

A Holistic Approach to Marine Ecosystems Biology

Eric Karsenti^{1*}, Silvia G. Acinas², Peer Bork¹, Chris Bowler^{3,4}, Colomban De Vargas^{3,5,6}, Jeroen Raes^{7,8}, Matthew Sullivan⁹, Detlev Arendt¹, Francesca Benzoni¹⁰, Jean-Michel Claverie^{3,11}, Mick Follows¹², Gaby Gorsky^{3,6,13}, Pascal Hingamp^{3,11}, Daniele Iudicone¹⁴, Olivier Jaillon¹⁵, Stefanie Kandels-Lewis¹, Uros Krzic¹, Fabrice Not^{3,5,6}, Hiroyuki Ogata^{3,11}, Stéphane Pesant^{16,17}, Emmanuel Georges Reynaud¹⁸, Christian Sardet^{3,6,19}, Michael E. Sieracki²⁰, Sabrina Speich²¹, Didier Velayoudon²², Jean Weissenbach¹⁵, Patrick Wincker¹⁵, the Tara Oceans Consortium[¶]

1 European Molecular Biology Laboratory (EMBL), Heidelberg, Germany, **2** Institut de Ciències del Mar (ICM), Consejo Superior de Investigaciones Científicas (CSIC), Barcelona, Spain, **3** CNRS, Paris, France, **4** Institut de Biologie de l'École Normale Supérieure (IBENS), Paris, France, **5** Station Biologique de Roscoff (SBR), Evolution du Plancton et Paleo Oceans (EPPO), Roscoff, France, **6** Université Pierre et Marie Curie (UPMC), Paris, France, **7** VIB, Brussels, Belgium, **8** Vrije Universiteit Brussel, Brussels, Belgium, **9** University of Arizona, Tucson, Arizona, United States of America, **10** University of Milano-Bicocca, Milan, Italy, **11** Aix-Marseille Université, Marseilles, France, **12** MIT, Cambridge, Massachusetts, United States of America, **13** Laboratoire d'Océanographie de Villefranche (LOV), Villefranche-sur-Mer, France, **14** Stazione Zoologica Anton Dohrn, Naples, Italy, **15** Commissariat à l'énergie atomique et aux énergies alternatives (CEA), Genoscope, Evry, France, **16** Universitaet Bremen, Bremen, Germany, **17** MARUM Center for Marine Environmental Sciences, Bremen, Germany, **18** University College Dublin, Dublin, Ireland, **19** BioDev, Villefranche-sur-Mer, France, **20** Bigelow Laboratory for Ocean Sciences, West Boothbay Harbor, Maine, United States of America, **21** Laboratoire de Physique des Océans (LPO), CNRS/Institut français de recherche pour l'exploitation de la mer (IFREMER)/Institut de Recherche pour le Développement (IRD)/Université de Bretagne Occidentale (UBO), Brest, France, **22** DVIPC, Paris, France

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

Citation: Karsenti E, Acinas SG, Bork P, Bowler C, De Vargas C, et al. (2011) A Holistic Approach to Marine Ecosystems Biology. *PLoS Biol* 9(10): e1001177. doi:10.1371/journal.pbio.1001177

Published: October 18, 2011

Copyright: © 2011 Karsenti et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The Tara Oceans expedition is generously funded by CNRS, EMBL, Genoscope/CEA, VIB, Stazione Zoologica Anton Dohrn, UNIMIB, ANR, FWO, BIOS, Biosphere 2, agnès b., the Veolia Environment Foundation, Region Bretagne, World Courier, Illumina, Cap L'Orient, the EDF Foundation EDF Diversiterre, FRB, the Prince Albert II de Monaco Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: karsenti@embl.de

¶ Membership of the Tara Oceans Consortium is provided in the Acknowledgments.

The Community Page is a forum for organizations and societies to highlight their efforts to enhance the dissemination and value of scientific knowledge.

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.

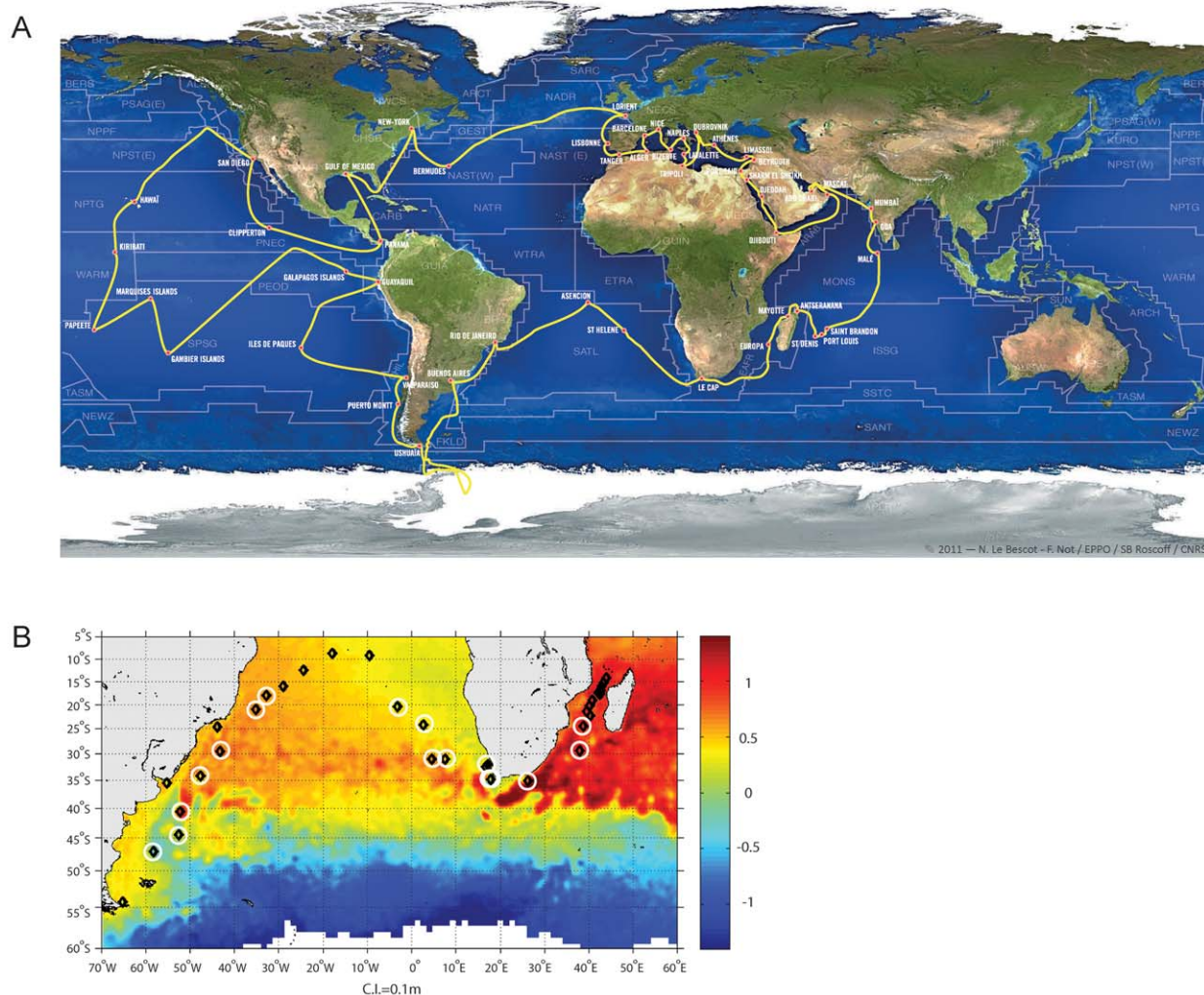


Figure 1. The *Tara Oceans* cruise. (A) Route of the *Tara Oceans* expedition. Sampling stations from surface to 1,000 m are carried out between ports of call guided by satellite data about the basin to sub-mesoscale structures. (B) *Tara Oceans* sampling sites in the Mozambique Channel and South Atlantic. 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 water. 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 five 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 heterogeneous 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 Ecosystems 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-land 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 PAN-GAEA 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 pro-

cess 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.

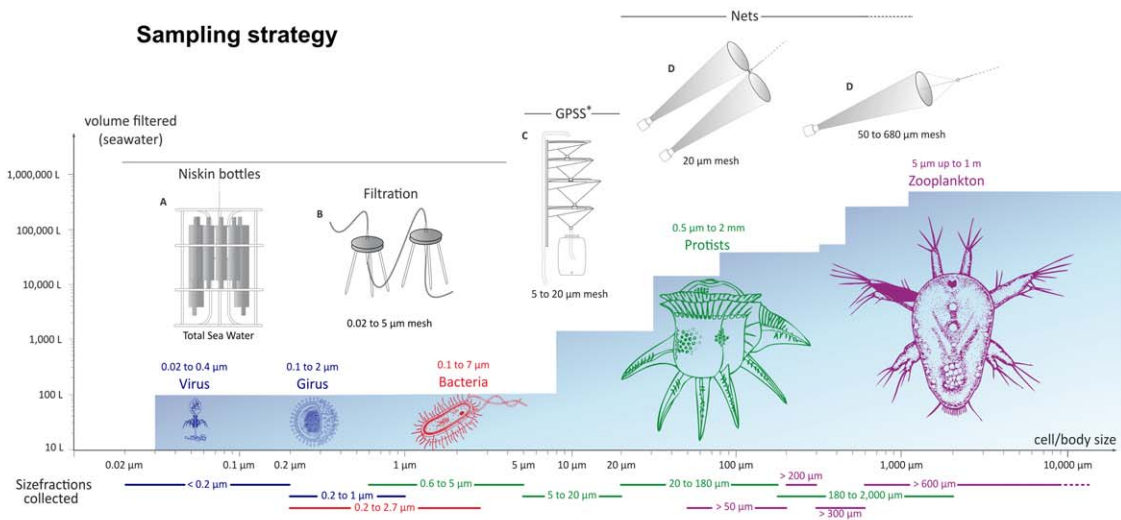
These data will transform our ability to link species diversity and metabolic potential/activity to environmental conditions and ecosystem outputs and promises to lead to the discovery of emergent ecological principles (Figure 2C). They will allow a deeper understanding of biodiversity gradients within 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.

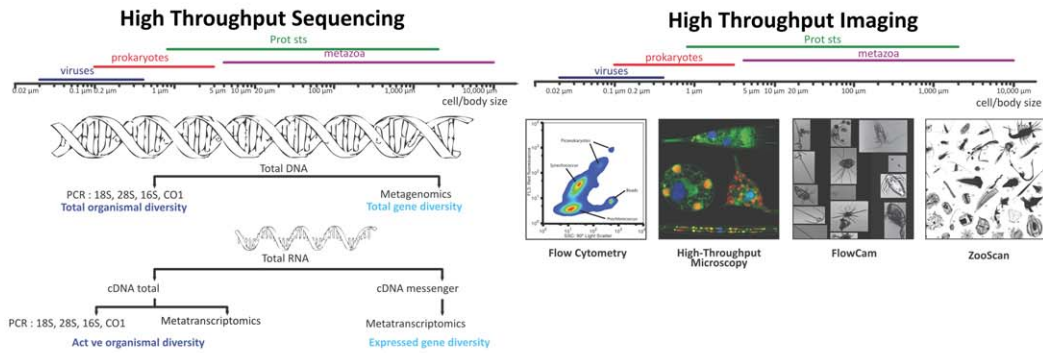
Acknowledgments

We are keen to thank the commitment of the following people and sponsors who made this singular expedition possible: CNRS, EMBL, Genoscope/CEA, VIB, Stazione Zoologica Anton Dohrn, UNIMIB, ANR (projects POSEIDON, BIOMARKS, PROMETHEUS, and TARA-GIRUS), FWO, BIO5, Biosphere 2, agnès b., the Veolia Environment Foundation, Region Bretagne, World Courier, Illumina, Cap L’Orient, the EDF Foundation EDF Diversiterre, FRB, the Prince Albert II de Monaco Foundation, Etienne Bourgois, the *Tara* schooner and its captain and crew. *Tara Oceans* would not exist without the continuous support of the participating 23 institutes (see <http://oceans.taraexpeditions.org>). Special thanks also goes to Fabrice Not and Noan Le

A



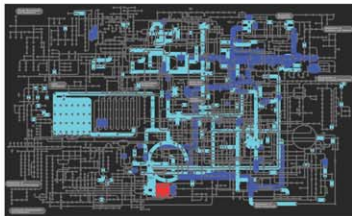
B



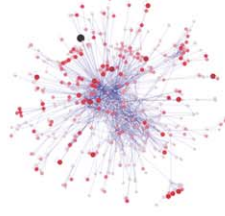
C

Eco- Morpho- Genetic modeling

Community metabolome mapping



Taxa / Gene network analysis



Ecosystem stochastic modeling

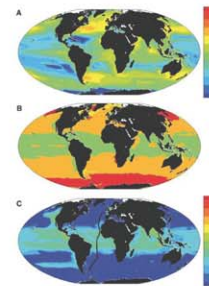


Figure 2. The Tara Oceans model. (A) Methods for sampling organisms by size classes and abundance. The blue background indicates the filtered volume required to obtain sufficient organism numbers for analysis. Actual volumes from which organisms are sampled are always recorded. (B) Methods for analyzing 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

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:

Abergel C^{3,11}, Arslan D^{3,11}, Audic S^{3,5,6}, Aury JM¹⁵, Babic N^{3,5,6}, Beaufort L^{3,11,23}, Bittner L^{3,5,6}, Boss E²⁴, Boutte C^{3,5,6}, Brum J⁹, Carmichael M^{3,5,6}, Casotti R¹⁴, Chambouvet A²⁵, Chang P^{3,6,13,26}, Chica C¹⁵, Clerissi C^{3,6,27}, Colin S^{3,5,6}, Cornejo-Castillo FM², Da Silva C¹⁵, De Monte S^{3,4}, Decelle J^{3,5,6}, Desdevises Y^{3,6,27}, Dimier C^{3,5,6}, Dolan J^{3,6,13}, Duhaime M⁹, Durrieu de Madron X^{3,28}, d'Ortenzio F^{3,6,13}, d'Ovidio F^{6,29}, Ferrera I², Garczarek L^{3,6,30}, Garet-Delmas MJ^{3,5,6}, Gasmi S³¹, Gasol JM², Grimsley N^{3,6,27}, Heilig R¹⁵, Ignacio-Espinoza J⁹, Jamet JL³², Karp-Boss L²⁴, Katinka M¹⁵, Khalili H¹⁵, Kolber Z³³, Le Bescot N^{3,5,6}, Le Goff H^{3,6,34}, Lima-Mendez G^{7,8}, Mahé F^{3,5,6}, Mazzocchi MG¹⁴, Montresor

M¹⁴, Morin P^{3,6,30}, Noel B¹⁵, Pedrós-Alió C², Pelletier E¹⁵, Perez Y^{31,35}, Picheral M^{3,6,13}, Piganeau G^{3,6,27}, Poirot O^{3,11}, Poulain J¹⁵, Poulton N²⁰, Prejger F^{3,6,13}, Prihoda J^{3,4}, Probert I^{3,5,6}, Rampal J³¹, Reverdin G^{3,6,34}, Romac S^{3,5,6}, Romagnan JB^{3,6,13,26}, Roullier F^{3,6,13}, Rouviere C^{3,6,13,26}, Samson G¹⁵, Santini S^{3,11}, Sarmento H², Sciandra A^{3,6,13}, Solonenko S⁹, Stemann L^{3,6,13}, Subirana L^{3,6,27}, Sunagawa S¹, Tanaka A^{3,4}, Testor P^{6,29}, Thompson A¹², Tichanné-Seltzer V^{3,11}, Tirichine L^{3,4,27}, Toulza E^{3,6,27}, Tozzi S³³, Veluchamy A^{3,4}, Zingone A¹⁴

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 Développement (BioDev), Villefranche-sur-Mer, France,

27 Observatoire Océanologique de Banyuls sur mer (OOB) Banyuls-sur-Mer, France,

28 Centre de Formation et de Recherche sur les Environnements Méditerranéens (CE-FREM), Université de Perpignan, Perpignan, France,

29 Laboratoire d'Océanographie et du Climat (LOCEAN), Paris, France,

30 Station Biologique de Roscoff (SBR), Roscoff, France,

31 Université de Provence, Marseille, France,

32 PROcessus de Transferts et d'Echanges dans l'Environnement (PROTEE), Université du Sud Toulon-Var (USTV), Toulon, France,

33 University of California Santa Cruz, Santa Cruz, California, United States of America,

34 Laboratoire d'Océanographie Dynamique et de Climatologie (LODYC), Paris, France,

35 Institut Méditerranéen d'Ecologie et de Paléocologie (IMEP), Marseille, France

References

- Karsenti E (2008) Self-organization in cell biology: a brief history. *Nat Rev Mol Cell Biol* 9: 255–262.
- Keurentjes JJ, Angenent GC, Dicke M, Santos VA, Molenaar J, et al. (2011) Redefining plant systems biology: from cell to ecosystem. *Trends Plant Sci* 16: 183–190.
- Krakauer DC, Collins JP, Erwin D, Flack JC, Fontana W, et al. (2011) The challenges and scope of theoretical biology. *J Theor Biol* 276: 269–276.
- Nédélec F, Surrey T, Karsenti E (2003) Self-organisation and forces in the microtubule cytoskeleton. *Curr Opin Cell Biol* 15: 118–124.
- Raes J, Bork P (2008) Molecular eco-systems biology: towards an understanding of community function. *Nat Rev Microbiol* 6: 693–699.
- Arrigo KR (2005) Marine microorganisms and global nutrient cycles. *Nature* 437: 349–355.
- Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, et al. (2004) The evolution of modern eukaryotic phytoplankton. *Science* 305: 354–360.
- Bowler C, Karl DM, Colwell RR (2009) Microbial oceanography in a sea of opportunity. *Nature* 459: 180–184.
- Falkowski PG, Fenchel T, DeLong EF (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science* 320: 1034–1039.
- Karl DM (2007) Microbial oceanography: paradigms, processes and promise. *Nat Rev Microbiol* 5: 759–769.
- Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281: 237–240.
- Lindell D, Jaffe JD, Johnson ZI, Church GM, Chisholm SW (2005) Photosynthesis genes in marine viruses yield proteins during host infection. *Nature* 438: 86–89.
- Sharon I, Tzahor S, Williamson S, Shmoish M, Man-Aharonovich D, et al. (2007) Viral photo-synthetic reaction center genes and transcripts in the marine environment. *ISME J* 1: 492–501.
- Sullivan MB, Lindell D, Lee J, Thompson LR, Bielawski JP, et al. (2006) Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PLoS Biol* 4: e234. doi:10.1371/journal.pbio.0040234.
- Lupp C (2009) Microbial oceanography. *Nature* 459: 179.
- Fuhrman JA (2009) Microbial community structure and its functional implications. *Nature* 459: 193–199.
- Irigoien X, Huisman J, Harris RP (2004) Global biodiversity patterns of marine phytoplankton and zooplankton. *Nature* 429: 863–867.
- Rohwer F, Thurber RV (2009) Viruses manipulate the marine environment. *Nature* 459: 207–212.
- Smith J, Wigginton N, Ash C, Fahrenkamp-Uppenbrink J, Pennisi E (2010) Changing oceans. *Introduction. Science* 328: 1497–1528.
- Sarmento H, Montoya JM, Vázquez-Domínguez E, Vaqué D, Gasol JM (2010) Warming effects on marine microbial food web processes: how far can we go when it comes to predictions? *Philos Trans Royal Soc London B Biol Sci* 365: 2137–2149.
- Arnaud-Haond S, Arrieta JM, Duarte CM (2011) Global genetic resources. Marine biodiversity and gene patents. *Science* 331: 1521–1522.
- Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, et al. (2007) The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol* 5: e77. doi:10.1371/journal.pbio.0050077.
- Gilbert JA, Meyer F, Antonopoulos D, Balaji P, Brown CT, et al. (2010) Meeting report: the terabase metagenomics workshop and the vision of an Earth microbiome project. *Stand Genomic Sci* 3: 243–248.
- John SG, Mendez CB, Deng L, Poulos B, Kauffman AKM, et al. (2010) A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environ Microbiol Rep* 3: 195–202.
- Beal LM, De Ruijter WPM, Biastoch A, Zahn R (2011) On the role of the Agulhas system in ocean circulation and climate. *Nature* 472: 429–436.
- Follows MJ, Dutkiewicz S, Grant S, Chisholm SW (2007) Emergent biogeography of microbial communities in a model ocean. *Science* 315: 1843–1846.
- Raes J, Letunic I, Yamada T, Jensen IJ, Bork P (2011) Toward molecular trait-based ecology through integration of biogeochemical, geographical and metagenomic data. *Mol Syst Biol* 7: 473.
- d'Ovidio F, De Monte S, Alvain S, Dandonneau Y, Lévy M (2010) Fluid dynamical niches of phytoplankton types. *Proc Natl Acad Sci U S A* 107: 18366–18370.
- Tittensor DP, Mora C, Jetz W, Lotze HK, Ricard D, et al. (2010) Global patterns and predictors of marine biodiversity across taxa. *Nature* 466: 1098–1101.
- Barton AD, Dutkiewicz S, Flierl G, Bragg J, Follows MJ (2010) Patterns of diversity in marine phytoplankton. *Science* 327: 1509–1511.
- Bragg JG, Dutkiewicz S, Jahn O, Follows MJ, Chisholm SW (2010) Modeling selective pressures on phytoplankton in the global ocean. *PLoS ONE* 5: e9569. doi:10.1371/journal.pone.0009569.
- Gianoulis TA, Raes J, Patel PV, Bjornson R, Korbel JO, et al. (2009) Quantifying environmental adaptation of metabolic pathways in metagenomics. *Proc Natl Acad Sci U S A* 106: 1374–1379.
- Temperton B, Gilbert JA, Quinn JP, McGrath JW (2011) Novel analysis of oceanic surface water metagenomes suggests importance of polyphosphate metabolism in oligotrophic environments. *PLoS ONE* 6: e16499. doi:10.1371/journal.pone.0016499.