Blue-Cloud

Piloting innovative services for Marine Research & the Blue Economy

Demonstrator #2 Plankton Genomics

Blue-Cloud workshop 23rd of March 2021

Blue-Cloud has received funding from the European Union's Horizon Programme call BG-07-2019-2020, topic: [A] 2019 - Blue Cloud services, grant Agreement number 862409.

Plankton Genomics

- Genetic potential and diversity of plankton is unlocked by sequencing projects

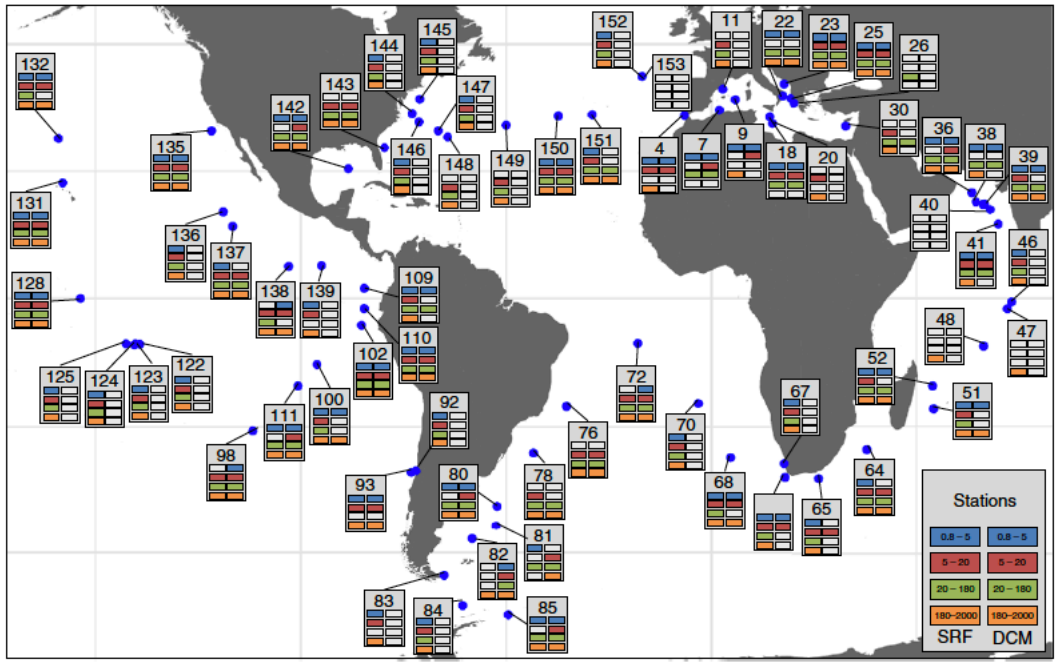
- However we are still only scratching the surface of microbial diversity in the Ocean

© Christian Sardet

Plankton Genomics

- No match in existing databases: **Marine Eukaryotes 50 %** (Carradec, Pelletier et al., 2018)

- Common practice to discard any data that cannot be annotated through public databases

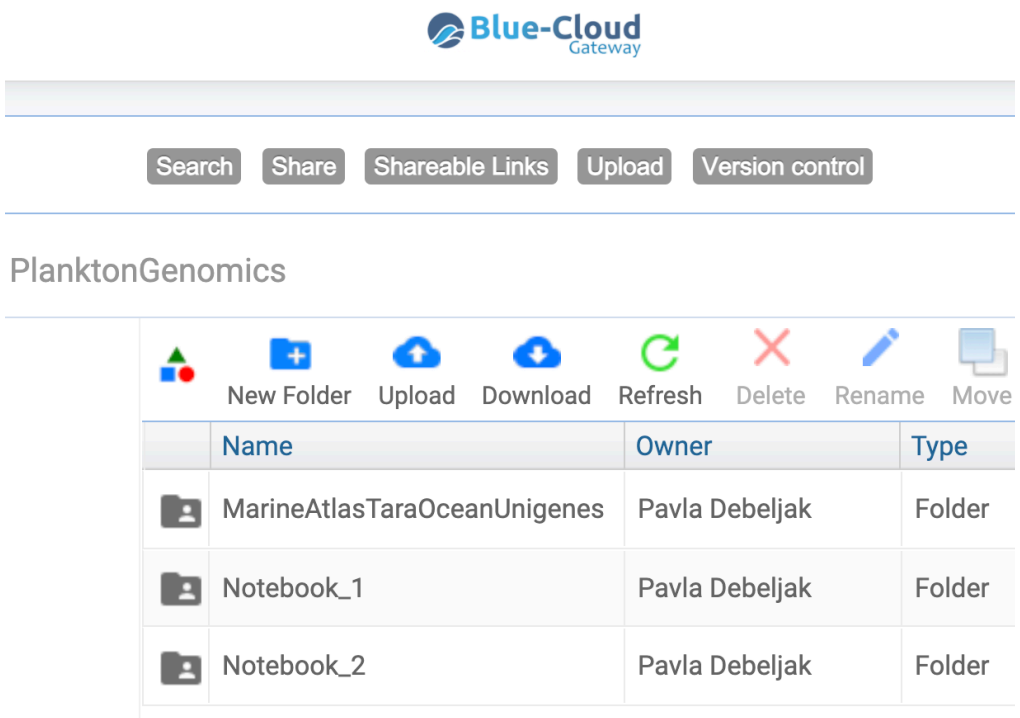


Tara Oceans: 68 different geographic locations across all the major oceanic provinces

Plankton Genomics Demonstrator

Notebooks in the Blue-Cloud VRE

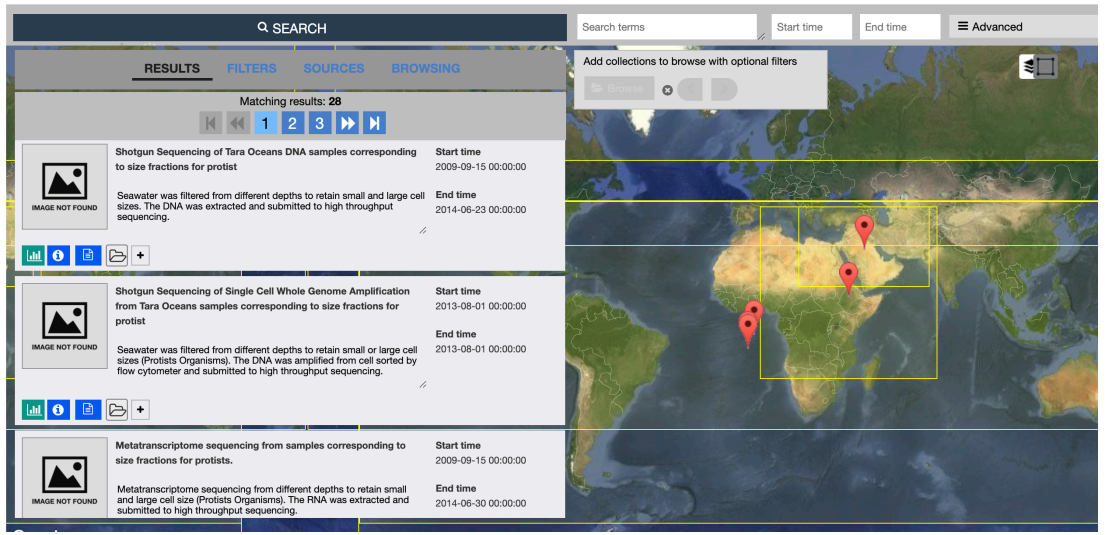
- **Notebook 1:** Discovery of unknown biodiversity from unassigned genetic signals, co-occurence with known biodiversity, and environmental context
- **Notebook 2:** Niche modelling and biogeography of as yet unknown biodiversity



Blue-Cloud
Piloting innovative services for Marine Research & the Blue Economy

Plankton Genomics Datasets

- Exploration of the biggest Eukaryotic plankton database built from **metagenomic** and **metatranscriptomic** sequences
- Combining information on biogeographic distribution and providing data for global ocean modelling



Data available through ENA European Nucleotide Archive elixir

http://gs-service-production.geodab.eu/gs-service/search?view=elixir-ena

Workflow - Notebook 1: Discovery of Unknowns

TARA OCEANS

- Annotated_Taxa
- Annotated_Function
- MetaG Genetic potential
- MetaT Gene expression

NOTEBOOK 1

- **Jupyter** Notebooks
  - KNOWN
  - UNKNOWN
- Product Datasets to be passed on to **R** for UNKNOWN/KNOWN Analysis
- Datasets published in the Blue-Cloud catalogue

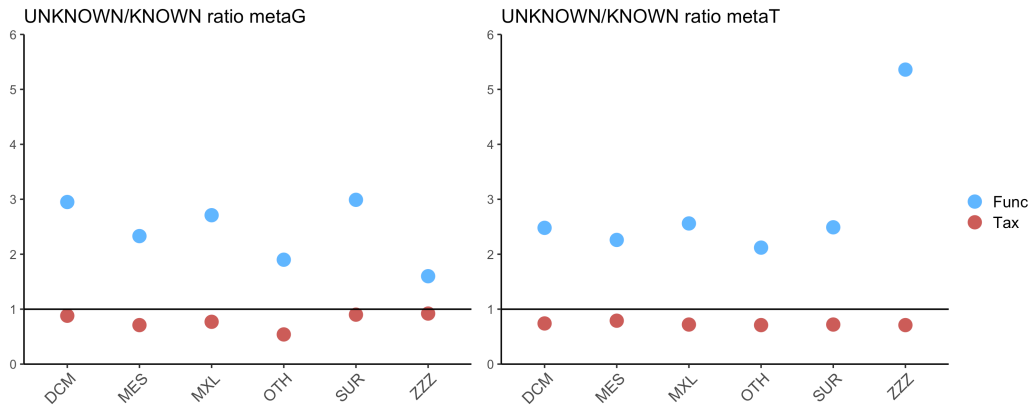
Workflow - Notebook 1: Discovery of Unknowns

TARA OCEANS

- Annotated_Taxa
- Annotated_Function
- MetaG
  Genetic potential
- MetaT
  Gene expression

NOTEBOOK 1

Ratio of UNKNOWN/KNOWN for MetaG and MetaT by depth layer



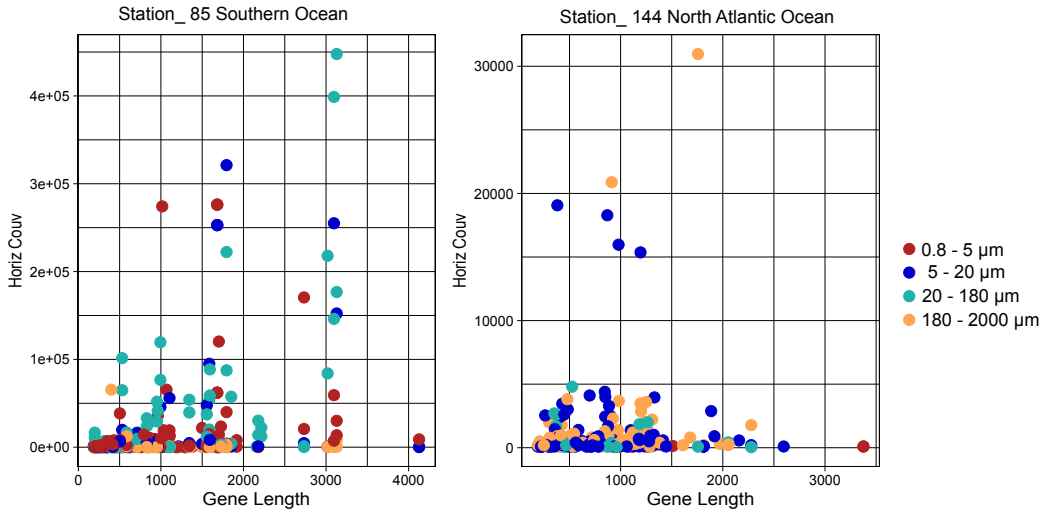
Workflow - Notebook 1: Discovery of Unknowns

TARA OCEANS

- Annotated_Taxa
- Annotated_Function
- MetaG Genetic potential
- MetaT Gene expression

NOTEBOOK 1

Horizontal coverage by sampling station and planktonic size class



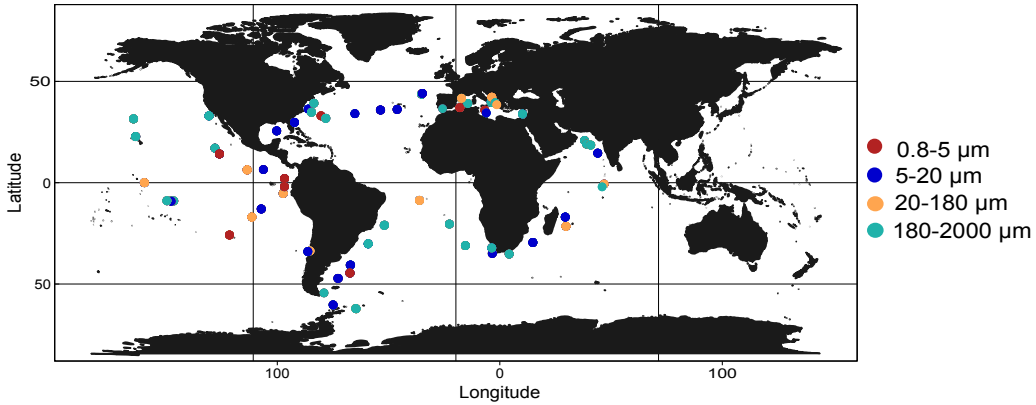
Workflow - Notebook 1: Discovery of Unknowns

TARA OCEANS

- Annotated_Taxa
- Annotated_Function
- MetaG Genetic potential
- MetaT Gene expression

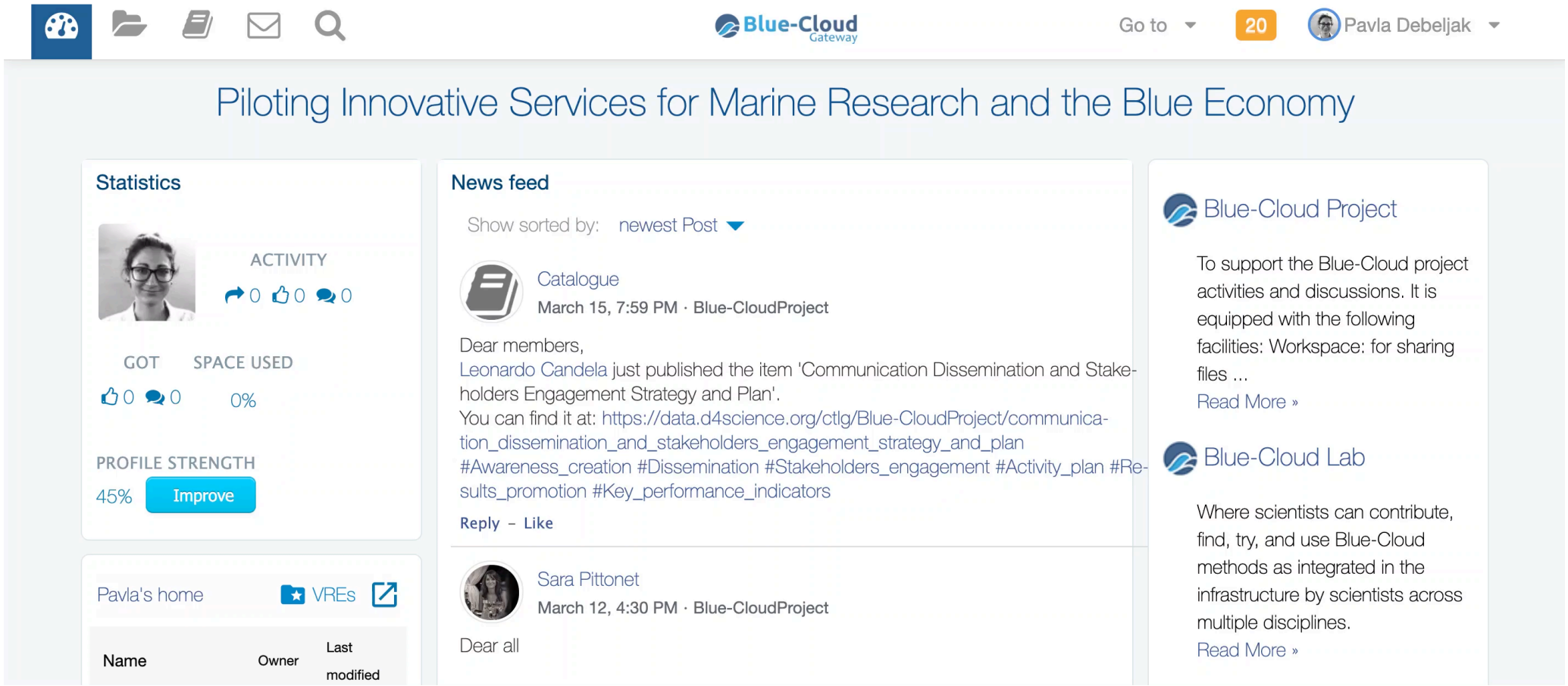
NOTEBOOK 1

Distribution by oceanic region and planktonic size class



Published as catalogue in the VRE to be
Passed on to Notebook 2 for niche modelling

Workflow - Notebook 1: Discovery of Unknowns



Blue-Cloud
Piloting innovative services for Marine Research & the Blue Economy

Thank you for your attention!

Demonstrator #2
Plankton Genomics

Pavla Debeljak, SU
Lucie Bittner, SU
Eric Pelletier, Genoscope
Sakina-Dorotée Ayata, SU
Jean-Olivier Irisson, SU
Stéphane Peasant, EMBL-EBI
Guy Cochrane, EMBL-EBI
Vishnukumar Balavenkataraman Kadhirvelu, EMBL-EBI

SCIENCES SORBONNE UNIVERSITÉ

EMBL-EBI

VLIZ