Sailing the Seven Seas for an Epic Plankton Study

The *Tara* research schooner carried scientists on a four-year expedition, gathering thousands of water samples for imaging and sequencing to study plankton diversity and assess the impact of climate change on ocean ecology.

Introduction

In 1995, Eric Karsenti, PhD, was leading a cell biology team at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany. While conducting cell cycle research, he decided to read Charles Darwin's account of his global survey expedition aboard the HMS Beagle. During the five-year journey, Darwin spent most of his time on land, making flora and fauna observations that would later fuel his theory of evolution. After reading the book, Dr. Karsenti began thinking about the benefits of a similar expedition to map ocean biodiversity. In 2009, the idea became reality and the 110 foot Tara research schooner set sail. Over a four-year journey, researchers sampled, analyzed, and sequenced > 35,000 salt water samples from around the globe for an extensive plankton study.

"Although tiny, these organisms are a vital part of the Earth's life support system, providing half of the oxygen generated each year on Earth by photosynthesis and lying at the base of marine food chains on which all other ocean life depends," Dr. Karsenti said. Yet only a small fraction of plankton has been cultured in the laboratory. Until recently the organization, evolution, and dynamics of marine ecosystems have been poorly understood.

The Tara Oceans Expedition united more than 200 scientists from various disciplines, combining ecology, oceanography, cell biology, genetics, and systems biology, to study plankton in a context that will improve understanding of their interactions with the environment. In addition to high-throughput microscopy imaging and flow cytometry, next-generation sequencing (NGS) studies were performed to uncover the presence of uncultivatable plankton, and assess community diversity and dynamics.

To learn more about the expedition, the sequencing data produced, and the discoveries that were made, iCommunity spoke with three central figures in the project. Dr. Karsenti is the expedition organizer and recipient of the 2015 Centre National de la Recherché Scientifique Gold Medal. Chris Bowler, PhD, is the Scientific Coordinator of the *Tara* Oceans expedition and leader of the Environmental and Evolutionary Genomics Section at the Institut de Biologie de l'Ecole Normale Superieure (IBENS). Patrick Wincker, PhD, is the Project Director for Sequencing Technology and Eukaryote Genomics at Genoscope.

Q: Why was the *Tara* Oceans Expedition undertaken and how quickly did it come together?

Eric Karsenti (EK): I originally began thinking about a scientific ocean voyage back in 1995. I thought it would be interesting to perform an *HMS Beagle*–like expedition to survey plankton diversity and assess their dynamic role in the ocean ecosystem. I also realized that the romanticism of a sailing voyage would catch people's attention, giving us an excuse to communicate the science of the expedition to nonscience audiences. I did not think about it again until 2000 when I decided to learn how to put together this type of expedition. I talked with Christian Sardet and Gaby Gorsky at the Observatoire Océanologique de Villefranche and they recognized the value immediately. I contacted Romain Troublé, the owner of the *Tara* schooner, and that's when we decided to form the expedition.

Q: What was the goal of the expedition?

EK: The idea of this expedition was first to do an evaluation of the state of life in the ocean. It started with collecting samples. Our goal was to sample ocean plankton in the upper sun-lit photic zone (660ft/200m). We also took samples deeper than the ocean twilight zone (> 6000ft/2000m). During the four-year expedition, the *Tara* sailed 90,000 miles (140,000 km), gathering samples from more than 200 sites at various depths.

From left to right: Patrick Wincker, PhD, Eric Karsenti, PhD, and Chris Bowler, PhD, are part of an interdisciplinary team that collected and analyzed hundreds of oceanic samples. Their goal was to assess the structure, diversity, and dynamics of oceanic plankton ecosystems.

The *Tara* research schooner collects samples near the Galapagos Islands in the Pacific Ocean. Photo courtesy of the Tara Oceans Project.

Q: Drs. Bowler and Wincker, when did you join the project?
Chris Bowler (CB): I got in touch with Eric when I heard about the project through the grapevine. My lab studies diatoms, organisms that form an important foundation of plankton communities in marine and freshwater ecosystems. Diatoms are photosynthetic unicellular eukaryotes and are believed to be as important as tropical rainforests for generating oxygen and removing carbon dioxide from the atmosphere. Yet, we know little about them. For the last 20 years, our diatom studies have been lab-based. The *Tara* Oceans Expedition offered a unique opportunity to obtain samples from the open ocean and coastal regions to test hypotheses that we had generated in the lab. We were excited about the project and invited a few others to join, slowly building the complete team.
Patrick Wincker (PW): Genoscope became involved in 2008 as the scientific consortium behind the project was being assembled. One of the early participants referred us, and we were eager to join. We offered to perform sequencing and data analysis of the samples.

> "We realized that the complexity of the community that we were sequencing was very high and we needed a higher throughput system to gain biological insights. Ultimately, we used six HiSeq Systems for this project."

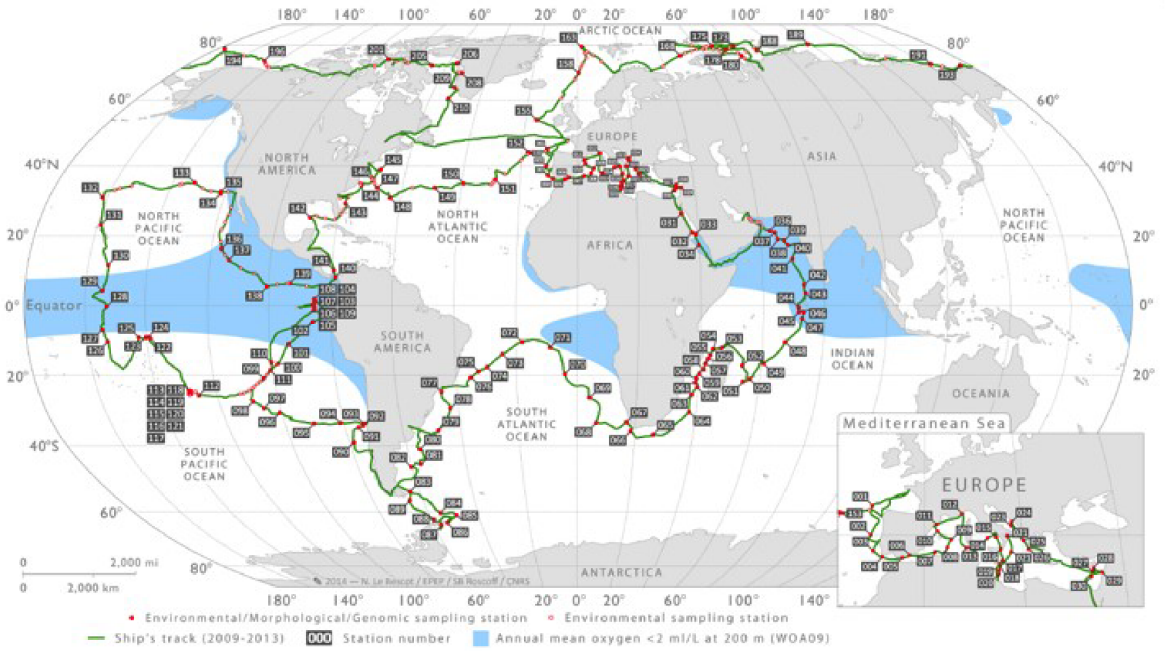
Q: How did you decide where to take the samples? What methods did you use to keep track of them?
PW: We worked with other members of the consortium to define and prioritize a list of sampling locations that best represented different oceanic conditions. We wanted to capture the full diversity of oceans on all sides of the globe.
EK: We spent about 60 hours at each of the 210 sampling stations, or approximately 12,600 hours collecting samples. We sampled to depths of 500–600 meters for biological specimens. For characterization of the water column, we used physicochemical instruments and routinely took samples to depths of 1000 meters. We developed a sophisticated logistics system where each sample was barcoded, itemized on a log sheet, and associated with date, time, environmental location, and other data.

> "We had samples that were fractionated and enriched in viruses, bacteria, or larger organisms, with each of those entering the sequencing pipeline. Through bioinformatics, we assembled all the data into something coherent."

CB: You can imagine the complexity of the tagging process, with several samples taken at the same sampling station and at the same depth, but for different analyses. We had samples that were fractionated and enriched in viruses, bacteria, or larger organisms, with each of those entering the sequencing pipeline. Each barcode enabled us to follow the sample all the way into the lab. Through bioinformatics, we assembled all the data into something coherent. It was a complex, logistical procedure to get all of that worked out. We managed to do it well, thanks to a dedicated team of people.

> "It's unbelievable that we have come so far in such a short time. Illumina sequencing systems were introduced soon after we started… We had no clue that they'd enable us to sequence all the planktonic life in the ocean."

Q: What sample sites were selected to assess differences in plankton life?
EK: This first expedition was a global phenographic sampling to characterize how the whole plankton ecosystem is structured and organized, how it functions, and how different parameters of depth and location impact it. In addition to taking samples along coasts and in the open ocean, we also gathered samples from mesoscale eddies, upwellings, acidifying waters, and anaerobic zones.
This is an initial sampling that we will use as a baseline for comparison with samples from future expeditions. Within a few years, we'll have analyzed all the data and have a clearer understanding of the ocean system. Then we can set research goals and return to specific sampling locations as part of a scientifically based plan.



Inspired by Darwin's voyage on the *HMS Beagle*, the *Tara* expedition took four years to sail 90,000 miles (140,000 km). Scientists collected over 35,000 salt water samples from more than 200 sites. Sampling stations are indicated in red.

Q: What types of analyses were performed on the samples?
CB: We performed microscopy, flow cytometry, nutrient measurements, and sequencing of some samples. Everything on the boat was prepared carefully for specific protocols when the samples got back on land. We decided to conduct sequencing for all types of organisms, from viruses to small metazoan, and to perform those studies on the same series of samples.

Q: What sequencing systems did you use for the studies?
PW: At the very beginning of the project, we were still working with the Roche 454 System. That didn't last long. We realized that the complexity of the community that we were sequencing was very high and we needed a higher throughput system to gain biological insights. We switched first to the Genome Analyzer™ System, and then quickly moved to the HiSeq™ 2000 System. Ultimately, we used six HiSeq Systems for this project.

> "We're now performing single-cell sequencing to understand the role of specific organisms … the next step is to assess gene expression and microbial interactions under different environmental conditions."

Q: What types of sequencing studies have you performed so far?
PW: We've performed ribosomal RNA sequencing to profile eukaryotic diversity in the photic zone and metagenomics to study viruses, prokaryotes, and picoeukaryotes. Not all the sequencing is finished. We've published the first series of research based on metagenomic data from 579 samples taken at 75 sampling stations. Five papers were published in Science covering: eukaryotic plankton diversity in the sunlit ocean;[1] the structure and function of the global ocean microbiome;[2] patterns and ecological drivers of ocean viral communities;[3] determinants of community structure in the global plankton interactome;[4] and the environmental characteristics of Agulhas rings and how they affect interocean plankton transport.[5]

Q: Were there any surprises in the initial data?
EK: We were surprised by the large diversity of eukaryotes that we found. In terms of species, the eukaryotes are much more diverse than bacteria and viruses. We also did not expect the amount of data that we obtained. We have almost a full characterization of the whole ecosystem (viruses, bacteria, and eukaryotes) between the surface and 100 meters, and quite a good characterization of the water layer between 100 and 500 meters.
CB: The eukaryotic diversity is astounding. There are many new organisms that have yet to be identified, which is exciting. About 90% of the sequences are from organisms that we cannot place with a specific name, and we can't put a third of them into the eukaryotic tree of life. The diversity in the ocean is large, but it is also finite. As we sequence more samples, we'll get pretty close to the final, complete set of information.
PW: We think that we have most of the bacterial gene types that are present in the ocean. There are probably tens of thousands of species. We now have a catalog of about 40 million genes and this catalog is close to covering most the genes present in the ocean.

Q: What type of data analysis are you performing?
CB: We're interrogating the data in a more in-depth fashion to integrate different data sets. We have many environmental parameter data, tens of thousands of images of organisms, and all this sequence information. We need to put that together to understand more about which organisms are doing what in the ecosystem. What do they look like and why are they doing what they do?
To address these questions, we need to build matrices with environmental parameters and gene expression to see what genes are corresponding to which parameters, or organisms. We have already begun to perform these analyses with the data set that we have.
PW: We're now performing single-cell sequencing to understand the role of specific organisms in certain locations. Now that we have described most of the species in terms of metagenomics, the next step is to assess gene expression and microbial interactions under different environmental conditions (location, ocean depth, water quality, etc). Ultimately, we'll have large metatranscriptomic data sets for all the gene collections.
EK: We also have some data on the correlation between the ecosystem composition and carbon flux. It appears that there are gene and ecosystem networks associated with carbon sequestration in the ocean.
CB: In addition, we're analyzing samples from specific ocean locales. For example, we are currently exploring a complete data set from the Arctic Ocean. We'll use this data set to compare Arctic ecosystems with those found in the rest of the world. It is exciting science.

> "The diversity in the ocean is large, but it is also finite. As we sequence more samples, we'll get pretty close to the final, complete set of information."

Q: You have millions of cells in your ocean samples. How do you determine which cells to use for single-cell sequencing?
PW: First we identify the cells using a ribosomal RNA biomarker, so we know what species they come from. Then we sequence cells from those organisms that are known to be important, but are also uncultured.

Q: What are the next steps to understand how the ocean ecosystems are changing?
CB: We will be assessing human impact on the oceans. We have baseline samples from coastal areas and the open ocean. Detailed databases exist of human impact on the ocean, covering the number of invasive species that have been recorded in certain areas, growth, or decline of fisheries, ocean acidification, temperature changes, coastline changes. In the future, we can analyze our results with these data sets to understand the links between different phenomena, including human impact.
New expeditions will enable us to compare sample data and understand how the ocean is changing as a consequence of human activities. The rich aspect of our sampling is that we have many samples from the open ocean. That's unusual, because it's not every day that boats go out to the middle of large oceans to sample plankton. These open ocean samples are a valuable resource.

Q: When you began your careers in science, could you have imagined being able to sequence hundreds of ocean samples so fully and completely?
CB: I'm from the generation where we were cloning genes one by one, and a PhD project would be sequencing a gene. It's unbelievable that we have come so far in such a short time. Illumina sequencing systems were introduced soon after we started thinking about the expedition. We had no clue that they'd enable us to sequence all
the planktonic life in the ocean.
PW: This project wouldn't have been possible without next-generation sequencing.
EK: Even when we started the project, I did not realize that we would be able to perform sequencing as quickly as we did. It was a big surprise.

Q: When will the next global science expedition set sail?
EK: There is a new expedition planned for 2016 that will focus on sampling the water and coral in the coral reefs of the West Pacific. Genomics will play a large role in this study as well.

References

1. de Vargas C, Audic S, Henry N, et al. Ocean Plankton: Eukaryotic plankton diversity in the sunlit ocean. *Science*. 2015;348(6237):1261605.
2. Sunagawa S, Coelho LP, Chaffron S, et al. Ocean Plankton: Structure and function of the global ocean microbiome. *Science*. 2015;348(6237):1261359.
3. Brum JR, Ignacio-Espinoza JC, Roux S, et al. Ocean Plankton: Patterns and ecological drivers of ocean viral communities. *Science*. 2015;348 (6237):1261498.
4. Lima-Mendez, Faust K, Henry N, et al. Ocean Plankton: Determinants of community structure in the global plankton interactome. *Science*. 2015;348 (6237):1262073.
5. Villar E, Farrant GK, Follows M, et al. Ocean Plankton. Environmental characteristics of Agulhas rings affect interocean plankton transport. *Science*. 2015;348(6237):1261447.

Learn more about the Illumina products and systems mentioned in this article:

HiSeq Systems:

www.illumina.com/systems/hiseq_2500_1500.html

Illumina, Inc. • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com