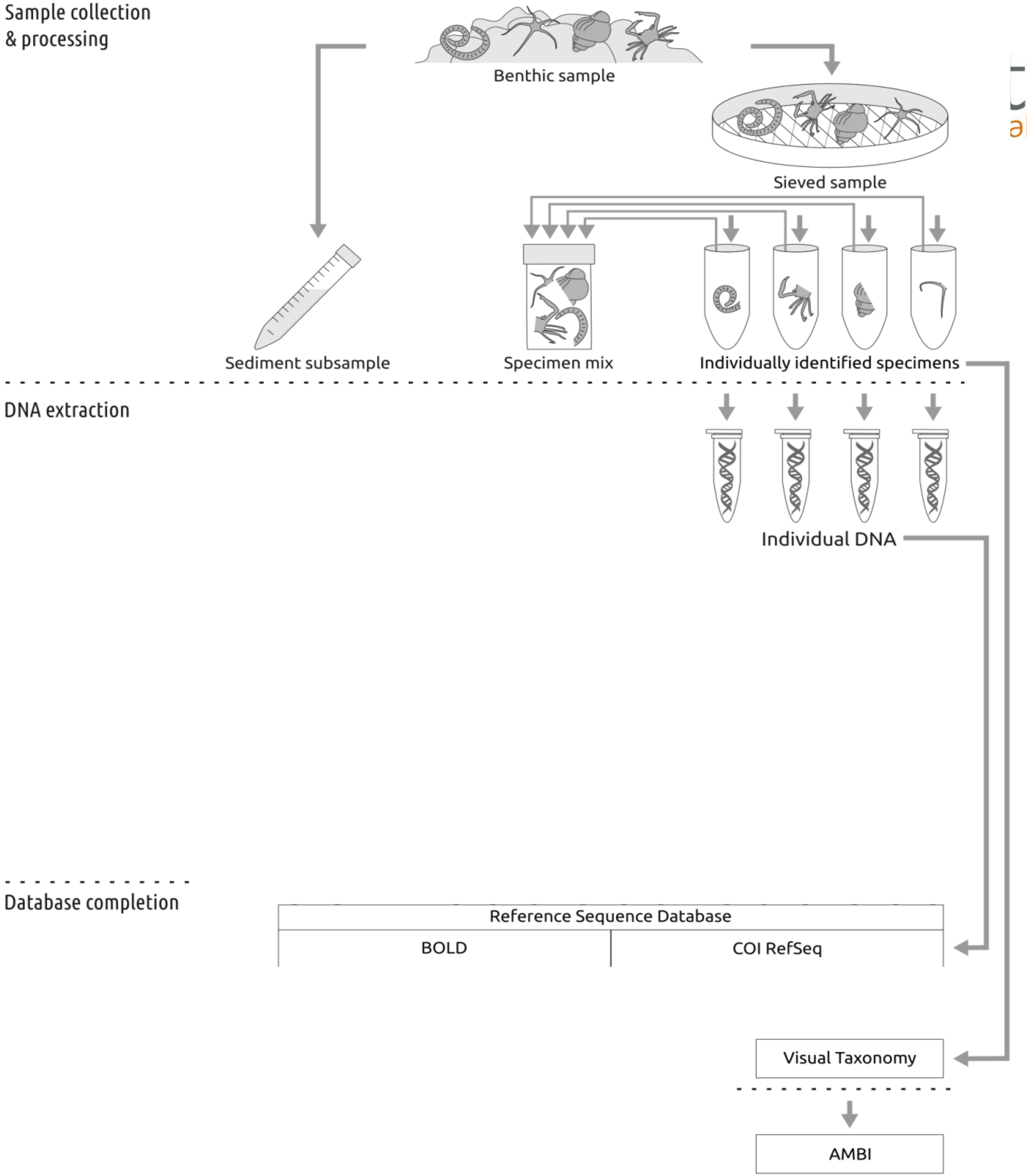
Sediment subsample

Specimen mix

Individually identified specimens

Visual Taxonomy

AMBI



Sample collection
& processing

Benthic sample

Sieved sample

Sediment subsample

Specimen mix

Individually identified specimens

DNA extraction

Individual DNA

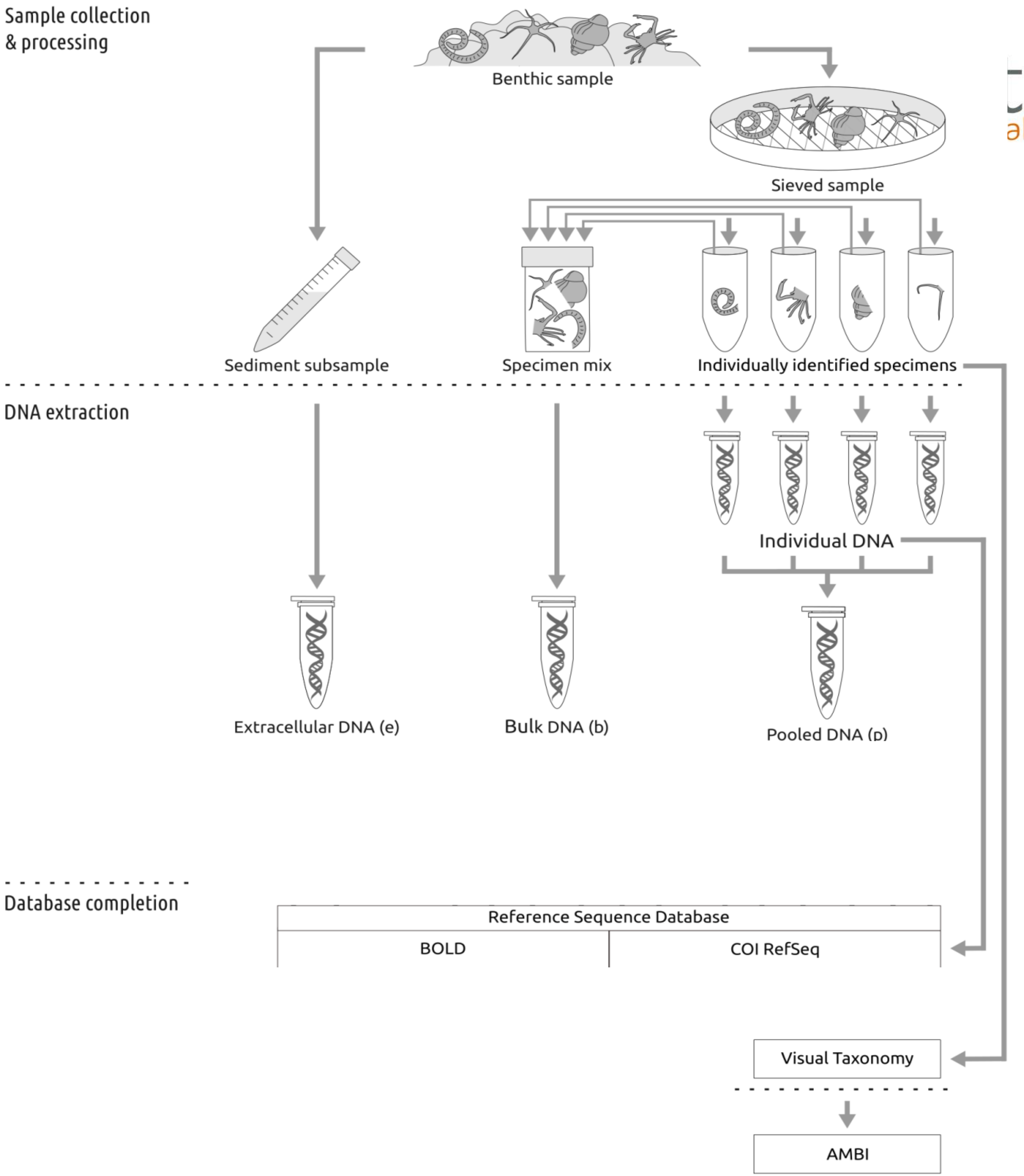
Database completion

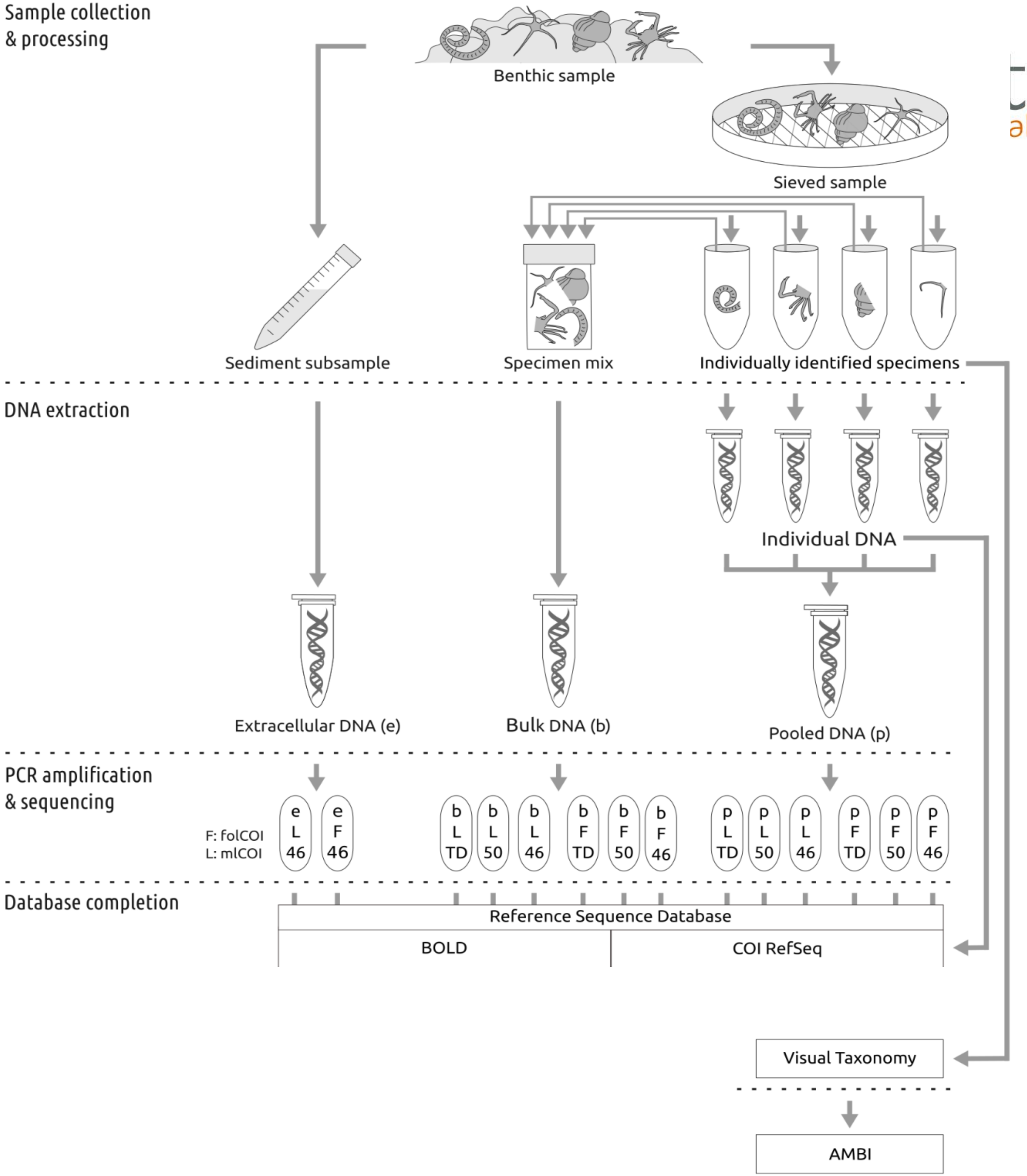
Reference Sequence Database

BOLD	COI RefSeq

Visual Taxonomy

AMBI


Sample collection
& processing
Benthic sample
Sieved sample
Sediment subsample
Specimen mix
Individually identified specimens
DNA extraction
Individual DNA
Extracellular DNA (e)
Bulk DNA (b)
Pooled DNA (p)
Database completion
Reference Sequence Database
BOLD
COI RefSeq
Visual Taxonomy
AMBI



Sample collection & processing

Benthic sample

Sieved sample

Sediment subsample

Specimen mix

Individually identified specimens

DNA extraction

Individual DNA

Extracellular DNA (e)

Bulk DNA (b)

Pooled DNA (p)

PCR amplification & sequencing

F: folCOI
L: mlCOI

Database completion

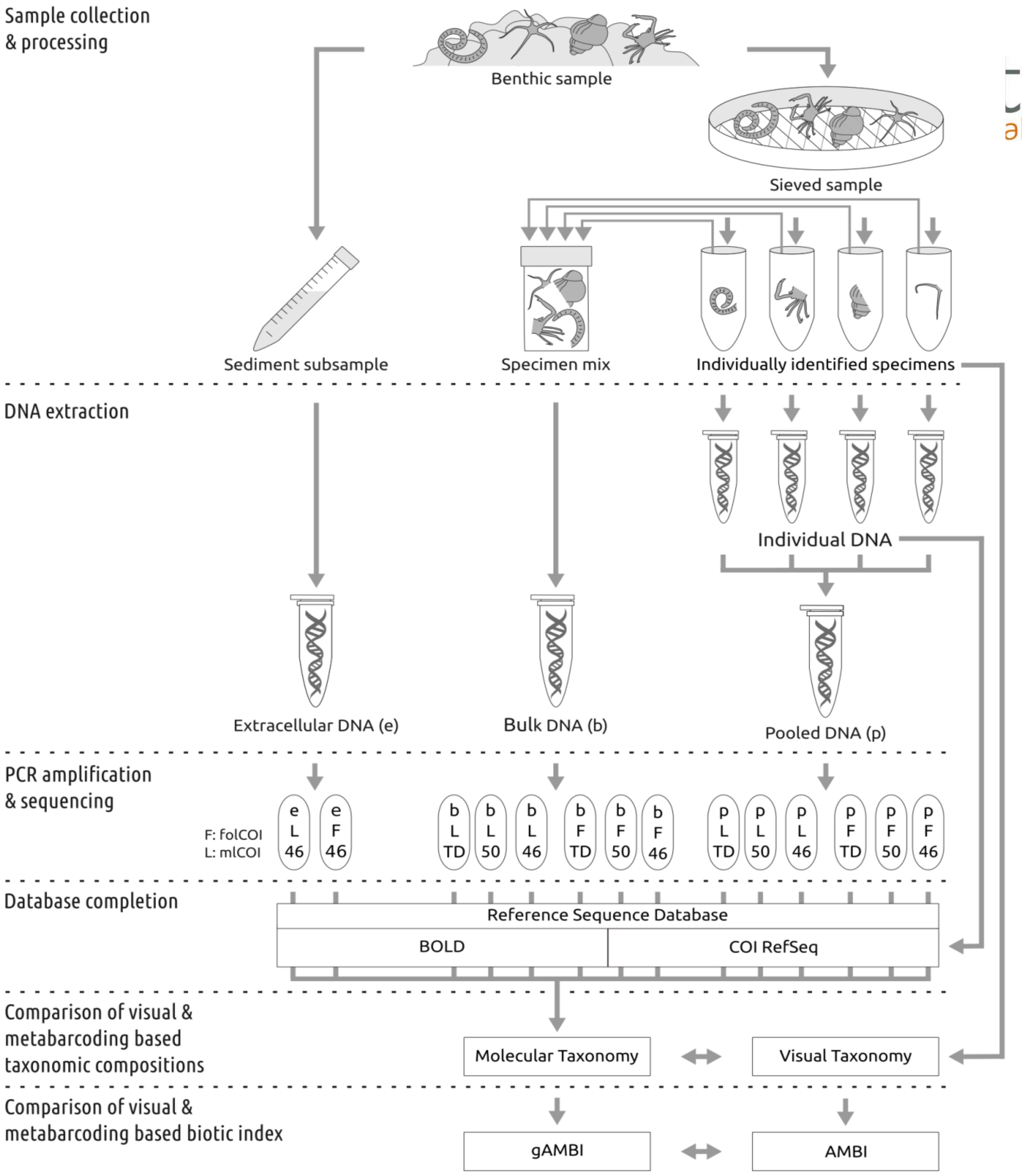
Reference Sequence Database

BOLD

COI RefSeq

Visual Taxonomy

AMBI



Sample collection & processing

Benthic sample

Sieved sample

Sediment subsample

Specimen mix

Individually identified specimens

DNA extraction

Individual DNA

Extracellular DNA (e)

Bulk DNA (b)

Pooled DNA (p)

PCR amplification & sequencing

F: folCOI
L: mlCOI

Database completion

Reference Sequence Database

BOLD

COI RefSeq

Comparison of visual & metabarcoding based taxonomic compositions

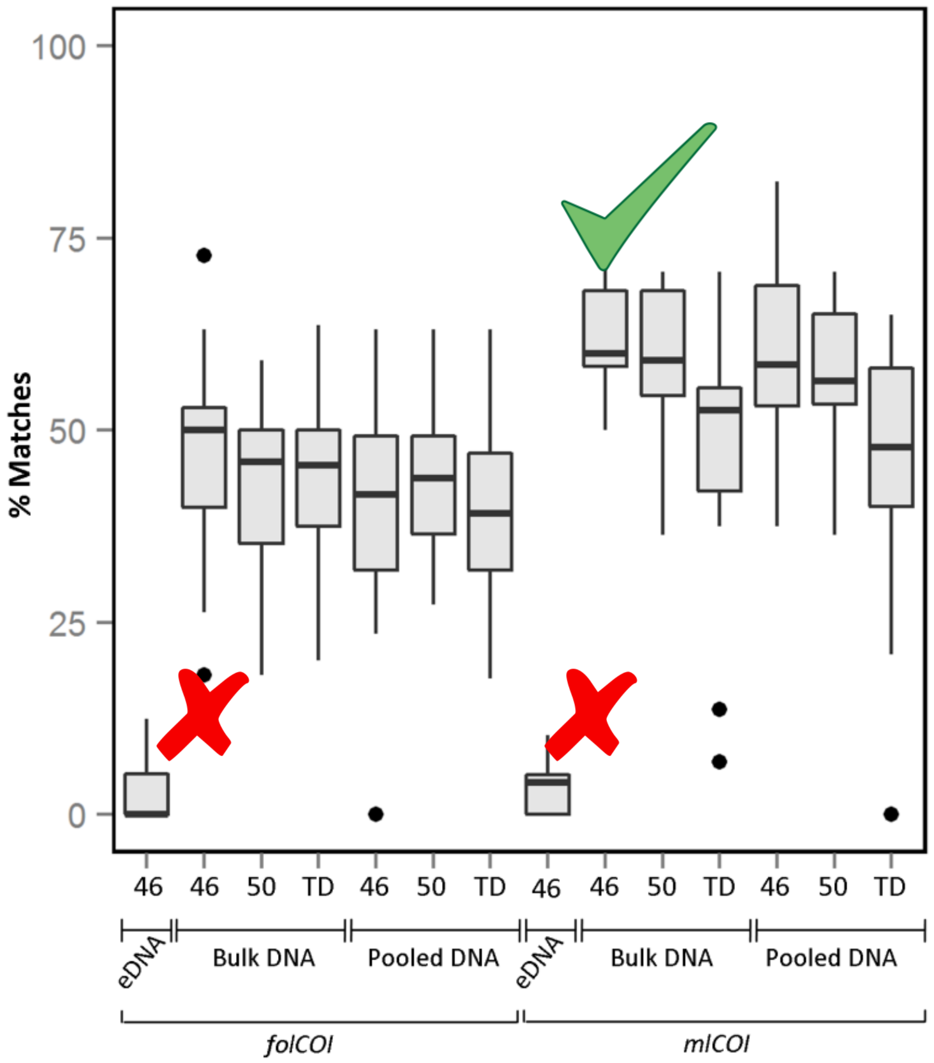
Molecular Taxonomy ↔ Visual Taxonomy

Comparison of visual & metabarcoding based biotic index

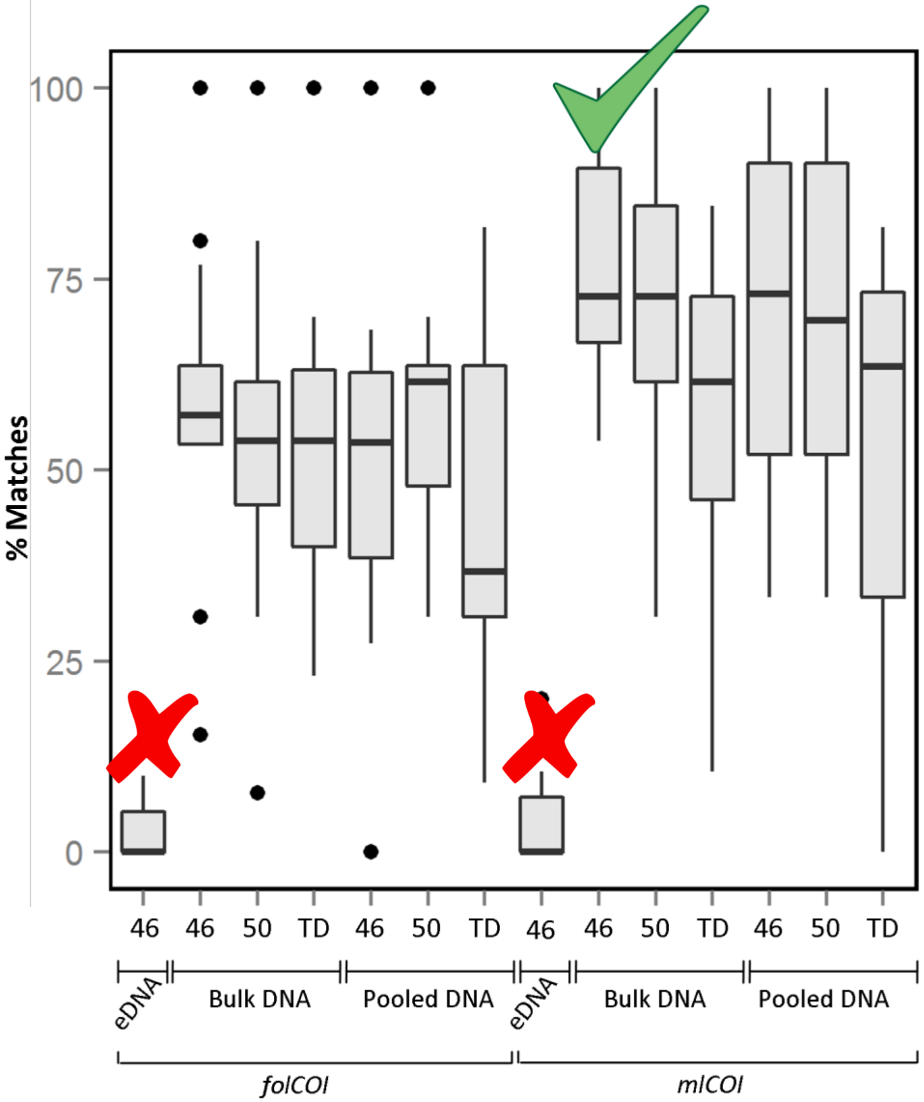
gAMBI ↔ AMBI

Transforming Science into Business

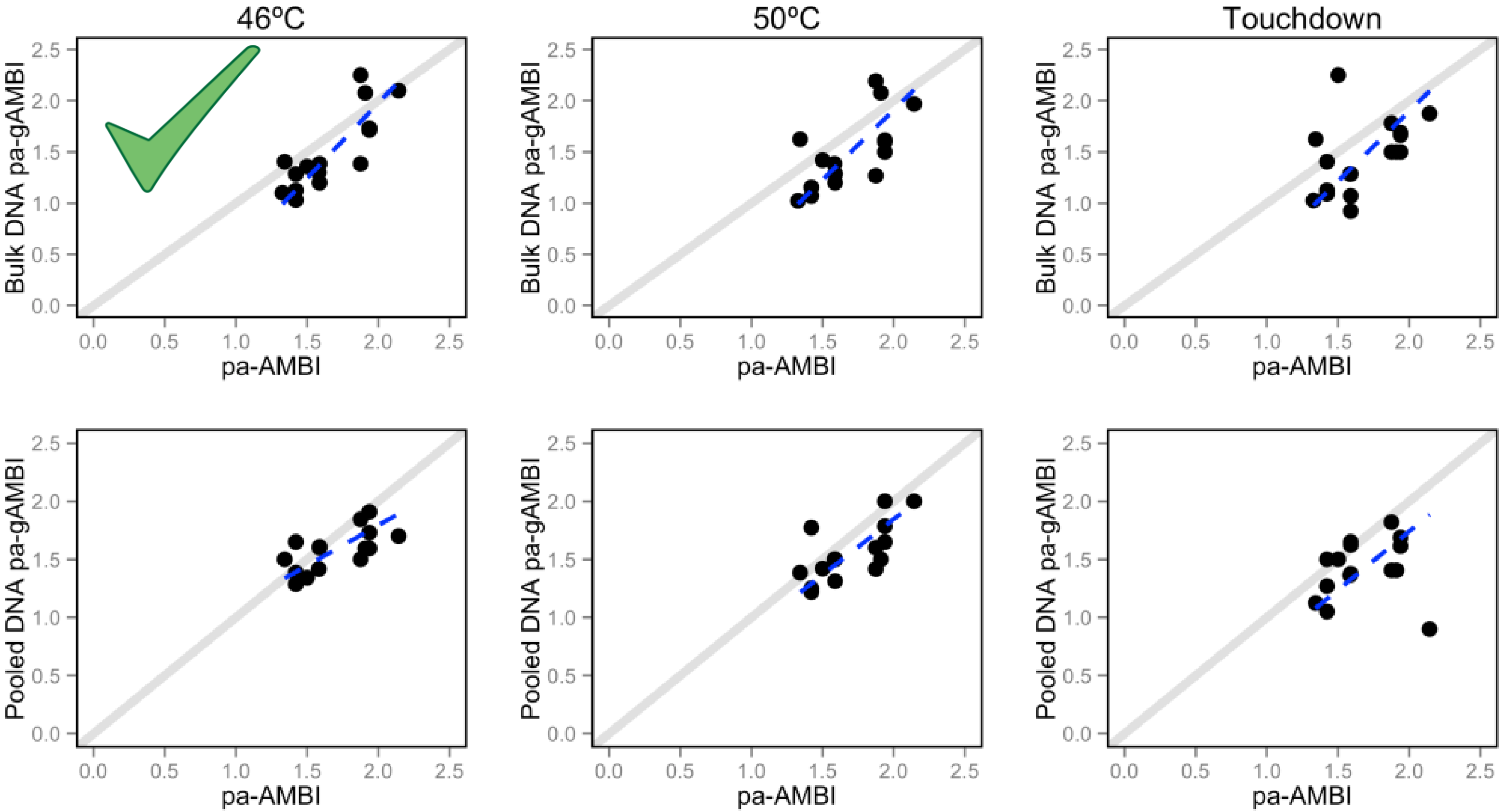
All taxonomic levels

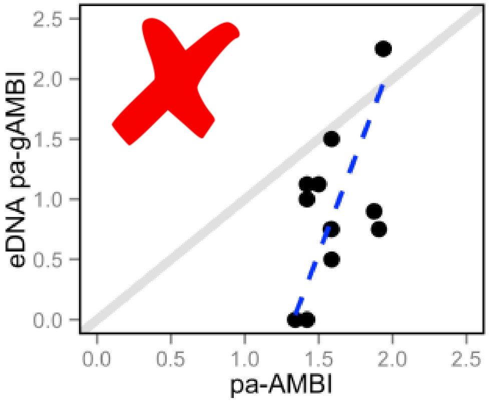


Only species level



AMBI vs gAMBI





mlCOI

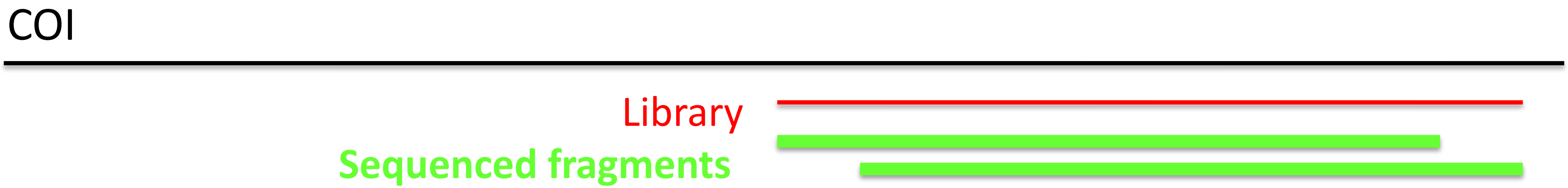
METHODOLOGICAL ASPECTS

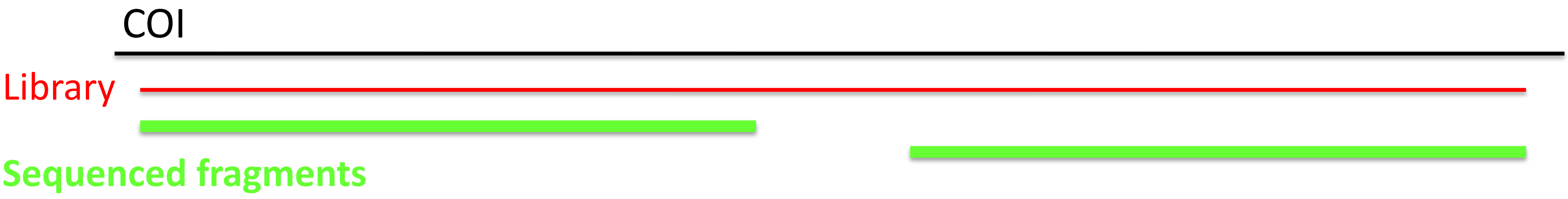
Amplify two different regions of the COI gene

Short – 313 bp

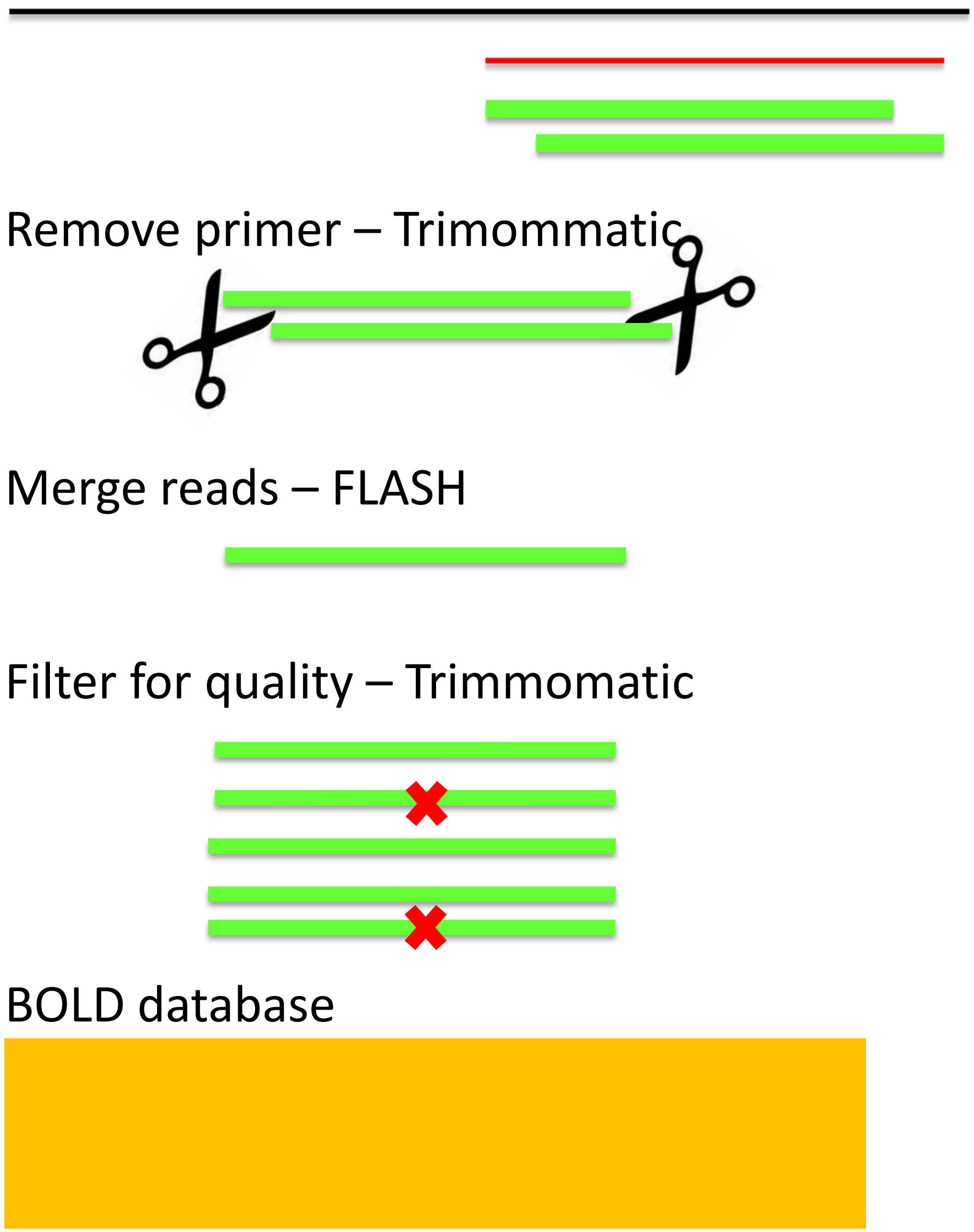
Long – 618 bp

Sequence on the MiSeq – 300 bp PE





APPROACH TAKEN

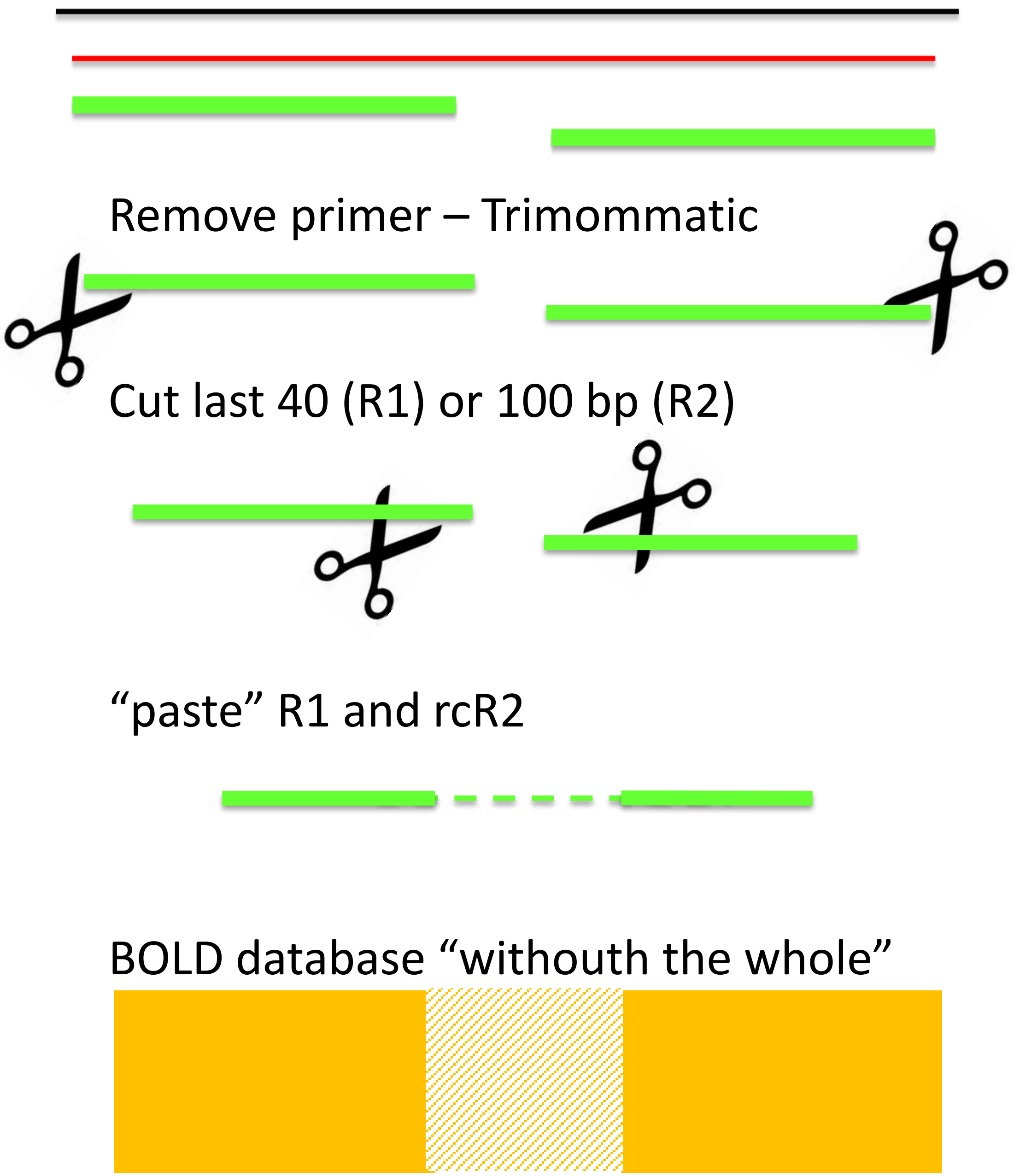


Remove primer – Trimommatic

Merge reads – FLASH

Filter for quality – Trimmomatic

BOLD database



Remove primer – Trimommatic

Cut last 40 (R1) or 100 bp (R2)

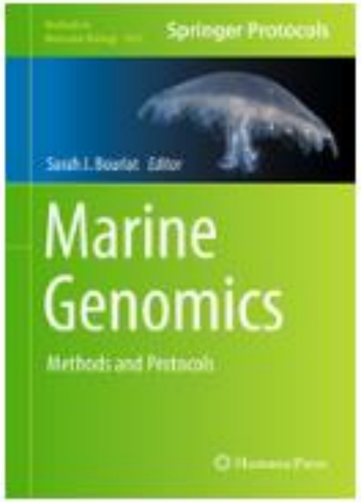
“paste” R1 and rcR2

BOLD database “withouth the whole”

Mothur

- Align to database
- Chimera removal
- Singleton removal
- Taxonomic assignment
- Classify into phylotypes

Transforming
Science into
Business



Springer Protocols

Sarah J. Bourlat Editor

Marine
Genomics

Methods and Protocols

Humana Press

Marine Genomics

Methods and Protocols

Editors: **Bourlat**, Sarah J. (Ed.)

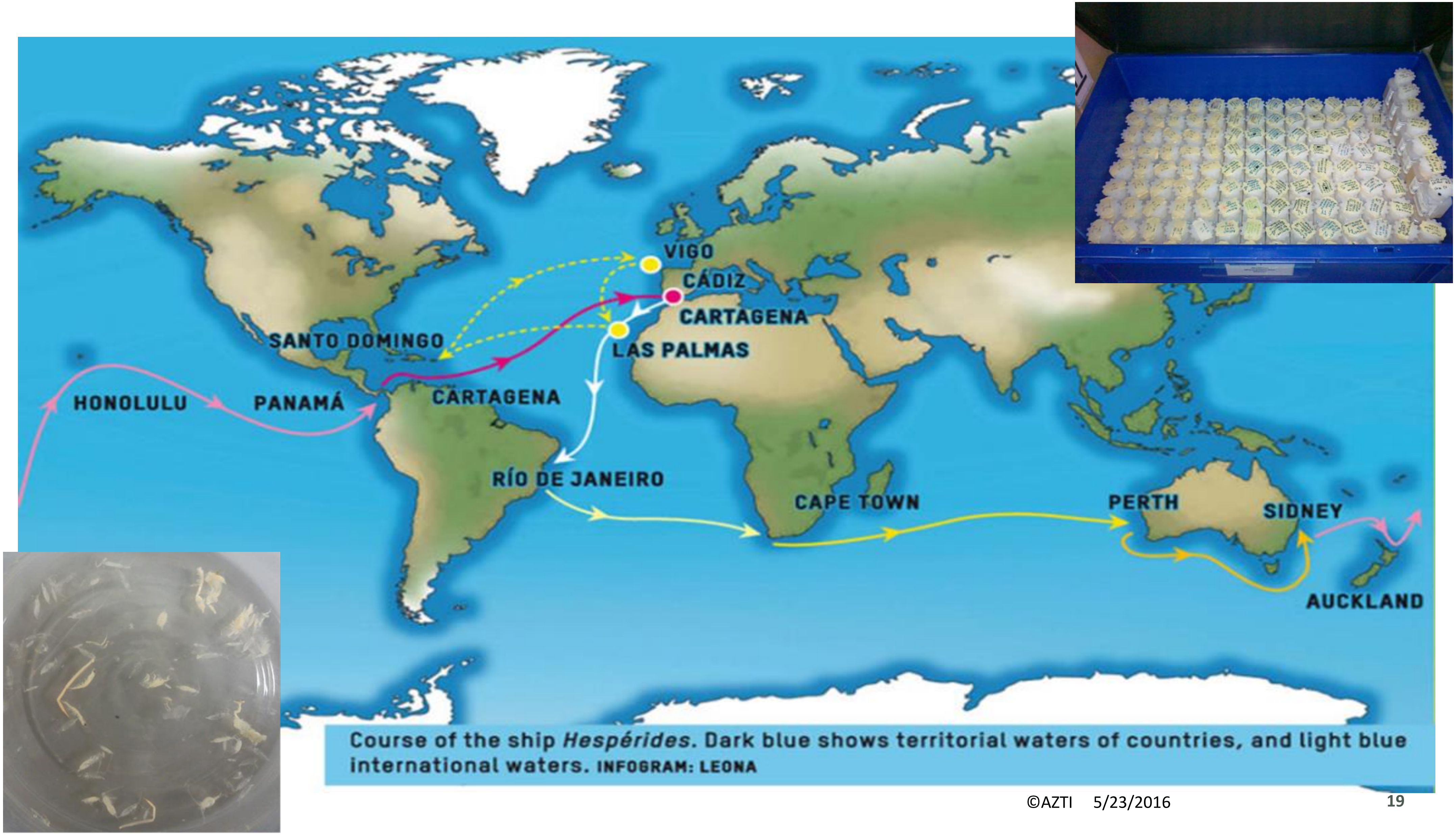
Chapter 16

Analysis of Illumina MiSeq Metabarcoding Data: Application to Benthic Indices for Environmental Monitoring

Eva Aylagas and Naiara Rodríguez-Ezpeleta

Metabarcoding of mesozooplankton

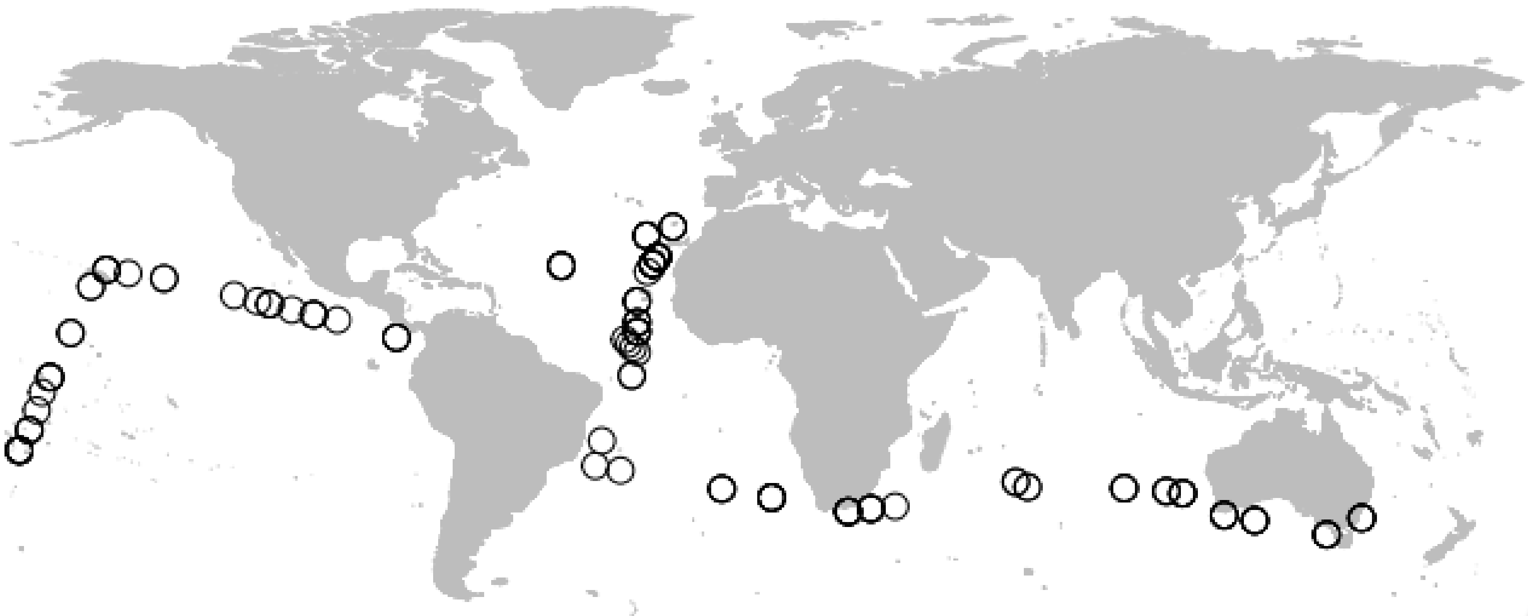
- Samples from the MALASPINA expedition

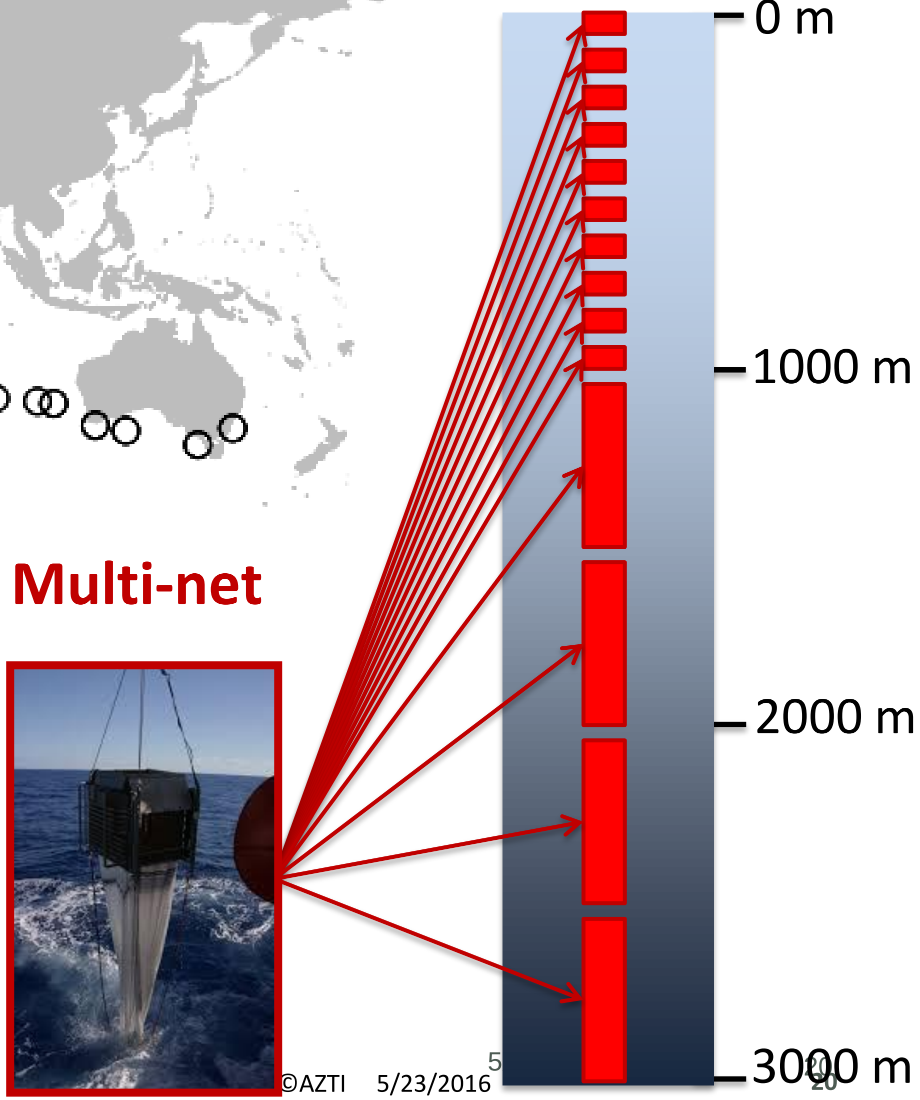


Course of the ship *Hespérides*. Dark blue shows territorial waters of countries, and light blue international waters. INFOGRAM: LEONA

Mesozooplankton

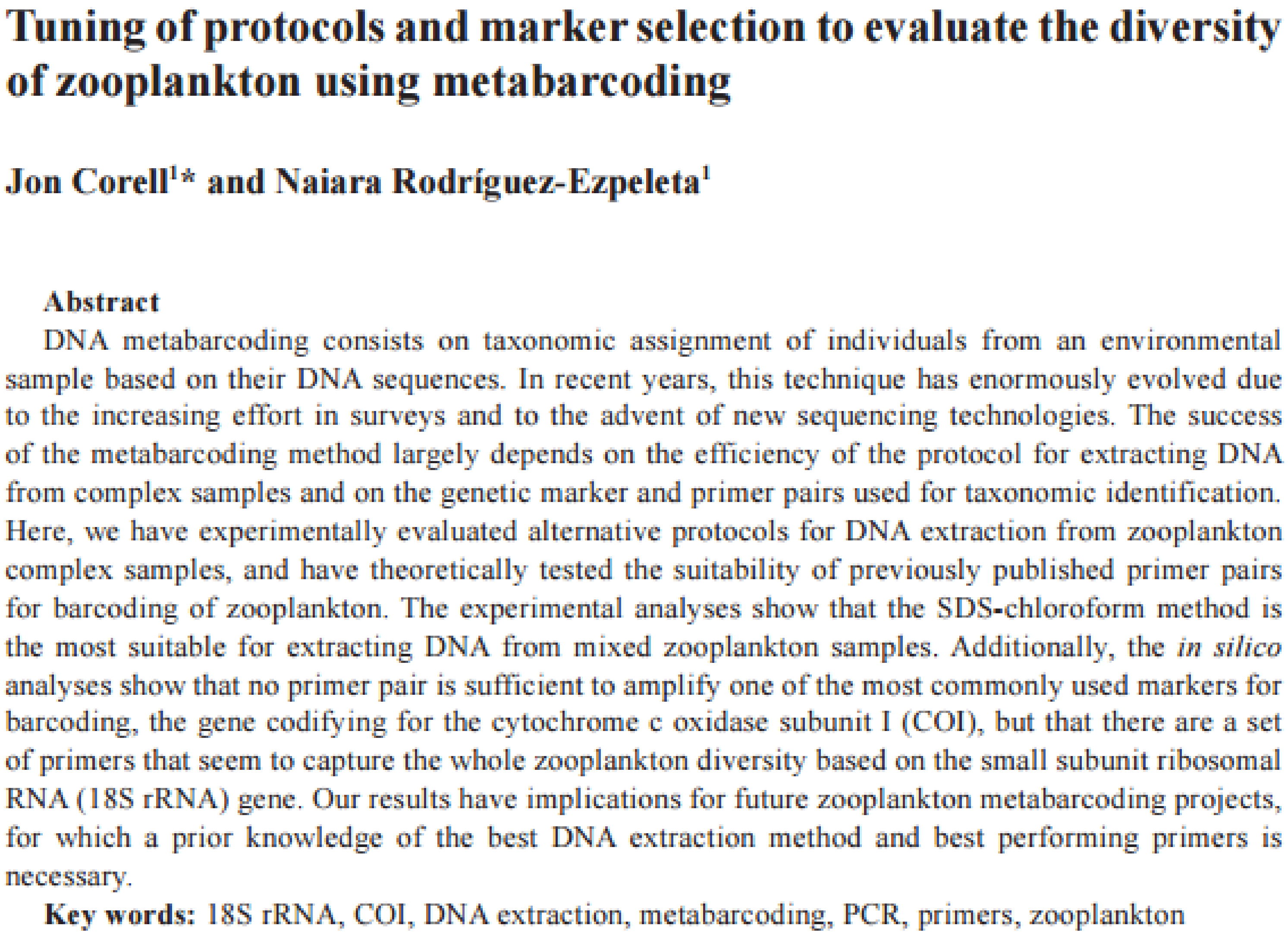
SAMPLES





- Mesozooplankton (133)
  - SIZE: 300 - 5000 µm
  - DEPTH: 0 to 3000 m

19 | Revista de Investigación Marina, 2014, 21(2)



Tuning of protocols and marker selection to evaluate the diversity of zooplankton using metabarcoding

Jon Corell[1]* and Naiara Rodríguez-Ezpeleta[1]

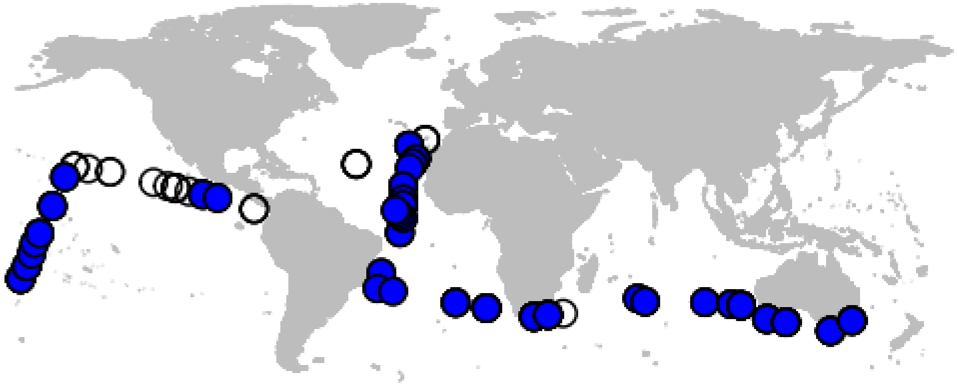
Abstract

DNA metabarcoding consists on taxonomic assignment of individuals from an environmental sample based on their DNA sequences. In recent years, this technique has enormously evolved due to the increasing effort in surveys and to the advent of new sequencing technologies. The success of the metabarcoding method largely depends on the efficiency of the protocol for extracting DNA from complex samples and on the genetic marker and primer pairs used for taxonomic identification. Here, we have experimentally evaluated alternative protocols for DNA extraction from zooplankton complex samples, and have theoretically tested the suitability of previously published primer pairs for barcoding of zooplankton. The experimental analyses show that the SDS-chloroform method is the most suitable for extracting DNA from mixed zooplankton samples. Additionally, the *in silico* analyses show that no primer pair is sufficient to amplify one of the most commonly used markers for barcoding, the gene codifying for the cytochrome c oxidase subunit I (COI), but that there are a set of primers that seem to capture the whole zooplankton diversity based on the small subunit ribosomal RNA (18S rRNA) gene. Our results have implications for future zooplankton metabarcoding projects, for which a prior knowledge of the best DNA extraction method and best performing primers is necessary.

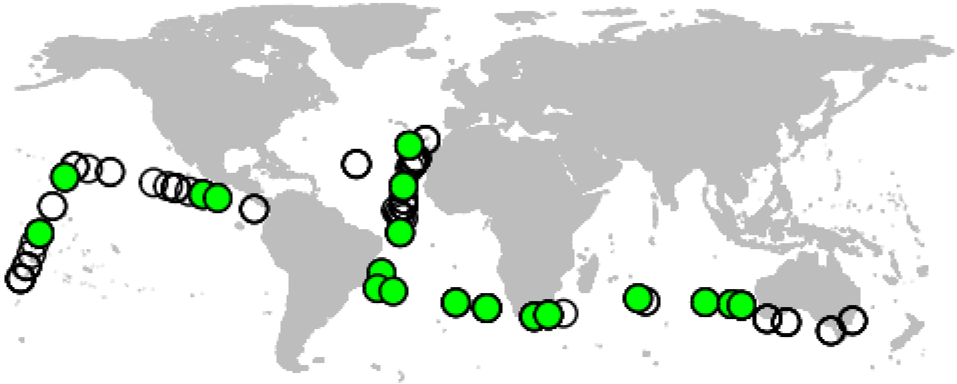
Key words: 18S rRNA, COI, DNA extraction, metabarcoding, PCR, primers, zooplankton

18S rRNA (Machida #1, #2RC) - 115

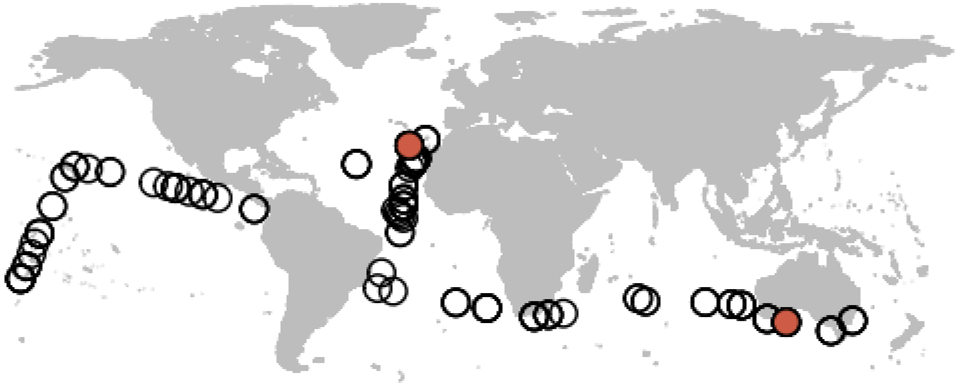
COI (Leray) - 89



COI (Meyer) - 28

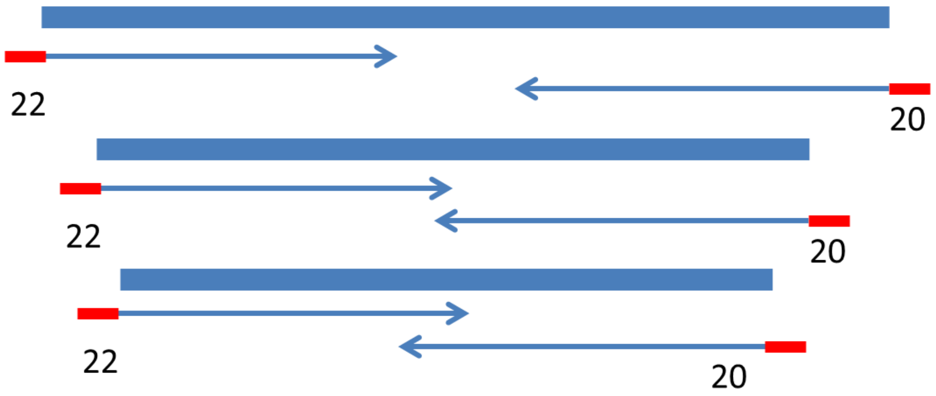


18S (metagenome) - 5



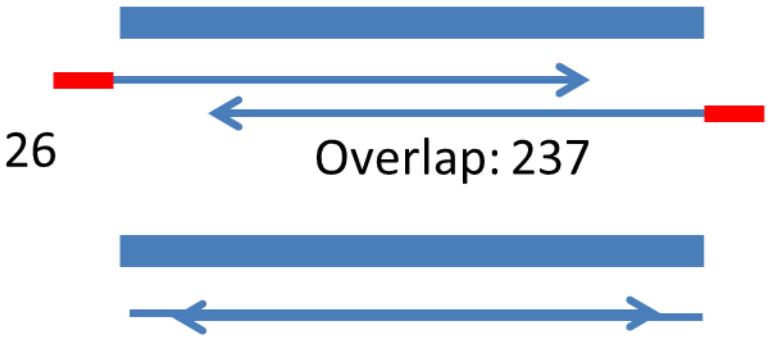
18S rRNA (Machida #1, #2RC) - 115

V4 region (variable length)



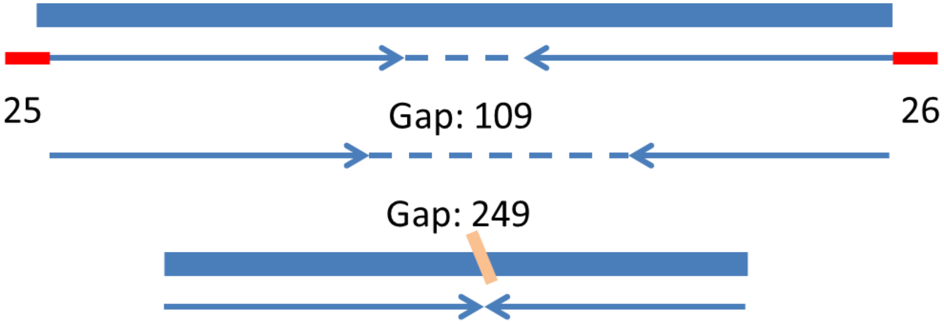
COI (Leray) - 89

Leray region (313bp)



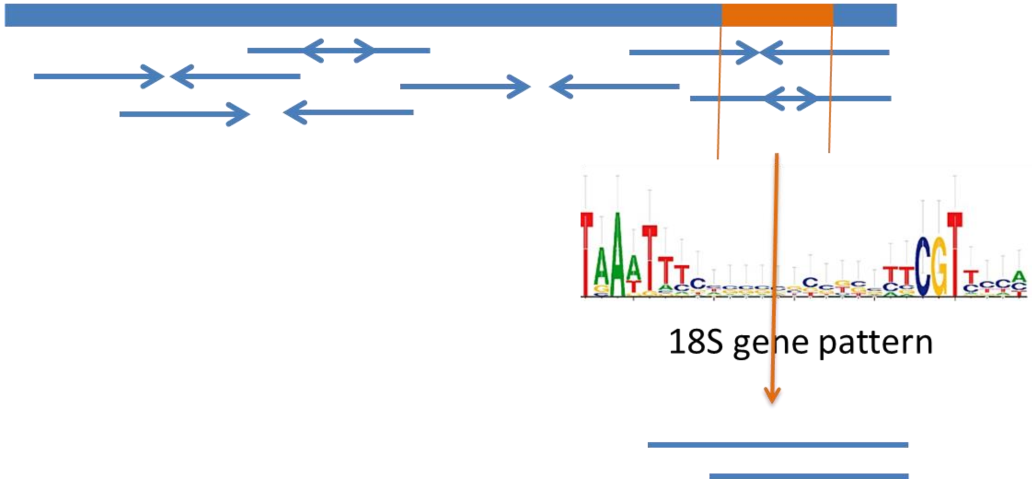
COI (Meyer) - 28

Folmer region (658 bp)



18S (metagenome) - 5

Genome



Some general conclusions

- Number of OTU very variable even if excluding miTags
- Number of taxa not related to number of OTU
- Lowest number of unassigned reads with Machida
- Highest number of non-Metazoa reads with Machida
- High number of unclassified Metazoa with COI (particularly with Leray)
    - Phyla not present in BOLD but present in SILVA?
    - New phyla amplified with COI but not with 18S?
    - More sequencing errors in COI than in 18S?
- Deepest taxonomic level with Machida

- Groups preferentially amplified by each marker
- Groups completely missed by some marker
- Once “clean”, comparison between markers
- Overlapp between markers?
- Need to combine to get a complete overview?
- Compare with visual taxonomy!
- **Biology!**

- Slight differences in the metabacoding protocol can translate into drastically different results
- Comparability requires standardization
- Absolute results are not possible
- Are relative results possible?
- **Metabarcoding is powerful, but power without control is nothing**
- **Use it, but be critical!**

azti tecnalia

Transforming Science into Business

www.azti.es | www.alimentatec.com | www.itsasnet.com
T. +34 94 657 40 00 | info@azti.es

Txatxarramendi ugartea z/g
48395 Sukarrieta, Bizkaia (SPAIN)

Herrera Kaia, Portualdea z/g
20110 Pasaia, Gipuzkoa (SPAIN)

Astondo Bidea, Edificio 609
Parque Tecnológico de Bizkaia
48160 Derio, Bizkaia (SPAIN)