EXPLORING THE FUNCTIONAL SIDE OF THE OCEAN SAMPLING DAY METAGENOMES

A. Fernàndez-Guerra [1]*, R. Kottmann [2], F. Malfatti [3], M. Bicak [4] and F. Glöckner [5]
[1] Jacobs University Bremen, Max Planck Institute Bremen, University of Oxford - afernand@mpi-bremen.de
[2] Max Planck Institute Bremen
[3] OGS, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale
[4] University of Oxford
[5] Jacobs University Bremen, Max Planck Institute Bremen

Abstract
In June 2014 the first Ocean Sampling Day took place with the objective to simultaneously sample the world's coastal oceans and provide insights into fundamental patterns in microbial diversity and function. In combination with a rich set of environmental and oceanographic measurements, 150 metagenomes were sequenced. OSD added more than three million genes to the Ocean-Microbial Reference Gene Catalog including data from mega-surveys like TARA Ocean expedition and Global Ocean Survey. The ongoing analysis of these metagenomes increases our knowledge of the Ocean Microbiome, and allows identifying hotspots of novelty in terms of function. Furthermore, with OSD we can now investigate the impact of human activities on oceans coastal areas where an intimate interaction between dense human populations and the marine world is ongoing.

Keywords: Coastal systems, Biodiversity, Biotechnologies, Mediterranean Sea

The Ocean Sampling Day (OSD) [1] is a simultaneous, collaborative, standardized, global mega-sequencing campaign to analyze marine microbial community composition and functional traits. 150 metagenomes were sequenced from the first OSD in June 2014 including a rich set of environmental and oceanographic measurements. Unlike other ocean mega-surveys such as Global Ocean Sampling (GOS) [2] or the TARA Oceans expedition [3] that mostly sampled open ocean waters most of the OSD samples are from coastal sampling sites, a previously not well studied area. Additionally, with the repetition of OSD in 2015 and beyond, these cumulative samples, related in time, space and environmental parameters, will provide insights into fundamental rules describing microbial diversity and function and contribute to the blue economy through the identification of novel ocean-derived biotechnologies.

The Ocean Sampling Day and the Ocean Microbiome
With 40,154,822 genes the Ocean Microbial-Gene Reference Catalog (OM-RGC) -- compiled for the TARA Oceans expedition data analysis -- is the largest marine gene dataset to date. It includes genes from different sources like the GOS, the Pacific Ocean Virome study [4] and from ocean microbial reference genomes [3] among others. Despite its extensive oceanic coverage, coastal waters are still underrepresented in OM-RGC. OSD is complementing this by providing metagenomics data from 124 coastal sampling sites (less than 10 km to the shore) (Figure 1).

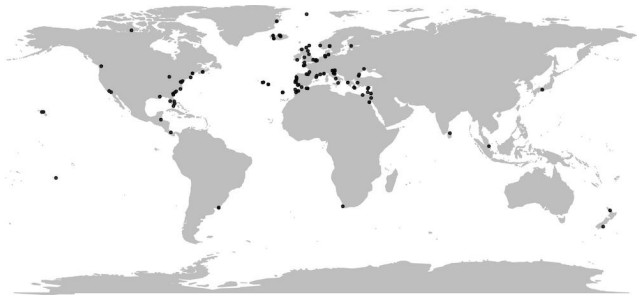

Fig. 1. Figure 1. Geographical distribution of the Ocean Sampling Day metagenomes

Coastal environments are a thin boundary between mainland and the ocean waters; they are highly dynamic environments receiving a high load from terrestrial environments, natural or human-based. This geographical heterogeneity is indicated by the low level of similarity in the raw genetic distances between the different OSD metagenomes. The dynamic and heterogeneous nature of the coastal environments is again underlined by exploring the functions based on the orthologous groups of genes (OG) classification defined by eggNOG [5]. A total of 60,644 OGs were identified on the 150 OSD metagenomes, but only 189 OGs were shared between all OSD metagenomes The low amount of OGs conforming the core of the OSD reinforce the idea of the heterogeneity in coastal environments. Analyzing the functional annotations for the exclusive OSD genes, the orthologous groups showing higher abundances are those related to the categories: "Amino acid transport and metabolism", "Energy production and conversion", "Replication, recombination and repair" and "Translation, ribosomal structure and biogenesis". These functional categories suggest that the organisms habiting in the OSD coastal sites are specialized and adapted to live in a changing environment, although we cannot assess the level of activity without metatranscriptomic data [6]. With one single sampling point in time we show the existence of a high geographic heterogeneity at the geographic level, and with the future OSDs we will be able to study the dynamics of the coastal environments in a global scale, allowing us to gain more insights about their resilience and be able to observe predictable patterns [7].

References
1 - Kopf, Anna, Mesude Bicak, Renzo Kottmann, Julia Schnetzer, Ivaylo Kostadinov, Katja Lehmann, Antonio Fernandez-Guerra, et al. 2015. The ocean sampling day consortium. *Gigascience* 4:27.
2 - Rusch, Douglas B, Aaron L Halpern, Granger Sutton, Karla B Heidelberg, Shannon Williamson, Shibu Yooseph, Dongying Wu, et al. 2007. The sorcerer II global ocean sampling expedition: Northwest atlantic through eastern tropical pacific. PLoS Biol 5 (3): e7.
3 - Sunagawa, Shinichi, Luis Pedro Coelho, Samuel Chaffron, et al. 2015. Structure and function of the global ocean microbiome. Science 348 (6237): 1261359.
4 - Hurwitz, Bonnie L, Li Deng, Bonnie T Poulos, and Matthew B Sullivan. 2013. Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics. Environ Microbiol 15 (5): 1428-40.
5 - Powell, Sean, Kristoffer Forslund, Damian Szklarczyk, Kalliopi Trachana, Alexander Roth, Jaime Huerta-Cepas, Toni Gabaldón, et al. 2014. EggNOG v4.0: Nested orthology inference across 3686 organisms. Nucleic Acids Res 42 (Database issue): D231-9
6 - Gifford, Scott M, Shalabh Sharma, and Mary Ann Moran. 2014. Linking activity and function to ecosystem dynamics in a coastal bacterioplankton community. Front Microbiol 5:185
7 - Fuhrman, J. A., Cram, J. A., & Needham, D. M. (2015). Marine microbial community dynamics and their ecological interpretation. Nature Reviews. Microbiology, 13(3), 133-46