ESR 8 - EXPLORING THE GLOBAL OCEAN BIOLOGICAL DARK MATTER USING SINGLE EUKARYOTIC CELL GENOMICS & TRANSCRIPTOMICS APPLIED TO WORLD META-OMICS DATASETS

Research project	Microbial eukaryotes (protists) appear to be the dark matter of oceans' Life, and are particularly prominent in planktonic ecosystems where they play the role of plants in terrestrial systems and allow organic matter transfer from the smallest microbes (prokaryotes and viruses) to animals (zooplankton). Over the last 7 years, 2 major circumglobal scientific expeditions, Tara Oceans (2009-2014) and Malaspina (2010-2011), have collected holistic plankton samples for high-throughput sequencing and imaging technologies, from viruses to fish larvae, across the world oceans from surface waters to meso- and bathy-pelagic layers. Hundreds of meta-genomes and meta-transcriptomes have been generated from these projects, in particular across the organismal size-fractions corresponding to eukaryotic cells (from 1 to >1,000 µm). If meta-omics data from prokaryotic or viral size-fractions are reaching saturation and are relatively well annotated (taxonomically and functionally), **the genes and contigs from eukaryotic meta-omics data are far from saturation and unknown for the great majority, pointing out the huge gap in reference gen(ome)s of marine protists and leaving us with a massive amount of orphan genes and senseless environmental data**. The PhD candidate will generate new Single-Amplified Transcriptomes (SATs) from key planktonic protists, as well as use the Single-Amplified Genomes/Transcriptomes (SAG/Ts) generated by other ETN partner teams (ESR1 and ESR7 among others) and within the Tara Oceans consortium, in order to recruit meta-omics reads from the Tara Oceans and Malaspina datasets. This will allow to **(i) annotate and assemble, taxonomically, a significant part of global oceans eukaryotic meta-omics data; (ii) assess the genomic and phylogenomic variability around key eukaryotic plankton taxa; and (iii) understand which genes are co-triggered by specific environmental conditions for key plankton taxa**. In particular, the functional and evolutionary hypotheses generated by ESR 1, ESR2, ESR9 and ESR11 will be tested in the real world across global ecological and geographical scales. A focus will be given on one of the most important but poorly understood ecological function in oceanic plankton: mixotrophy or the capacity that most if not all eukaryotic phytoplankton have to both photosynthesize and feed by endo- or phagocytosis.
Supervisor	Name: Colomban de Vargas Email: vargas@sb-roscoff.fr Website: www.sb-roscoff.fr. www.oceanomics.eu/, www.planktonplanet.org
Host institution	Station Biologique de Roscoff (SBR-CNRS) Roscoff, France

PhD program	ED227, National Museum of Natural History/University Pierre et Marie Curie (http://www.mnhn.fr/fr/enseignement-formation/enseignement-superieur/ecole-doctorale).
Expected results	Taxonomic annotation of massive amount of orphan meta-omics data from global oceans. Assessment of genomic variability over global biogeographic scales for key eukaryotic ecological players (adaptation). Understanding of acclimation process for key ecological players at the genic expression level. Assessment of mixotrophic genomes expression over large geographic and ecological scales.
Planned secondments	CEA – Months 12, 24 and 36, one month each (3 m) – Targeting key SAGs data from Tara Oceans CSIC-ICM – Month 18 (3 m) – Linking functional & phylogenetic diversity in heterotrophic protists VIB – Month 29 (3 m) – Targeting specific genes for metagenomic searches
Required profile	Graduate in Bioinformatics or Cell/Molecular Biology. The candidate must have demonstrable skills with R, Python or Perl (in a Unix /Linux environment). Theoretical and technical skills in molecular biology/ecology will be an advantage.