



## Digital Marine Resources Community

---

**BioData.pt Talks**  
**Online Presentation**  
**Cymon J. Cox & João Machado**

10<sup>th</sup> December 2020

# Status of Community Organisation

- ▶ BioData.pt is the national ELIXIR node, ELIXIR being the ESFRI RI for Bioinformatics in the EU
- ▶ BioData.pt Marine Metagenomics to mirror ELIXIR Marine Metagenomics Community
- ▶ In 2021 becoming the ELIXIR Microbiome Community, BioData.pt to follow change
- ▶ BioData.pt Digital Marine Resources Community will support the marine domain

# Overview

## Objectives

- i) Provide a forum for researchers working on digital data in the marine environment in PT
- ii) Provision of digital services, training, computational resources, and working spaces for marine biologists
- iii) Provide an on-going assessment and re-evaluation of the needs of the community



# Hardware Resources



Lenovo Flex System:  
12 Blade servers each 28 cores @2.6GHz, 128GB/1TB memory



OceanStor 2200 Storage:  
30TB redundant

## Provisioning

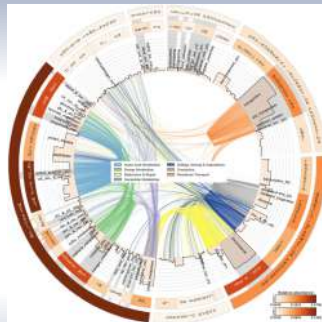
- \* Virtual machines / Docker, Singularity, BioConda etc
- \* HPC queues (SLURM)
- \* Project website hosting

# High Performance Computing

- ▶ Software for genomics, transcriptomics, meta-genomics, meta-transcriptomics, meta-barcoding, phylogenetics, etc...
- ▶ Full prokaryotic/eukaryotic genome assembly
- ▶ Genome assembly of *Tetraselmis striata* unicellular algae, 197G PacBio data, 56 cpus, 7 days, 700GB memory; Sea-cucumber 108 RNASeq samples
- ▶ Phylogenetic Bayesian Markov-Chain Monte Carlo simulations > 3 months common

# Marine Metagenomics

- ▶ ELIXIR-EXCELERATE – marine metagenomics WP6
- ▶ 6 semi-synthetic metagenomes – (20% eukaryotic, 16% random)
- ▶ 175 million reads each with functional and taxonomic information
- ▶ 4 data sets analysed with MG-RAST (3.0) and EMG (4.1)
- ▶ Both correctly identified similar numbers of genera 90%
- ▶ Both had both similar community composition accuracy 50% and numbers of false positives



Gianluca De Moro, CCMAR

# Biome-Shiny



Ricardo Leite, IGC

# Ocean Sampling Day / Genomic Observatories

- ▶ OSD: worldwide seawater sampling campaign for microbiome analysis
- ▶ Promoted by Micro B3/Assemble+/EMBRC
- ▶ Started 2014 – 22 sites in Portugal (sampling uneven across years)
- ▶ Originally 16S/18S rDNA, from 2019 full metagenomes
- ▶ From 2020 - Genomic Observatories, Assemble+/EMBRC



Rodrigo Costa, IST







# Marine Forests

## Marine Forests : Important Habitats Worldwide

Macroalgae, plants and animals like corals and sponges create complex habitats that provide resources, shelter and nursery grounds for many marine organisms.



**Macroalgae**

Forests of macroalgae are mainly formed by large brown algae such as kelp or fuoids, but also by red and green algae (© Douglas Kluge)



**Seagrass meadows**

Seagrasses are plants that evolved and adapted to live in the sea. They have roots, leafs and produce flowers and fruits (© Miro Senfick)



**Animal gardens**

Marine animal forests are mostly formed by invertebrates that live attached to rock, such as coral reefs and gorgonian gardens (© NOAA)

## Help the world monitoring Marine Forests

Citizens are contributing with records of seaweed, seagrass, corals, sponges and other forest-forming species to reach worldwide maps with distribution records.



Ester Serrão, CCMAR  
pause...