

## Test the Virtual Labs - Plankton Genomics

Webinar – 13 May 2022

Sara Pittonet, [Trust-IT](#)



# Blue-Cloud: your Open Science platform for collaborative marine research

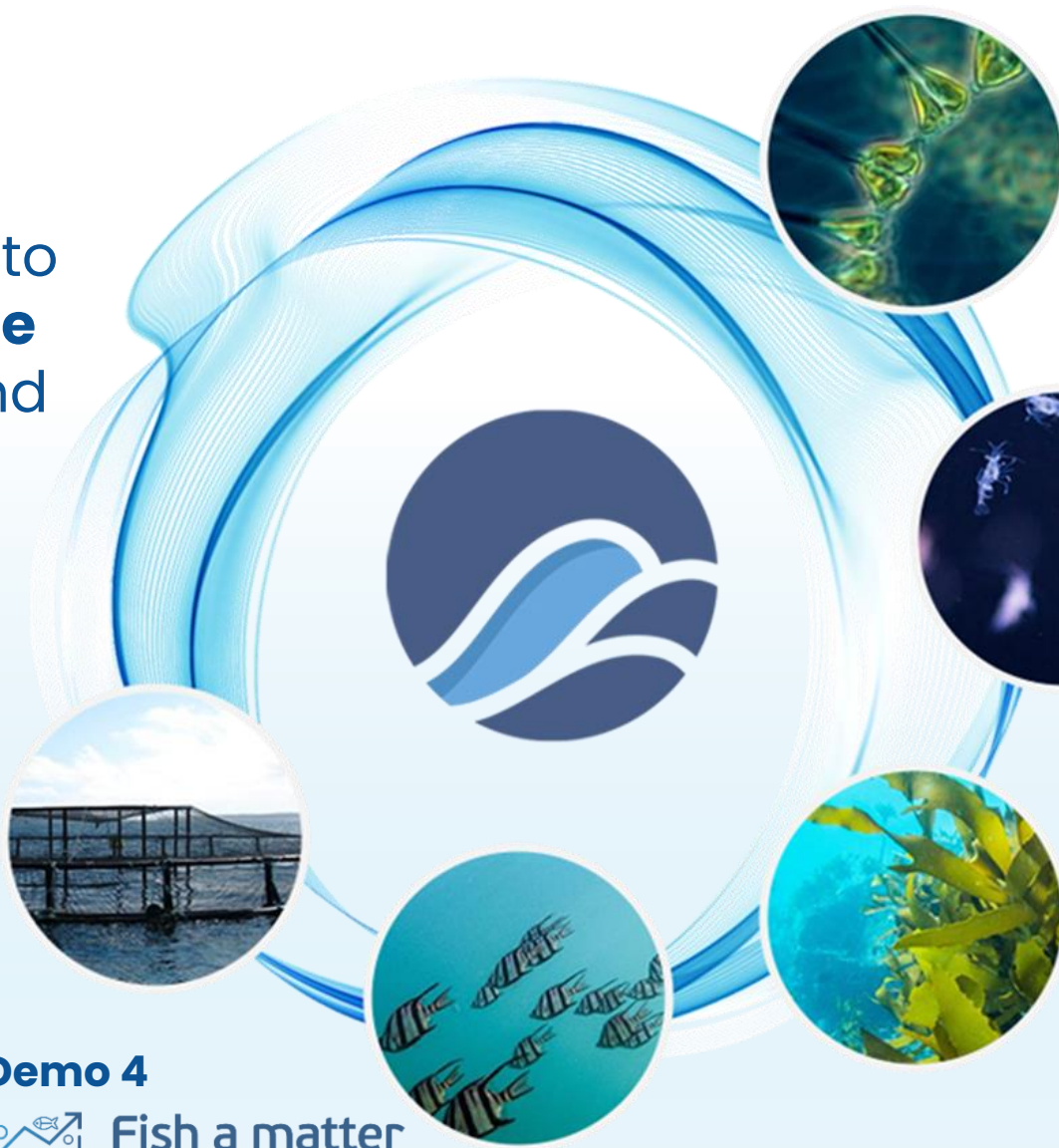
Blue-Cloud aims to promote the **sharing** of *data, processes and research findings* in the marine domain by delivering a **collaborative web-based environment** that enables *open science*, underpinned by simplified access to an **unprecedented wealth of marine data resources** and **interoperable added-value services and products**”

**Funding:** H2020: The ‘Future of Seas and Oceans Flagship Initiative’  
(BG-07-2019-2020) topic: [A] 2019 - Blue Cloud services  
October 2019- September 2022  
20 partners + **13 Blue federated Infrastructures**



Blue-Cloud

**5 demonstrators** to showcase the **Blue Cloud** Services and its potential in promoting the Blue economy



**Demo 1**



Zoo & Phytoplankton  
EOV products

**Demo 2**



Plankton  
Genomics

**Demo 3**



Marine Environmental  
Indicators

**Demo 5**



Aquaculture  
Monitor

**Demo 4**



Fish a matter  
of scales





## Exploring and mapping plankton genomics data with Blue-Cloud

**Authors:**  
 Pavla Debeljak, Alexandre Schickele,  
 Sakina-Dorothee Ayata, Lucie Bittner,  
 Jean-Olivier Irissou, Sorbonne Université,  
 Federico Drago, Trust-IT Services

Blue-Cloud has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017744. Topic: [A] 2019 - Blue Cloud services, Grant Agreement No 101017744.

The dataset includes 4 different size classes (Figure 1) based on the different filter sizes used at on-board filtration: 0.8-5 µm (pico- to nanoplankton), 5-20 µm (nanoplankton), 20-180 µm (microplankton), 180-2000 µm (mesoplankton). The Jupyter Notebook allows for the extraction of Unknowns based on Function or Taxonomy retrieved from Carradec et al. The codes in the notebook find the unknown sequences and calculate the ratio of unknowns to knowns for different size fractions of the Tara Ocean data. Furthermore, giant scaffolds can be excluded and mean sequence length and standard deviation calculated and plotted in R (using the "ggplot" package).

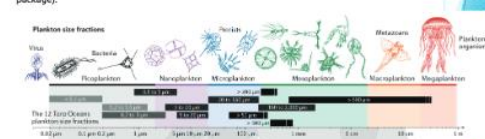


Figure 1. Size Fractions from Pico- to Megaplankton by filter size from the Tara Oceans expeditions. Picture modified from Sunagawa et al. 2020.

The different size classes (Fig. 1) ranging from Nanoplankton to Mesoplankton analysed together or separately can then be plotted (Figure 2) using environmental data in R (leaflet package).



Figure 2. Example of the visualisation of Tara Oceans data and environmental parameters in the leaflet package in R implemented in Notebook 1.1.

### 1.2) Creating protein functional clusters

Notebook 1.2. allows for the creation of protein functional clusters from FASTA files derived from metagenomic and metatranscriptomic sequencing. These clusters contain annotated as well as unknown sequences that can be passed on to the IT service: mapping the geographic distribution of plankton functional gene clusters using habitat prediction models (described in the next section).

The necessary data can be retrieved from <https://www.genoscope.cns.fr/tara> under "Tara Oceans Eukaryotic Genomes ("MAGs") (Delmont et al. 2021 accessible through <https://www.biorxiv.org/content/10.1101/2020.10.15.341214v2>).

# Plankton Genomics Virtual Lab

Led by the [European Bioinformatics Institute](#) and created by the [Faculty of Sciences at Sorbonne University](#), with contribution from [Flanders Marine Institute](#).

Access the VLab



**DECODE THE OCEAN!**

7, 8 and 9 February 2022  
[blue-cloud.org/hackathon](https://blue-cloud.org/hackathon)





# Agenda

- 11:00 **Sara Pittonet Gaiarin**, Trust-IT Services, Blue-Cloud  
Project Coordinator – Blue-Cloud introduction
- 11:05 **Eric Pelletier**, Genoscope, CEA – Genes and functions
- 11:15 **Jean-Olivier Irisson**, Laboratoire d'Océanographie de  
Villefranche, Sorbonne Université – Habitat modeling
- 11:25 **Alexandre Schickele**, Laboratoire d'Océanographie  
de Villefranche, Sorbonne Université – Plankton  
Genomics Virtual Lab
- 11:45 – **Q&A**