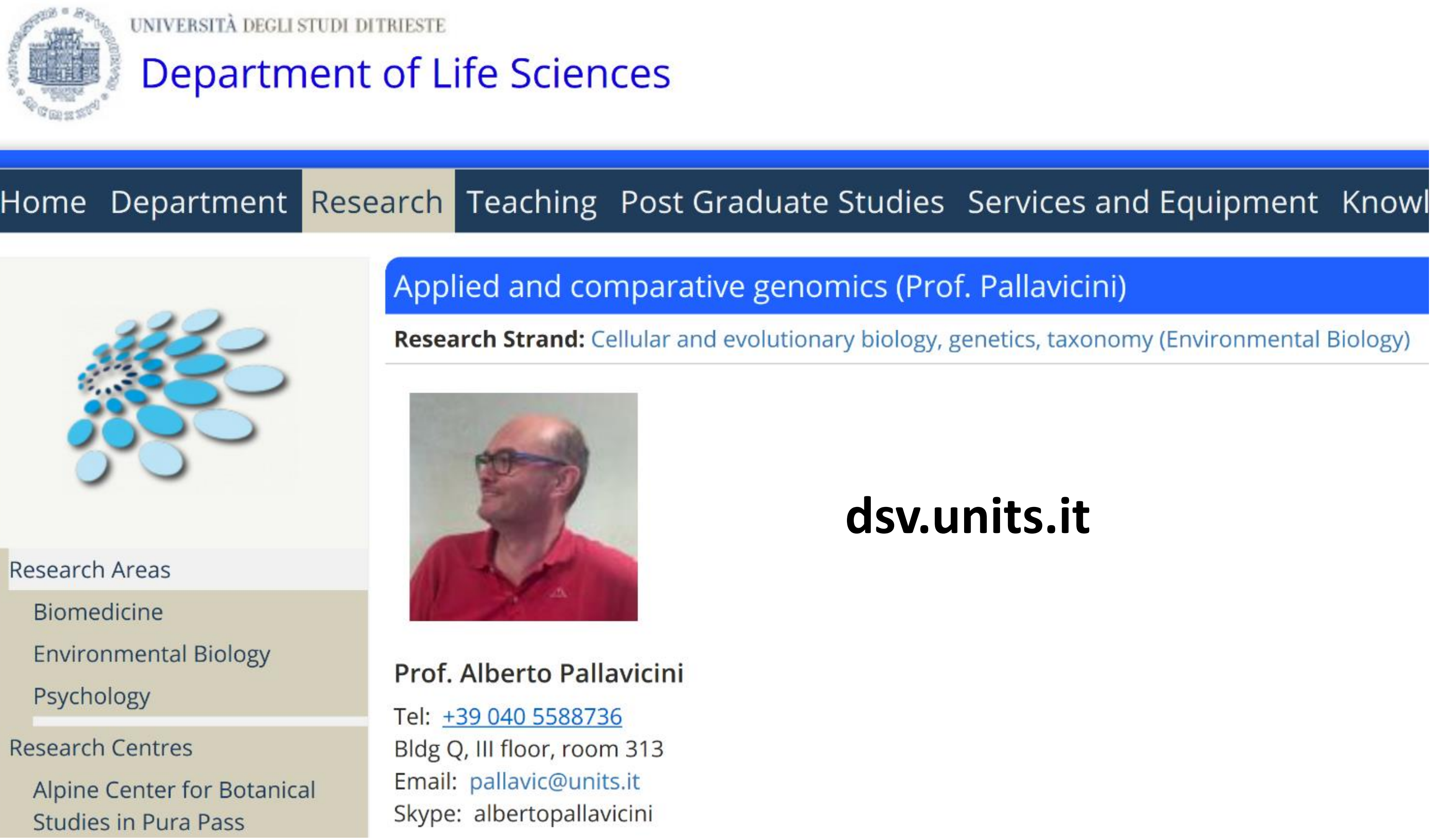


Department of Life Sciences

Applied and comparative genomics (Prof. Pallavicini)

Research Strand: Cellular and evolutionary biology, genetics, taxonomy (Environmental Biology)

Prof. Alberto Pallavicini

Tel: +39 040 5588736
Bldg Q, III floor, room 313
Email: pallavic@units.it
Skype: albertopallavicini

dsv.units.it

- medium size university ▶ 16,000 students | 1,000 International
- 10 Departments
- +75 Study programmes (bachelor, master, Ph.D)
- 4 Cities - 4 campuses

Area

Social Science and Humanities
Technology and Science
Life Science

ANVUR ▶ National Accreditation

Set up in 1877 as School of Commerce, in 1924 granted the title of Royal University of Economics and Business Studies of Trieste, in 1938, its name was changed to “Regia Università degli Studi” (Royal University of Trieste).

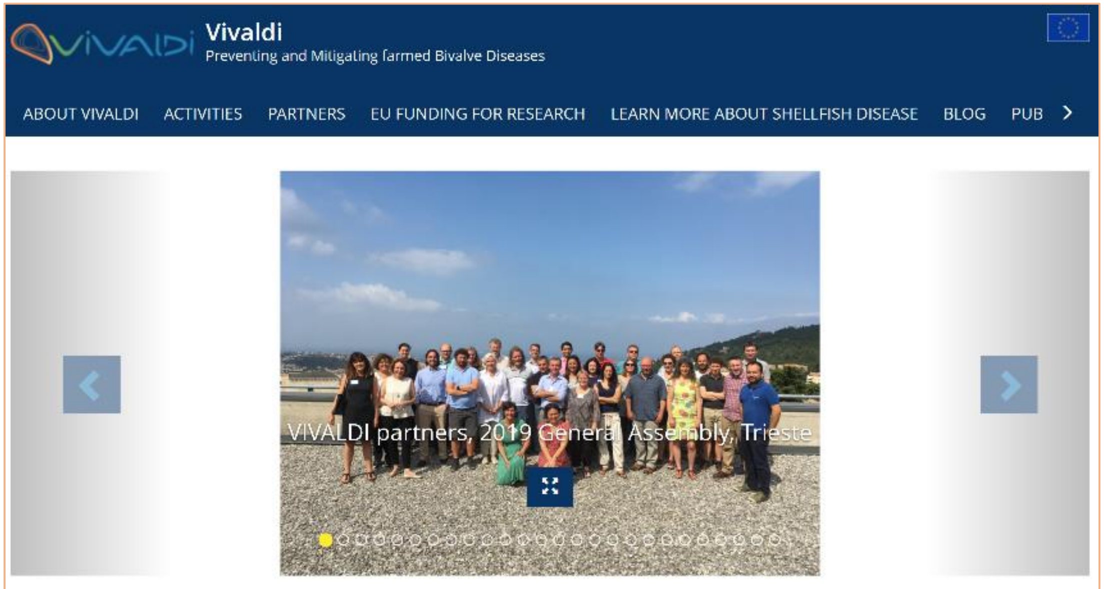
UNIVERSITÀ
DEGLI STUDI DI TRIESTE

ESOF2020
EUROSCIENCE OPEN FORUM
TRIESTE

Genomics of bivalve immunity

- Our research focuses on the comparative genomics of the animal immune system.

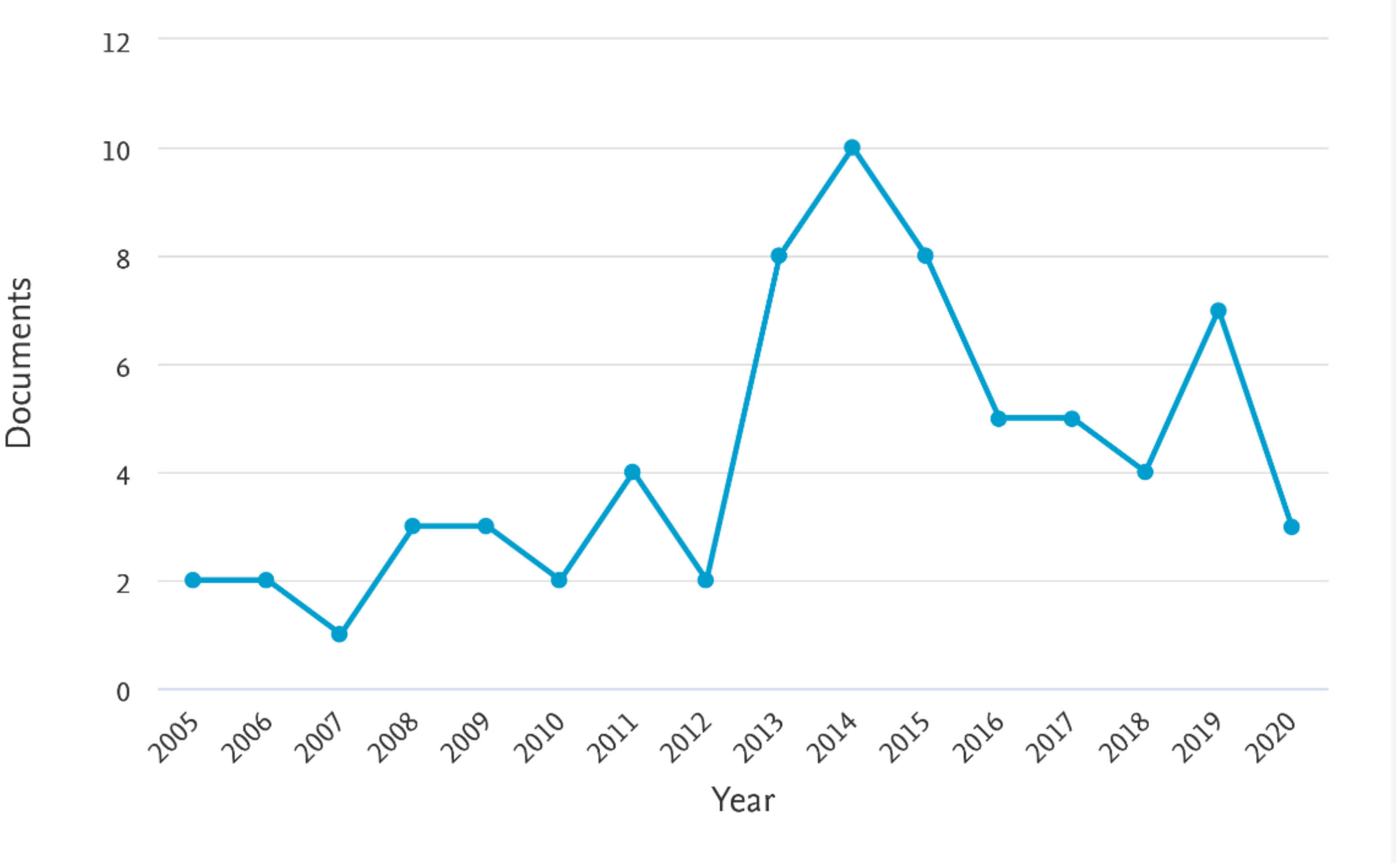
VIVALDI (Preventing and Mitigating Farmed Bivalve Diseases) is a European scientific project, financed in the framework of the EU Research and Development programme called Horizon 2020



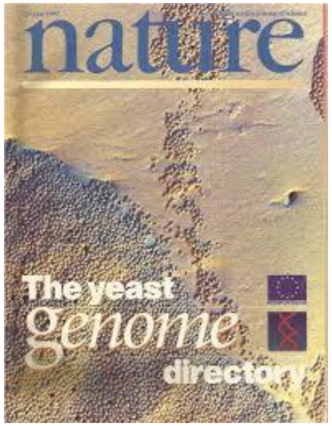
Vivaldi
Preventing and Mitigating farmed Bivalve Diseases

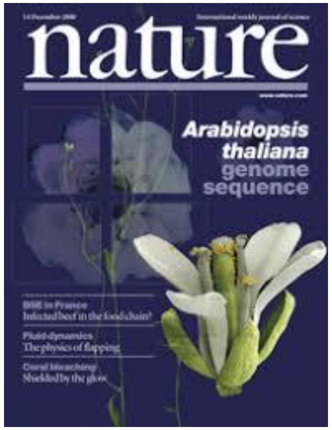
VIVALDI partners, 2019 General Assembly, Trieste

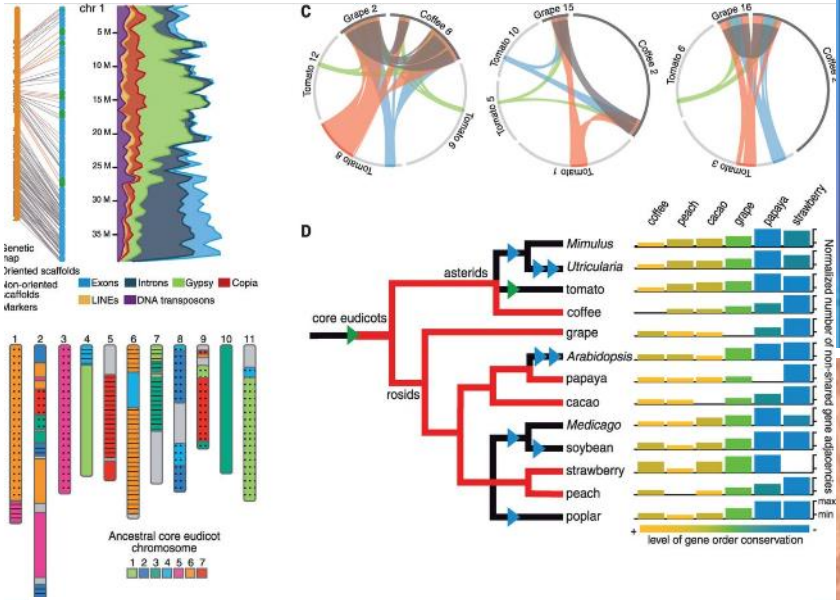
(TITLE-ABS-KEY (transcriptome OR "gene expression") AND AUTHOR-NAME (pallavicini AND a))

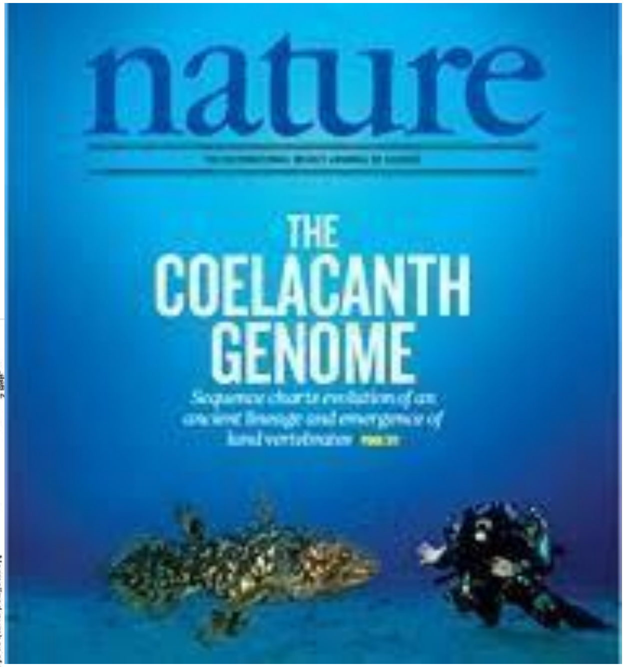


Comparative and evolutionary genomics









Coffee Genome Hub
a genomics and genetics resource for coffee

Prokaryotic and eukaryotic metabarcoding

Zooplankton metabarcoding

SCIENTIFIC REPORTS

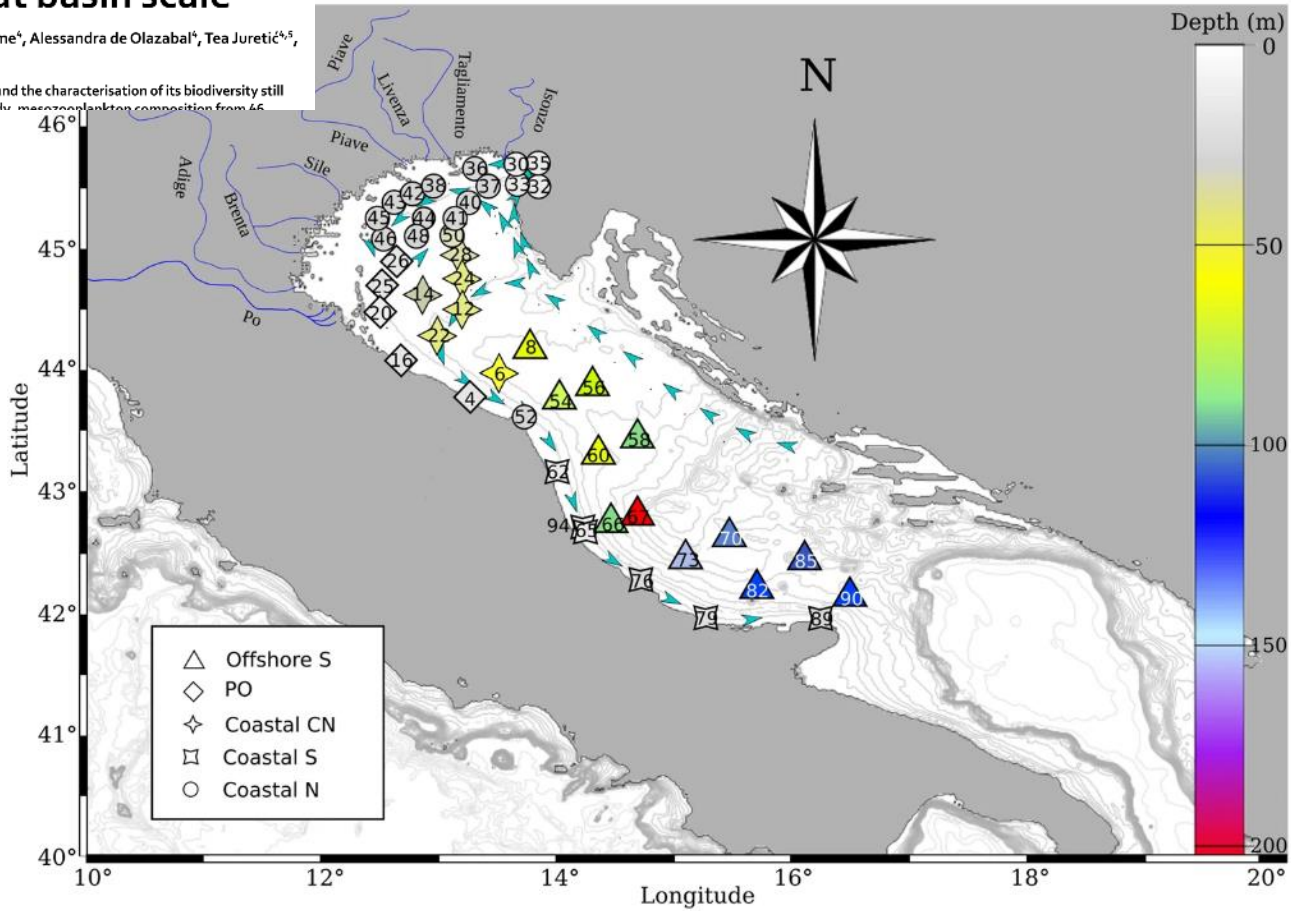
OPEN

Multi-marker metabarcoding approach to study mesozooplankton at basin scale

Received: 15 March 2018
Accepted: 25 July 2018

Sergio Stefanni[1], David Stanković[2,3], Diego Borme[4], Alessandra de Olazabal[4], Tea Juretić[4,5], Alberto Pallavicini[2] & Valentina Tirelli[4]

Zooplankton plays a pivotal role in marine ecosystems and the characterisation of its biodiversity still represents a challenge for marine ecologists. In this study, mesozooplankton composition from 46

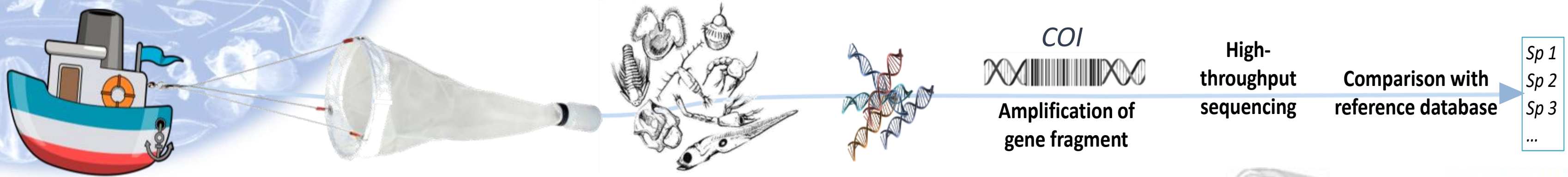


DNA metabarcoding for zooplankton biodiversity assessment

- important role in marine food webs and in marine carbon cycle
- most marine species have a planktonic life stage

- morphological biodiversity assessments are labour intensive

➢ increasing need of faster, more cost-effective and more objective methods to characterize the spatio-temporal variability



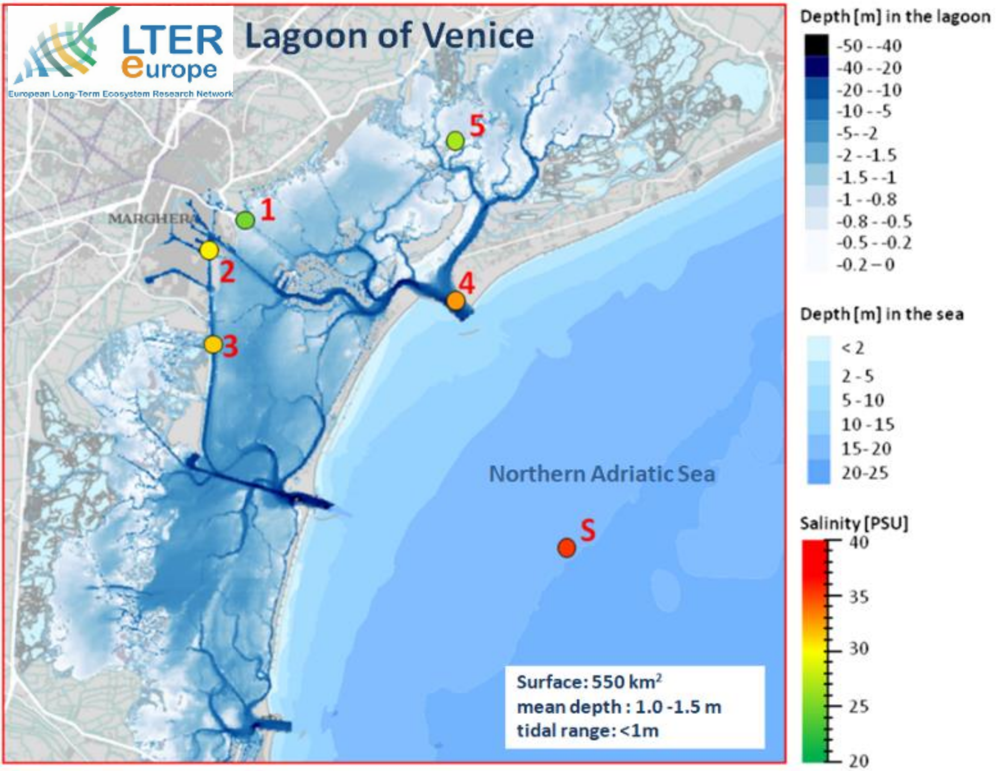
COI

Amplification of gene fragment

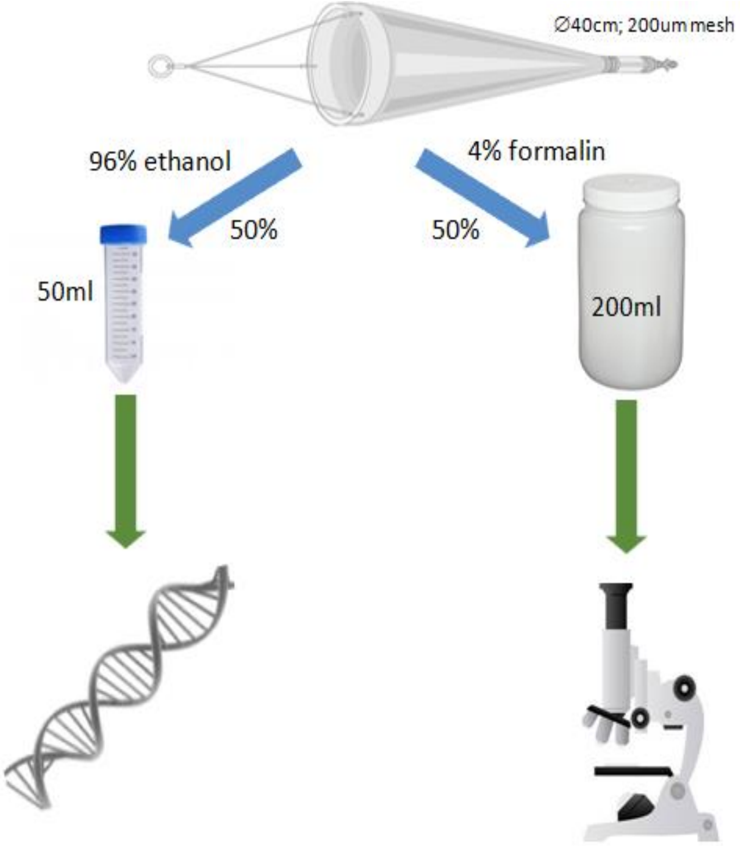
High-throughput sequencing

Comparison with reference database

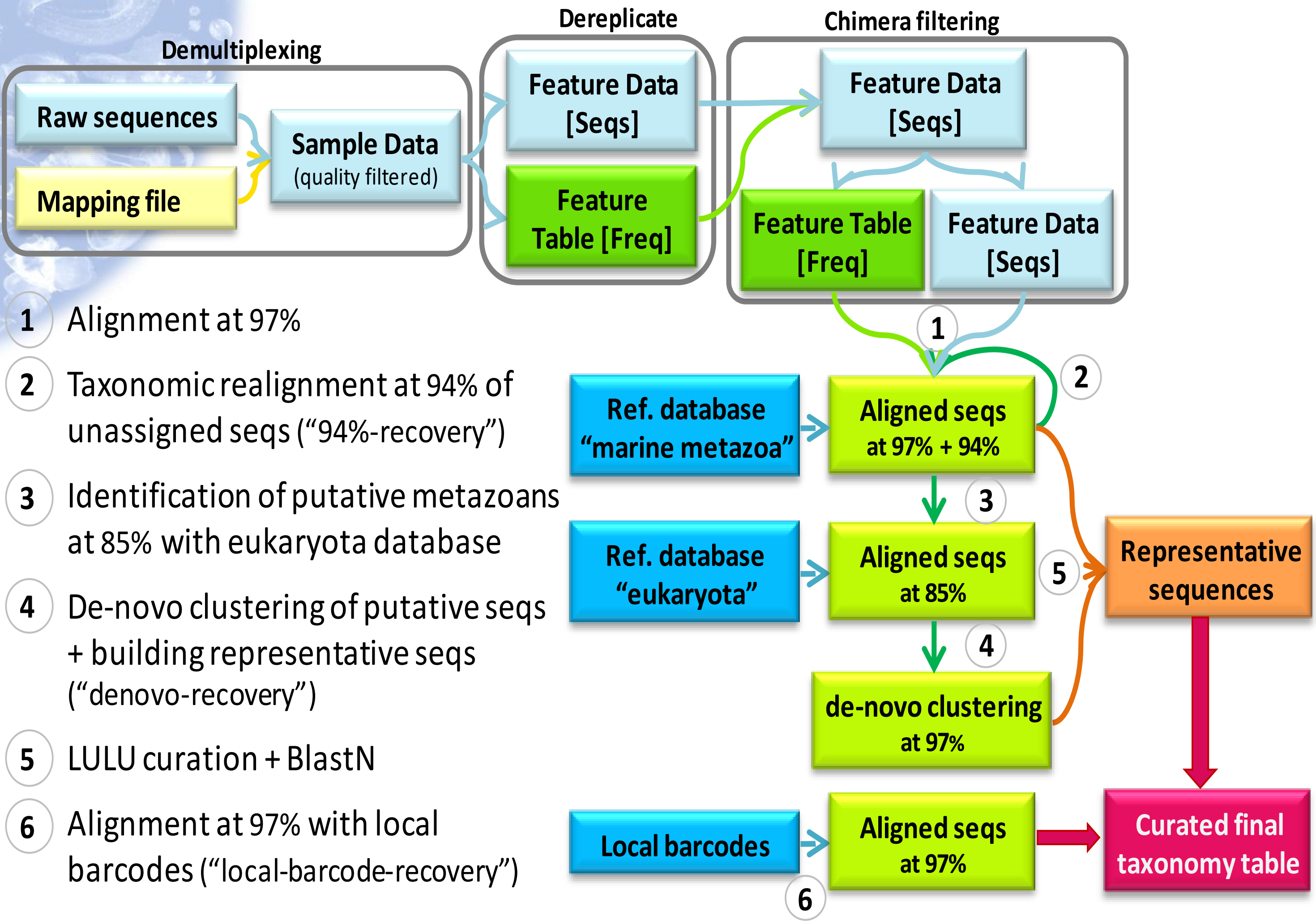
Sp 1
Sp 2
Sp 3
...



Surface: 550 km^2
mean depth: 1.0 -1.5 m
tidal range: <1m



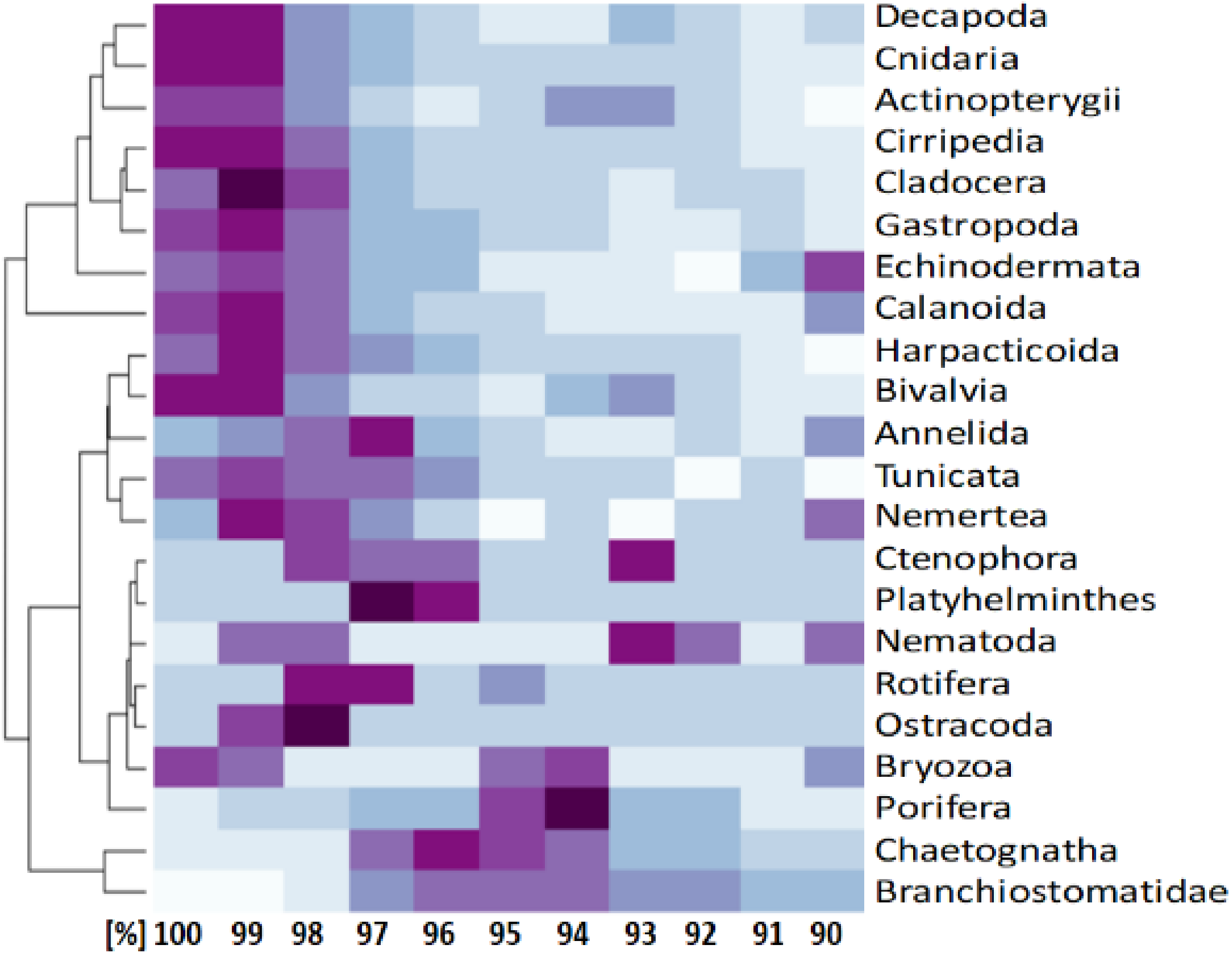
DNA metabarcoding for zooplankton biodiversity assessment



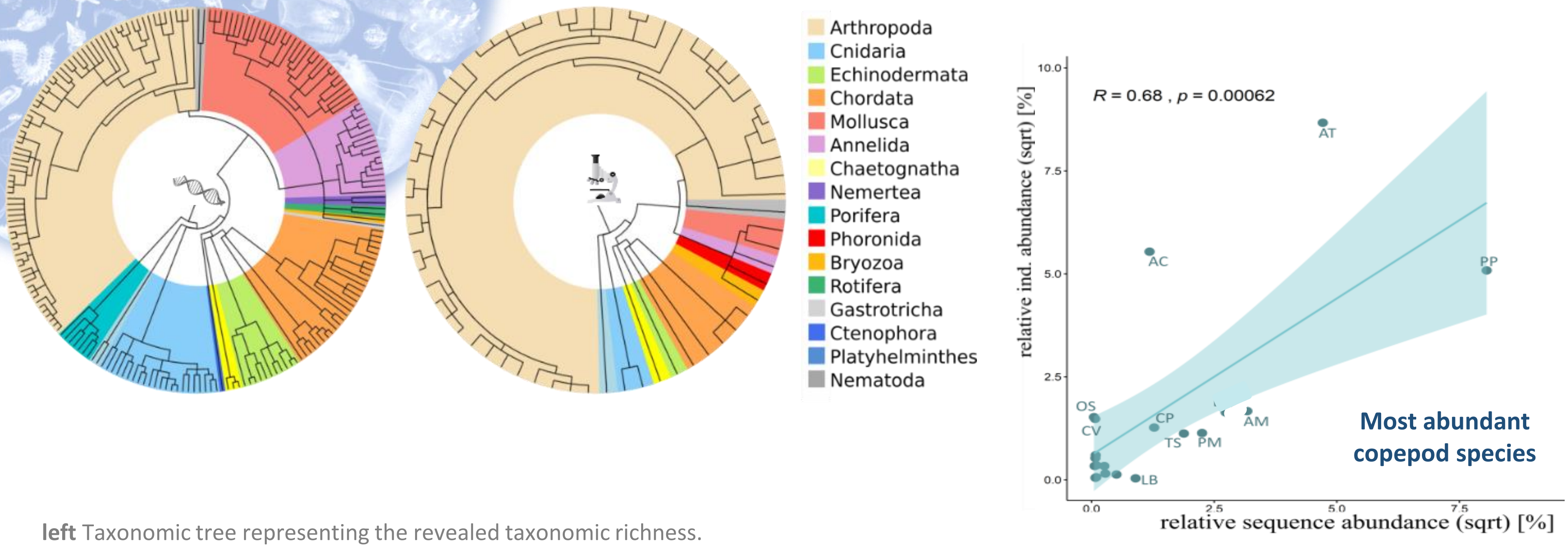
1. Alignment at 97%
2. Taxonomic realignment at 94% of unassigned seqs (“94%-recovery”)
3. Identification of putative metazoans at 85% with eukaryota database
4. De-novo clustering of putative seqs + building representative seqs (“denovo-recovery”)
5. LULU curation + BlastN
6. Alignment at 97% with local barcodes (“local-barcode-recovery”)

Flowchart of the bioinformatic multistep approach.

Heatmap based on the abundance of assignments at similarity thresholds from 100% to 90% for different taxonomic groups.



DNA metabarcoding for zooplankton biodiversity assessment



Most abundant copepod species

left Taxonomic tree representing the revealed taxonomic richness.

right Relative abundance of copepod reads of metabarcoding data and of individual counts of morphological data (% based on square rooted data)

+

- **Higher estimation of taxonomic richness with DNA metabarcoding compared to morphological analysis (especially mero- and ichthyoplankton and cryptic species).**
- **Both analyses provide similar spatio-temporal patterns of zooplankton community.**
- **No. of sequences and abundance counts show correlations for various taxonomic groups.**
- **Multistep approach based on the recovery of all putative sequences enables a more reliable taxonomic assignment.**

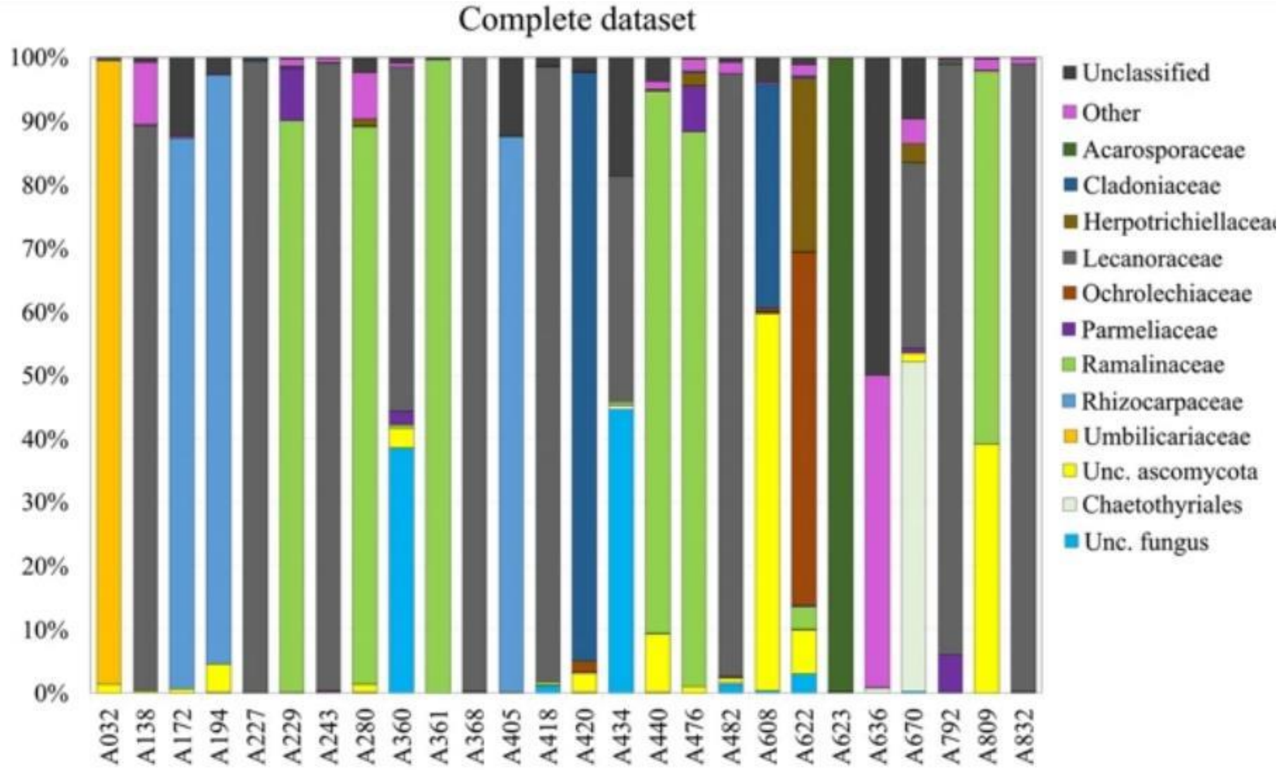
−

- **information about population structure gets lost (sex, larval stage).**
- **need of reliable, geographically comprehensive reference database**

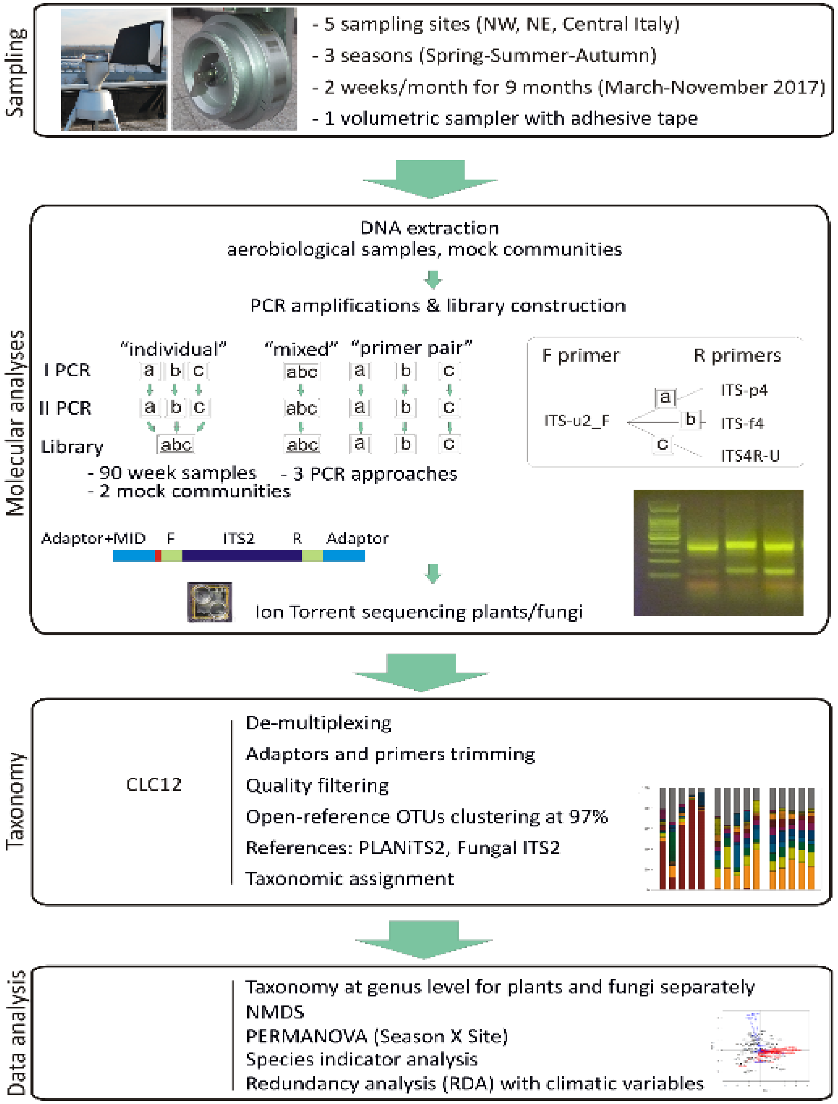
ITS2 metabarcoding analysis complements lichen mycobiome diversity data

Elisa Banchi[1] · David Stankovic[1,2] · Fernando Fernández-Mendoza[3] · Fabrizia Gionechetti[1] · Alberto Pallavicini[1] · Lucia Muggia[1,3]

Received: 9 April 2018 / Revised: 12 May 2018 / Accepted: 4 June 2018 / Published online: 22 June 2018



DNA metabarcoding for the characterization of aerobiological samples



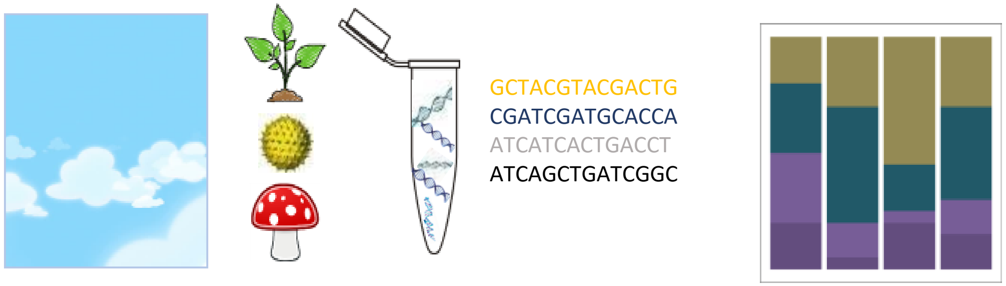
- 5 sampling sites (NW, NE, Central Italy)
- 3 seasons (Spring-Summer-Autumn)
- 2 weeks/month for 9 months (March-November 2017)
- 1 volumetric sampler with adhesive tape

- 90 week samples
- 2 mock communities
- 3 PCR approaches

De-multiplexing
Adaptors and primers trimming
Quality filtering
Open-reference OTUs clustering at 97%
References: PLANiTS2, Fungal ITS2
Taxonomic assignment

Taxonomy at genus level for plants and fungi separately
NMDS
PERMANOVA (Season X Site)
Species indicator analysis
Redundancy analysis (RDA) with climatic variables

FRA 2016-17 UNITS
COST ACTION



- **DNA extracted from tape with airborne particles**
- **First time of simultaneous amplification of plants and fungi**
- **Mock plant and fungal communities**
- **CLC Workbench and QIIME2 → similar results**
- **Up to 10X improvement of the standard microscopy detection**
- **Monitoring of pathogen, IAS, allergenic taxa**

Banchi et al. 2018; Banchi et al. 2019; Banchi et al. in review

Development of a new Plant database

DNA metabarcoding needs reliable databases

Most of the works dealing with plants use **NCBI downloaded sequences**, **clustered** with CD-HIT (Fu et al. 2012)

but

The sequences are not checked and can contain **errors in taxonomic assignment**, that affect the clustering, that group them and chose as representative the longest one, with its taxonomy, introducing biases in the databases.

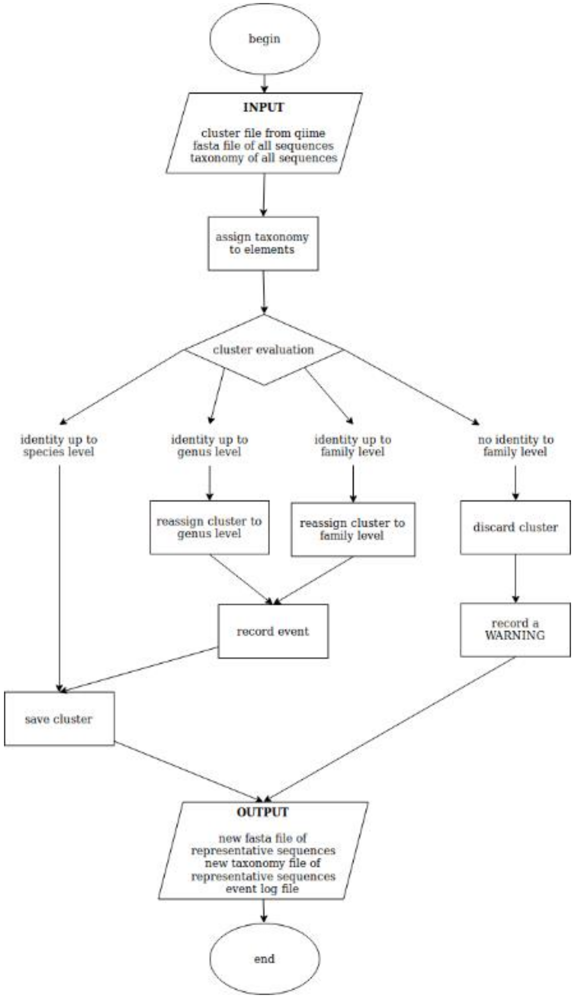
so

We have manually curated a **Plant ITS database** and developed a script ("better clustering for QIIME", **bc4q**) that in each cluster of similar sequences considers also the taxonomy

↓

PLANiTS: a curated sequence reference dataset for plant ITS DNA metabarcoding

- ✓ Up-to-date and curated (**no FUNGAL sequences!!**) ITS1, ITS2 and ITS database
- ✓ Available in a QIIME format
- ✓ Successfully applied in aerobiological samples



Banchi et al. 2019; Banchi et al. 2020

OXFORD
ACADEMIC

DATABASE
The Journal of Biological Databases and Curation

PLANiTS: a curated sequence reference dataset for plant ITS DNA metabarcoding

Elisa Banchi[1,2], Claudio G. Ametrano[1], Samuele Greco[1], David Stanković[1,3], Lucia Muggia[1,*] and Alberto Pallavicini[1,2,4,*]

[1]Department of Life Sciences, University of Trieste, via Giorgieri 5, 34127, Trieste, Italy [2]Division of Oceanography, National Institute of Oceanography and Applied Geophysics, via Piccard 54, 34151, Trieste, Italy [3]Marine Biology Station, National Institute of Biology, Fornače 41, Piran, Slovenia [4]Department of Biology and Evoliution of Marine Organisms, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Naples, Italy

*Corresponding author: Tel: +39 040 558 8825; Email: lmuggia@units.it
Correspondence may also be addressed to Alberto Pallavicini. Tel: +39 040 558 8736; Email: pallavic@units.it

Citation details: Banchi,E., Ametrano,C. G.., Greco,S. *et al.* PLANiTS: a curated sequence reference dataset for plant ITS DNA metabarcoding. *Database* (2020) Vol. 2020: article ID baz155; doi:10.1093/database/baz155

Received 5 September 2019; Revised 11 December 2019; Accepted 23 December 2019

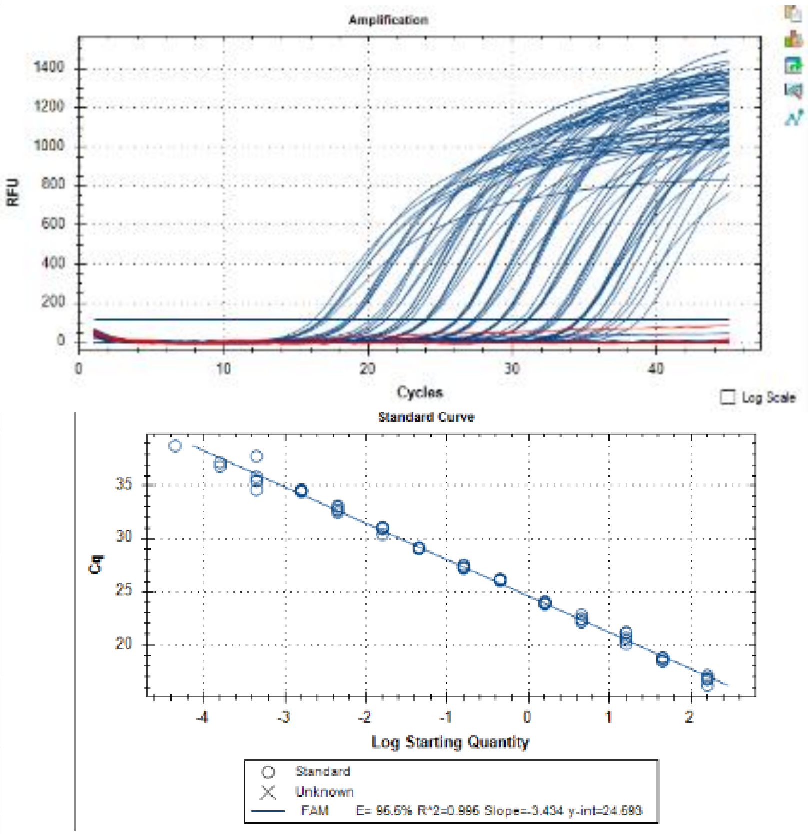
Abstract

DNA metabarcoding combines DNA barcoding with high-throughput sequencing to identify different taxa within environmental communities. The ITS has already been proposed and widely used as universal barcode marker for plants, but a comprehensive, updated and accurate reference dataset of plant ITS sequences has not been available so far. Here, we constructed reference datasets of Viridiplantae ITS1, ITS2 and entire ITS sequences including both Chlorophyta and Streptophyta. The sequences were retrieved from NCBI, and the ITS region was extracted. The sequences underwent identity check to remove misidentified records and were clustered at 99% identity to reduce redundancy and computational effort. For this step, we developed a script called ‘better clustering for QIIME’ (bc4q) to ensure that the representative sequences are chosen according to the composition of the cluster at a different taxonomic level. The three datasets obtained with the bc4q script are PLANiTS1 (100 224 sequences), PLANiTS2 (96 771 sequences) and PLANiTS (97 550 sequences), and all are pre-formatted for QIIME, being this the most used bioinformatic pipeline for metabarcoding analysis. Being curated and updated reference databases, PLANiTS1, PLANiTS2 and PLANiTS are proposed as a reliable, pivotal first step for a general standardization of plant DNA metabarcoding studies. The bc4q script is presented as a new tool useful in each research dealing with sequences clustering.

CA18226 - New approaches in detection of pathogens and aeroallergens

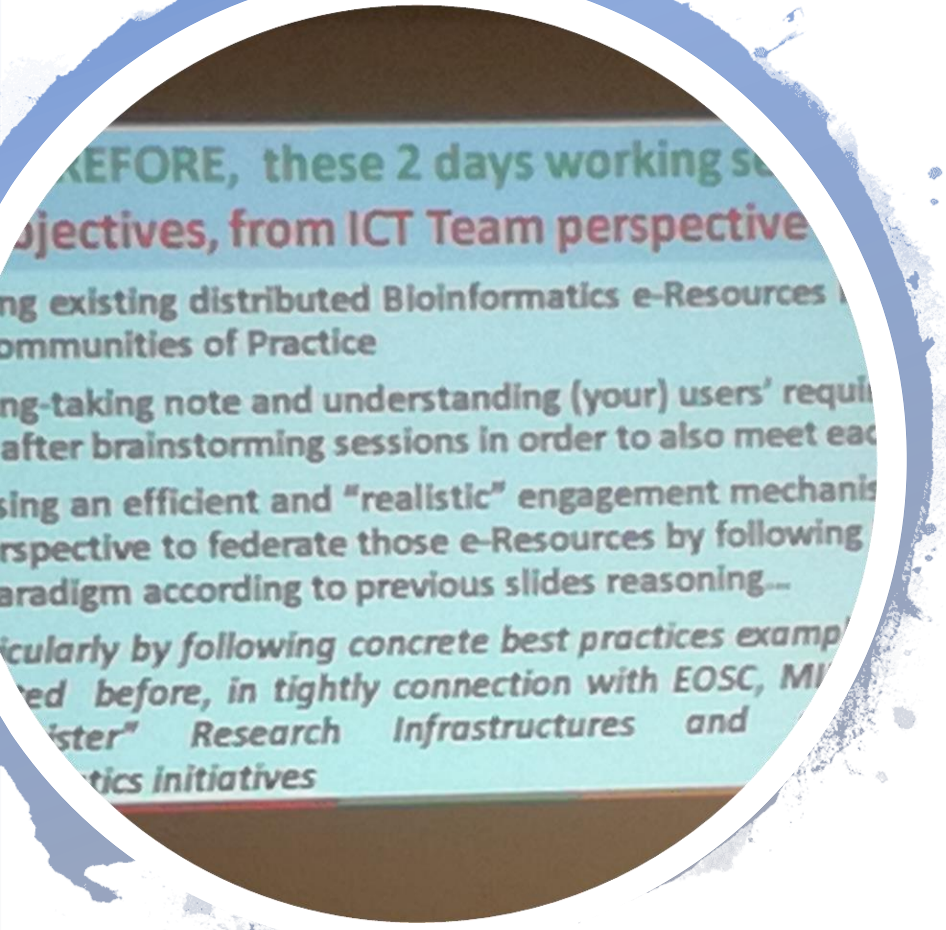
eDNA analysis for conservation and environmental monitoring purposes

ID Species	Group	Status of the system
Austropotamobius pallipes	Crustacean (native)	✓ validated
Austropotamobius torrentium	Crustacean (native)	✓ validated
Procambarus clarkii	Crustacean (invasive)	✓ validated
Faxonius limosus	Crustacean (invasive)	✓ validated
Pacifastacus leniusculus	Crustacean (invasive)	✓ validated
Pontastacus leptodactylus	Crustacean (invasive)	Under validation
Procambarus fallax fallax	Crustacean (invasive)	Under validation
Pseudorasbora parva	Fish (invasive)	✓ validated
Alosa fallax	Fish (native)	Under validation
Acipenser naccarii	Fish (native)	Under validation
Haplosporidium pinnae	Protist (pathogen)	Under validation



Reference database of 22 freshwater fish species inhabiting FVG

ICT team perspective



EFORE, these 2 days working s

jectives, from ICT Team perspective

ng existing distributed Bioinformatics e-Resources
ommunities of Practice

ng-taking note and understanding (your) users' requi
after brainstorming sessions in order to also meet eac

sing an efficient and "realistic" engagement mechanis
rspective to federate those e-Resources by following
aradigm according to previous slides reasoning...

icularly by following concrete best practices examp
ed before, in tightly connection with EOSC, MI
ister" Research Infrastructures and
tics initiatives

- A management of HPC infrastructures more «bioinformatic friendly»
- For the actual “biodiversity trough DNA metabarcoding” analysis ICT resources are not limiting
- Standardization of reference databases
- The development of IA methods for the automatic taxonomic assignment