### A Holistic Approach to Marine Eco-Systems Biology

The MIT Faculty has made this article openly available. *Please share* how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1371/journal.pbio.1001177">http://dx.doi.org/10.1371/journal.pbio.1001177</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Public Library of Science</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Wed Mar 16 21:44:03 EDT 2016</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/69086">http://hdl.handle.net/1721.1/69086</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td><a href="http://creativecommons.org/licenses/by/2.5/">http://creativecommons.org/licenses/by/2.5/</a></td>
</tr>
</tbody>
</table>
A Holistic Approach to Marine Eco-Systems Biology

Eric Karsenti1*, Silvia G. Acinas2, Peer Bork1, Chris Bowler3,4, Colomban De Vargas5,6, Jeroen Raes7,8, Matthew Sullivan9, Detlev Arendt1, Francesca Benzoni10, Jean-Michel Claverie11,12, Mick Follows12, Gaby Gorsky3,13, Pascal Hingamp3,14, Daniele Iudicone14, Olivier Jaillon15, Stefanie Kandels-Lewis1, Uros Krzic1, Fabrice Not3,5,6, Hiroyuki Ogata3,11, Stéphane Pesant16,17, Emmanuel Georges Reynaud18, Christian Sardet16,19, Michael E. Sieracki20, Sabrina Speich21, Didier Velayoudon22, Jean Weissenbach15, Patrick Wincker15, the Tara Oceans Consortium*

Introduction

With biology becoming quantitative, systems-level studies can now be performed at spatial scales ranging from molecules to ecosystems. Biological data generated consistently across scales can be integrated with physico-chemical contextual data for a truly holistic approach, with a profound impact on our understanding of life [1–5]. Marine ecosystems are crucial in the regulation of Earth’s biogeochemical cycles and climate [6,7]. Yet their organization, evolution, and dynamics remain poorly understood [8,9]. The Tara Oceans project was launched in September 2009 for a 3-year study of the global ocean ecosystem aboard the ship Tara. A unique sampling programme encompassing optical and genomic methods to describe viruses, bacteria, archaea, protists, and metazoans in their physico-chemical environment has been implemented. Starting as a grassroots initiative of a few scientists, the project has grown into a global consortium of over 100 specialists from diverse disciplines, including oceanography, microbial ecology, genomics, molecular, cellular, and systems biology, taxonomy, bioinformatics, data management, and ecosystem modeling. This multidisciplinary community aims to generate systematic, open access datasets usable for probing the morphological and molecular makeup, diversity, evolution, ecology, and global impacts of plankton on the Earth system.

Viruses, bacteria, archaea, protists, and planktonic metazoans form the bulk of biomass throughout the oceans and drive the global biogeochemical cycles that regulate the Earth system [6,9,10]. For instance, marine microbes produce nearly as much oxygen through primary production as land plants [11]. This system is driven by a complex ecological network of autotrophic, heterotrophic, and mixotrophic organisms, where trophodynamics and biogeochemical interdependencies are determining factors for primary production rates in marine systems. In addition, ocean viruses modulate primary production by inducing organism mortality and by encoding core photosynthesis genes that are expressed during infection [12–14]. Therefore, only an ecosystem-wide approach, from viruses to metazoans, will enable us to start disentangling the functioning of the Earth system. This approach ranges from mapping organismal diversity across scales spanning five orders of magnitude to developing empirical datasets that inform conceptual models about the complex interplay between organisms driving fluxes of energy, biogeochemical, and molecular “currencies” in ocean ecosystems [15].

A global-scale study of morphological, genetic, and functional biodiversity of plankton organisms in relation to the changing physico-chemical parameters of the oceans [8,16–18] is now critical to understanding and managing our fragile oceans. Specifically, such a dataset will improve our understanding of the principles governing marine ecosystems and the evolution of life in the ocean, thus enhancing our capacity of assessing ecosystem services and enabling a better prediction of fish stock distribution and impacts of...
global climate variations [19]. Planktonic organisms are also an enormous but largely untapped source [8,20] of bio-active compounds for the pharmaceutical, food, and cosmetics industries, as well as metabolic pathways that may provision our future energy needs [21]. In this context, the Tara Oceans consortium was founded, which embarked on a 3-year research cruise across the world’s oceans.

Tara Oceans is not the first group to explore global ocean biodiversity. For example, previous global initiatives include satellite-based chlorophyll measurements, the Census of Marine Life, long-term observation sites, and arrays of remote sensors on floats that provide physical, chemical, and biological data [15]. Other global genomics studies have been launched, e.g., Global Ocean Sampling (GOS) expedition [22] and the Earth Microbiome [23] project, as well as integrative projects focusing on specific biomes (e.g., Malaspina, http://www.expedicionmalaspina.es/). However, Tara Oceans takes such investigations one step further by integrating the genetic, morphological, and functional diversity in its environmental context at global ocean scale and at multiple depths (Figure 1), from viruses to fish larvae. While such a “study it all” approach is not novel (e.g., NSF Long Term Ecological Research sites), it has remained science fiction until technology and informatics became enabling. Now, high throughput sequencing, quantitative imaging methods, dedicated bio-informatics and modeling tools are poised to assess the complexity of the global ocean systems.

To achieve such integration, Tara Oceans is driven by researchers with expertise in biological and physical oceanography, ecology, microbiology, systematics, molecular, cellular and systems biology, bioinformatics, data management, and modeling.

Figure 1. The Tara Oceans cruise. (A) Route of the Tara Oceans expedition. Sampling stations from surface to 1,000 m are carried out between ports of call guided by satellite data about the basin to sub-mesoscale structures. (B) Tara Oceans sampling sites in the Mozambique Channel and South Atlantic. The images show near real time sea surface height (SSH) from satellite. Each sampling station is indicated by a black diamond; those that are currently being targeted for priority studies are encircled by a white halo. The altimetry data is from September 16, 2011, when Tara was sampling inside an Agulhas ring. Several other rings are also apparent in the figure, as is the Malvinas Current off the Argentinean coast that injects cold Antarctic water into the resident waters.
doi:10.1371/journal.pbio.1001177.g001
Pragmatically, to accomplish such an ambitious goal, Tara Oceans consortium scientists have necessarily been intimately involved in every aspect of the expedition. This includes planning, preparation, and running of the on-board sampling protocols, as well as the development of sample analysis and bioinformatics pipelines, data management, and modeling projects. This involvement ensures a coherent worldwide data collection and analysis strategy, which is reinforced through regular workshops. The consortium has an open access policy concerning the data that will be made available to the scientific community as soon as possible after validation. Finally, a significant outreach effort is made to show life on board of Tara as well as translate results to the broader public (see http://oceans.taraexpeditions.org/ and http://www.planktonchronicles.org).

The Expedition and Sampling Strategy

To collect and fractionate plankton communities on the basis of organism size, spanning orders of magnitude (Figure 2A), the sampling combines traditional (Niskin bottles and plankton net tows) and novel methods (e.g., a gravity-driven plankton sieve and chemistry-based concentration of viruses [24]) that feed into analytical pipelines appropriate for each size class. These analyses range from immediate visualization and quantification on-board Tara with a diversity of imaging tools, to collection and preservation of samples for genomic and morphological analysis back on land.

Given the global nature of the expedition and spatial heterogeneity in the oceans, a concern is the “snapshot” sampling strategy employed here at a single time point and at relatively sparse stations relative to the global ocean, which represents the inherent challenge of global ocean studies. In Tara Oceans we leverage near real-time remote sensing and other data to locate oceanographically interesting features (e.g., eddies, fronts, upwellings) and strengthen ecosystem comparisons (Figures 1B and 2A). The vast ocean basins are relatively homogeneous on seasonal to decadal time scales, whereas smaller-scale systems are more dynamic. For example, the Agulhas leakage system (Figure 1B) transports water from the Indian to the Atlantic Ocean and across to South America leaving heterogeneous ocean features behind that persist for weeks to months [25]. The Agulhas system represents an ideal case for applying the Tara Oceans near real-time sampling strategy and downstream analysis pipeline to deeply characterize the biology of these ecosystems. Through systematic study of such heterogenous systems coupled to broader “normal” ocean sampling, we hope to unveil the rules that govern the structure and dynamics of ocean ecosystems and to extrapolate such local observations to develop basin-scale process models as predictive tools [26]. Undoubtedly, such measurements and predictions will provide a starting point for hypothesis testing by more focused, follow-up campaigns.

The Tara Oceans Integrated Pipeline: Towards Eco-Systems Biology

The scientific programme of Tara Oceans requires an integrated pipeline that combines semi- or fully automated data acquisition methods, new bioinformatics tools, and standardized data organisation (Figure 2C). The high throughput imaging platform includes instruments tuned to organisms of particular size classes. They include (i) on-board and on-land flow cytometers to monitor virus particles, bacteria, and small protists, (ii) on-board digital and confocal microscopy for detailed 2D/3D imaging of cells within the 5–20-μm range, (iii) on-board and on-land FlowCams and ZooScans for quantitative recognition of organisms ranging from 20 μm to a few cm, light sheet and confocal microscopes for 3D imaging, and (iv) on-land electron microscopes for detailed ultrastructural analyses of small protists and viruses. In parallel, we use high throughput sequencing methods to obtain both deep phylogenetic rDNA/ rRNA tag data and metagenomic and metatranscriptomic functional profiles from size fractions covering the entire plankton community from viruses to fish larvae (Figure 2B and 2C). To bring all these data together for analyses, we leverage recently developed, dedicated computational approaches [27]. In addition, Tara Oceans is archiving meteorological, oceanographic, biogeochemical, and plankton morphology data in the PANGAEA database (http://www.pangaea.de/), linking to larger European and international data infrastructures. Thus, Tara Oceans can visualize, quantify, and genetically characterize ocean biodiversity within entire plankton ecosystems, as well as find patterns across unprecedentedly comprehensive data types.

With such a powerful dataset and toolkit, we anticipate testing the predictions of biodiversity hotspots from stochastic modelling [28–31], as well as mapping functional gene ecology and activities throughout the world’s oceans. This process has been initiated, predominantly for 0.1–0.8-μm-sized surface ocean microbes, using data provided by the GOS expedition [22,27,32,33], but will be dramatically extended here by sampling throughout the water column, across organismal size scales and beyond metagenomic sequencing, tightly coupled to global ocean-modelling efforts.

These data will transform our ability to link species diversity and metabolic potential/activity to environmental conditions and ecosystem outputs and promises to lead to the discovery of emergent ecological principles (Figure 2C). They will allow a deeper understanding of biodiversity gradients within and among systems and contrasting environments. We also anticipate establishing rules governing the self-organization of organism networks (e.g., to address which types of viruses, bacteria, protists, and zooplankton live together in a given environment) [16,27], and develop predictions about how these rules and communities will be affected by a changing environment.

In summary, the Tara Oceans project leverages powerful new technologies and analytical tools to develop the first planetary-scale data collection effort that links biogeography with ecology, genetics, and morphology. Guided by the cross-disciplinary philosophy the pay-offs can be immense, considering the massive number of samples and data that have been collected, archived, and interconnected for scientific study, only half-way through the expedition. A lesson from this project is that, when it comes to addressing broad and complex issues of general interest to mankind, competition between scientists may not be the best model. The Tara Oceans project is a pioneering enterprise towards a truly worldwide, systems-level characterization of the largest and most fundamental ecosystem on our planet.

Acknowledgments

We are keen to thank the commitment of the following people and sponsors who made this singular expedition possible: CNRS, EMBL, Genoscope/CÉA, VIB, Stazione Zoologica Anton Dohrn, UNIMIB, ANR (projects PO- SEIDON, BIOMARKS, PROMETHEUS, and TARA-GIRUS), FWO, BIO5, Biosphere 2, agnes b., the Veolia Environment Foundation, Region Bretagne, World Courier, Illu- mina, Cap L’Orient, the EDF Foundation EDF Diversiterie, FRB, the Prince Albert II de Monaco Foundation, Etiene Bourgeois, the Tara schooner and its captain and crew. Tara Oceans would not exist without the continuous
Figure 2. The Tara Oceans model. (A) Methods for sampling organisms by size classes and abundance. The blue background indicates the filtered volume required to obtain sufficient organism numbers for analysis. Actual volumes from which organisms are sampled are always recorded. (B) Methods for analyzing samples. Data on the right are from Tara Oceans sampling stations. (C) Models that will benefit from Tara Oceans data. High throughput genome sequencing and quantitative image analysis provide evolution, metabolic, and interaction data to build community metabolome maps, taxa/gene networks, and spatial ecosystem models.

doi:10.1371/journal.pbio.1001177.g002
support of the participating 23 institutes (see http://oceans.taraexpeditions.org). Special thanks also goes to Fabrice Not and Noan Le Bescot for creating the figures of this manuscript. This is contribution number 0001 of the Tara Oceans Expedition 2009–2012.

In addition to the main authors, the Tara Oceans consortium includes the following individuals:


23 Centre Européen de Recherche et d’Enseignement des Géosciences de l’Environnement (CEREGE), Marseille, France,
24 The University of Maine, Orono, Maine, United States of America,
25 Natural History Museum, London, United Kingdom,
26 Biologie de Developpement (BioDev), Villefranche-sur-Mer, France,
27 Observatoire Oceanologique of Banyuls sur mer (OBB), Banyuls-sur-Mer, France,
28 Centre de Formation et de Recherche sur les Environnements Mediterraneens (CERFEM), Université de Perpignan, Perpignan, France,
29 Laboratoire d’Oceanographie et du Climat (LOCLEAN), Paris, France,
30 Station Biologique of Roscoff (SBR), Roscoff, France,
31 Université of Provence, Marseille, France,
32 PROCure of Transferts et d’Echanges dans l’Environnement (PROTECE), Université du Sud Toulon-Var (USTV), Toulon, France,
33 University of California Santa Cruz, Santa Cruz, California, United States of America,
34 Laboratoire d’Oceanographie Dynamique et de Climatologie (LODVC), Paris, France,
35 Institut Méditerranéen d’Ecologie et de Paléocologie (IMEP), Marseille, France,

References