

Trawling the ocean virome

Even without a stint on an ocean-faring vessel, scientists can trawl through data to explore marine viruses and address new puzzles and cultural shifts.

Vivien Marx

Many scientists focused on marine biology split their time between periods on ocean-faring vessels and lab time on land. Others stay entirely dry with in silico deep dives. Computational ocean exploration might mean sifting through ocean data on fish, coral reef populations or physicochemical measurements of the global oceans. To some, what's most tempting is to computationally swim among marine viruses. A [separate story](#) in this issue describes some of the scientific questions researchers pursue about the ocean virome.

Computational ocean virome explorers do not have ready-made software pipelines. Trawling through these data is “just as daunting as handling a big metagenomic or metatranscriptomic dataset,” says Kathryn Campbell, a member of the Texas A&M University at Galveston lab of marine biologist Jessica Labonté, currently at sea. When mining any number of online resources, Campbell recommends a focused search, “because otherwise it becomes overwhelming very quickly,” she says. “It’s just too much data.”

In his experience, “all bioinformatics related to RNA viruses is pretty manual unfortunately,” says Guillermo Dominguez-Huerta, a former postdoctoral fellow in the lab of Matthew Sullivan at Ohio State University who is now a scientific consultant for the lab and who was a joint first author on two recent studies of marine RNA viruses^{1,2}. The team has developed ways to automate teasing out RNA viral sequence from the data, but beyond that, everything is indeed rather manual. Thus, he says, researchers should inspect their work continuously “to make sure it makes sense.”

Where to get ocean data

Microbial biodiversity surveys have often been done in a number of generally better-studied regions³, as with the [San Pedro Time Series](#) from the San Pedro Channel off the coast of Southern California. Global surveys have also been emerging, such as the [Sorcerer II Global Ocean Sampling Expedition](#) from 2004 to 2006 launched by J. Craig Venter. There are also data and samples from the [Malaspina](#)



Tara Mission Microbiomes is a marine expedition currently underway to collect data on how climate change and pollution affect ocean microbes and ocean ecosystems. Here, Milena Cerda rinses a piece of microplastic. Different types of microplastics are substrates for distinct microbiomes. Credit: M. Bardy, Fondation Tara Océan

[circumnavigation](#), an expedition devoted to data collection on ocean biodiversity and climate change that was led by the Spanish Ministry of Science and Innovation.

As microbiome researcher Shinichi Sunagawa of the ETH Zurich and colleagues point out⁴, sequencing technologies have advanced such that they now enable systematic and quantitative global ocean surveys. These advances, in turn, made it possible to find and assess marine double-stranded DNA virus populations. This latest work on marine RNA viruses, says Sunagawa, in which he was also involved, embeds new phylum-level findings into a “robust taxonomic framework.” In his view, this research ranks in importance with the reconstruction a few years ago of a group of bacterial genomes representing more than 35 phyla that the researchers call “the candidate phyla radiation”⁵. If one counts viruses in with other taxonomic groups, the finding might be the largest single expansion of established microbial taxonomy, he says. And he especially likes the definition of a new basal Orthornavirae

megataxon, the proposed phylum “Taraviricota. This proposed phylum is one of several findings from recently published analyses of sampling data from Tara Oceans^{1,2}, a global expedition supported by the Tara Ocean Foundation, or Fondation Tara Océan, based in France and with many partner organizations and supporters. The foundation is a major source of global data about the ocean and ocean microbes and, as its president Étienne Bourgois says, it’s a “family project.” The family business is the French fashion house agnès b., founded by his mother Agnès Troublé.

Because the family cares about the sea, they bought a 36-meter schooner from Lady Pippa Blake, widow of yachtsman and explorer Sir Peter Blake, after pirates killed him during an environmental expedition in the Amazon delta, and turned it into the expedition vessel and floating science laboratory *Tara*, devoted to understanding and protecting the world’s marine environment. It’s a way to continue what Peter Blake started, to

continue the conversation about the ocean and do research as well, says sailor-scientist Romain Troublé, executive director of the foundation and nephew of Agnès Troublé. The boat had been previously owned by explorer Jean-Louis Étienne. The foundation has supported several expeditions with *Tara* including the Tara Oceans and Tara Oceans Polar Circle expeditions, as well as Tara Mission Microbiomes, which is currently underway. The equilibrium of the planet “depends on the microbiome of the ocean in the same way we depend on our own microbiome,” says Romain Troublé. Viruses are part of the larger picture of how life is supported on the planet. It’s “a great mystery of the century” to decipher the roles, behaviors and functions of the ocean microbiome, including its beneficial effects. Over the last decade, he says, the expeditions have, for example, collected plankton samples from coastal waters, coral reefs and the high seas around the world for scientists to ask questions of. Microplastics in the ocean concentrate chemical pollutants such as pesticides, and microplastics appear to be substrates for distinct microbiomes. Polystyrene and polypropylene, for example, harbor different microbial communities. “We call it the plastisphere,” he says. All sample collection, not just of microplastics, happens with a view to scientific rigor to assure data quality, says Troublé. Many institutes are part of and support the expeditions through the Tara Ocean Foundation, including AtlantECO, the French Ministry of Research, the Swiss National Science Foundation, the US National Science Foundation, the European Molecular Biology Laboratory and the French National Centre for Scientific Research.



The schooner *Tara* is traveling 70,000 kilometers to assess marine microbial ecosystems and the teams are capturing metagenomic, metatranscriptomic, metabolomic and metaproteomic data. Here, *Tara* is near Deception Island off the coast of Antarctica. Scientists are also studying the positive effects of the ocean microbiome, including its virome. Credit: M. Le Roux/PolaRYSE, Fondation Tara Océan



Tara Oceans was an expedition initiated by EMBL researcher Eric Karsenti, here in the foreground. He is checking a rosette of Niskin bottles that collect water, and ocean microbe samples, at various depths. Sensors capture parameters such as temperature. Credit: Fondation Tara Océan

Its expedition *Tara Oceans* was initiated by cell and marine biologist Eric Karsenti of the European Molecular Biology Laboratory. The expedition ran from 2009 to 2013 and covered 125,000 kilometers of ocean, taking ocean water and samples. It collected nearly 35,000 samples of viruses, algae and plankton and delivered more than 60 terabases of DNA and RNA sequences.

The research community strives to follow FAIR data principles, the principles of findability, accessibility, interoperability and reusability, says Sunagawa. *Tara Oceans*' data troves can be found, for instance, in the [European Nucleotide Archive \(ENA\)](#), [Pangeaea](#), [Cyverse](#), [iVIRUS](#) and on [Genoscope](#). Other data-collection efforts target users with less programming experience and offer various types of data relevant to marine microbial research, he says: for example, the [Ocean Gene Atlas](#), a portal to search for a gene or protein sequence to see, for instance, its abundance on an ocean map. The [Ocean Barcode Atlas](#) lets users explore, for example, operational taxonomic units (OTU) data and plankton communities from *Tara Oceans* and OTUs from *Malaspina* prokaryote data. Sunagawa also points to the [Ocean Microbiomics Database](#) and its high-quality genome-resolved information about the global microbiome, which has sequencing data from 2003 onwards and which includes *Tara Oceans* data as well as datasets such as the [Hawaii Ocean Time-Series \(HOT\)](#), the [Bermuda Atlantic Time-series Study](#)

([BATS](#)), with its collection of ocean data dating back to 1988, and [BioGeotraces](#), with hydrographic and marine geochemical data from various expeditions.

The recent publications on RNA viruses^{1,2}, in which Sunagawa was also involved, have expanded the known diversity of these viruses, he says. They build on efforts by, for example, the research team that created and applied a cloud-based infrastructure called [Serratus](#)⁶, with which researchers can perform sequence alignment using [bowtie2](#) for nucleotide sequences and [DIAMOND2](#) for protein sequences in ‘ultra-high throughput’ on a petabase scale. Using [Serratus](#), the team identified more than 130,000 previously unknown RNA viruses, both on land and in the oceans. The wealth of resources for microbial and viral data about the oceans is helpful to the research community, but “we could still improve the connectivity between various datasets though,” says Sunagawa. That would help, for example, with searching and finding data products that are derived from primary data, such as identifiers of individual genome assemblies, genes and metagenome assembled genomes, which are all presented in different online locations. But connecting data resources is a project that itself takes resources, and such projects are hard to get funding for.

Going forward, it will be challenging, says Sunagawa, to update and keep up to date both past projects and ongoing projects such as the [Global Ocean](#)

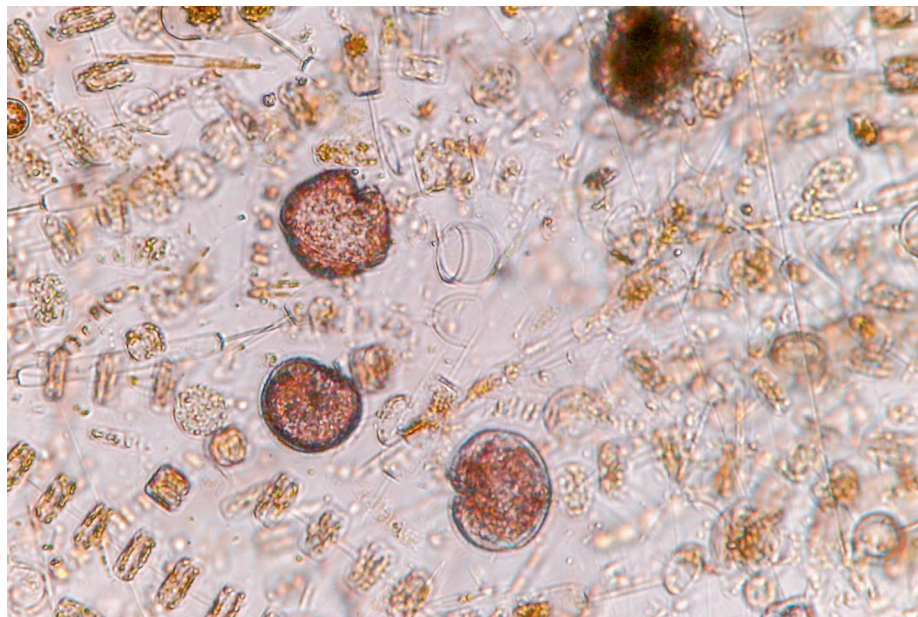
Ship-based Hydrographic Investigations program (GO-SHIP), which is focused on physical oceanography; the Antarctic Circumnavigation Expedition (ACE), on carbon-cycle marine biogeochemistry; Mission Microbiomes; and many more. “And ultimately, we will need to cross boundaries that currently separate biome-focused research to better understand processes at the sea–land–atmosphere interfaces.”

Tara Mission Microbiomes has been underway for nearly two years and wraps up in October 2022. At press time, the schooner *Tara* was off the Angolan Coast. At the end of the expedition, it will have traveled a total of 70,000 km of ocean area around South America, Africa, Europe and Antarctica. Mission Microbiomes is part of the EU-funded *AtlantECO* and also includes 42 research organizations from 13 countries. The microbiome mission is collecting data on how climate change is affecting the marine microbiome, on how pollution, microplastics pollution in particular, affects the marine environments and on the beneficial impact of the ocean microbiome.

Chris Bowler, from the Institut de Biologie de l'École Normale Supérieure, is scientific director of the Tara Oceans consortium, was scientific coordinator of the Tara Oceans expedition and was onboard in Antarctica during the Tara Mission Microbiomes expedition to collect data on the impact of icebergs on the Weddell Sea ecosystem. The project's scientists in Tara Mission Microbiomes, he says, are studying specific processes, including the Amazon plume, the Malvinas confluence, the impact of tabular icebergs in the Weddell Sea, the Benguela upwelling and more. The data from this expedition will be similar to those from Tara Oceans but, he says, “we will have much more contextual data related to the specific processes we have been studying.” The applied techniques are all ones that have undergone much advancement since Tara Oceans, he says. They include long-read



Krill are small ocean crustaceans that mainly eat phytoplankton and are a food source for animals such as whales and seals. Krill play a crucial role in biogeochemical cycles. Credit: F. Aurat, Fondation Tara Océan



Through photosynthesis, phytoplankton deliver oxygen to the planet. They are food for zooplankton, which are food for other marine organisms. This food web and its associated decomposition are part of the ocean's carbon pump, in which marine viruses play an important role that scientists have only begun exploring. Credit: M. Bardy, Fondation Tara Océan

sequencing, Hi-C sequencing to capture chromatin organization on a genome-wide basis and various types of microscopy.

Data and results from previous and ongoing expeditions are impressive, says Sunagawa but “we are still data-limited in our field of research.” Geographically, sampling stations are usually still separated by hundreds of kilometers, and often they are even further apart than that. This means that what is missing is both temporal and seasonal resolution, “and we keep detecting new organisms,” he says. Tara Mission Microbiomes will help to fill in some of these gaps. The mission is unlike Tara Oceans, with its focus more on coastal areas and environmental pollutants such as microplastics. Sunagawa and his group are not currently involved with Tara Mission Microbiomes, “but we look forward to seeing the first results coming out soon.”

Ocean cultures

Outside the ‘bubble’ of those working on marine viruses, researchers tend to be surprised that viruses exist in the ocean, says James Wainaina, a postdoctoral fellow in the Sullivan lab, a co-author of the recent *Science* papers^{1,2} on marine RNA viruses and a joint first author of one of them. It's also true that what is more generally familiar is due to what has been funded and studied before. “For example, human, animal and plant viruses have extensively been studied for their economic importance, therefore

they will be well known,” he says. Other virologists might be surprised to learn that RNA viruses have a connection to CO₂ assimilation and ocean function, says Dominguez-Huerta.

Up until around 20–25 years ago, when high-throughput sequencing technologies entered the scene, the RNA viruses that mattered to most were those that have economic importance, mainly because they infect humans, crops or livestock. “But we need to understand that that is a tiny, miniscule fraction of the RNA virosphere,” he says. And this virosphere encompasses all RNA viruses on Earth, both known and unknown. The vast majority of RNA viruses, and this is true for all viruses, infect microbes. Microbes sit at the base of food webs and drive the ecologically key biogeochemical cycles, which include, for example, the carbon cycle, the oxygen cycle, the nitrogen cycle and others.

Marine RNA viruses, says Dominguez-Huerta, mainly infect plankton, which drive the processes involved in biogeochemical cycles, including the biological carbon pump, “hence those viruses are necessarily impacting the ocean ecosystem through the infection of the very microbes that make it run.” From the first surveys of marine RNA viruses at specific sampling sites in the early 2000s, this was suspected. “But our study was meant to answer this question at a global level for the oceans.”



Marine viruses lyse the cells they infect and release organic matter, a process that plays a role in the Earth's carbon cycle, says Hiroiyuki Ogata. Credit: Kyoto University

In his experience, says Kyoto University microbiome researcher Hiroiyuki Ogata, who is another co-author of one of the *Science* papers¹, researchers outside marine virology or oceanography are open to the idea that marine viruses deserve consideration, and they appreciate the connection between marine viruses and biogeochemical cycles, including the biological carbon pump. This is also true from his encounters with the general public, where he senses open-mindedness about “the interaction between viruses and the Earth's health.”

When he interacts with physical and chemical oceanographers, however, their understanding “seems different,” which is also likely to be because they tend to think quantitatively, says Ogata. In many ways, the current metagenomic results are qualitative, he says, so many researchers in physical and chemical oceanography “tend to be unsatisfied with the lack of quantitative demonstration.” They ask, for instance, “How much carbon export is driven by viruses?” He agrees that future research will need to delve into this and other questions from a quantitative viewpoint. “However, if we consider the species interactions and experimental observations, then it seems clear to me that we should not ignore the effect of viruses in the carbon export.”

The correlations that can currently be drawn don't work quite so well for physical oceanographers who seek to work with quantitative models. Those models do have their fair share of “black boxes,” says Ogata, “while we biologists wish to shed light on individual species and their roles.” It's a slight difference in research perspective.

Marine virologists know the limitations of these kinds of studies and understand the inability to demonstrate causality, he says. Metagenomic studies imply causality between viruses and the carbon pump, but this has not yet been demonstrated. For now it's just correlation. “We need to dissect more detailed processes with a more focused sampling of sinking particles,” he says. And this direction is one he and his group are now undertaking.

Viruses and the carbon pump

The carbon pump is a global process by which CO₂ from the atmosphere is transferred to the ocean depths. In their analysis of Tara Oceans data¹, the Sullivan lab members and their colleagues at research institutes around the world note that plankton, the wide range of tiny plants and animals, dominate the ocean and sustain life on Earth. As the base of the food web, they drive the planet's biogeochemical cycles. Phytoplankton provide much of the oxygen delivered to the planet through photosynthesis, and they are food for zooplankton, which, in turn, are food for marine organisms such as fish. This food web and its accompanying decomposition of organisms are key to the ocean's carbon pump and to a process called carbon export. Organic matter, and with it carbon, sinks to ocean depths where bacteria further decompose it. Deep ocean currents bring only some of the carbon back toward the surface.

There have been studies on how DNA viruses affect ocean carbon export, and this recent work indicates that RNA virus abundances in the ocean are “strongly predictive” of ocean carbon flux, the researchers note¹. Ogata and his colleagues previously showed a correlation between DNA and RNA viruses and the carbon pump⁷. This latest work, he says, connects RNA viruses to the carbon pump with a more comprehensive catalog of marine RNA viruses than he and his team had previously captured.

For example, in the newer study, over 1,200 viral operational taxonomic units (OTUs) were identified as part of RNA virus subnetworks that are “predictive community biomarkers for carbon export.” But these findings are still in the realm of correlations. As Ogata explains, the metagenomics studies

from his group and others have indeed implied a connection between viruses and the carbon pump, but causality has not yet been established: “It is still in the range of correlations,” he says.

Not just disease

Invertebrates such as arthropods have been thought of mainly as vectors for diseases such as dengue, and there has been less understanding of “natural” viromes beyond these pathogens. Metagenomics has changed this approach by showing that much RNA diversity exists in animals, which indicates, the authors of one commentary⁸ note, that disease-causing viruses “are now the exception rather than the rule.”

The work on arthropods “opened the gates of the new world of invertebrate RNA viruses,” says Dominguez-Huerta. The team captured highly divergent taxa⁹, and this information made it possible to establish several new taxonomic groups. That team showed “the immensity of the unknown space of invertebrate RNA viruses,” he says, compared to the space of arthropod-carried RNA viruses that are human pathogens and that cause diseases such as dengue, Zika, West Nile and others. In the case of marine viruses, he and his colleagues worked to “illuminate this unexplored, unknown RNA virosphere” involving plankton in the global oceans.

Plankton includes many types of organisms, including a wide variety of protists. He hopes the team's work opens the gates to a new world of RNA viruses that infect microbial eukaryotes — protists and fungi — and that, says Dominguez-Huerta, these viruses “are absolutely different than anything else seen before.” □

Vivien Marx 

Nature Methods.

 e-mail: v.marx@us.nature.com

Published online: 29 July 2022

<https://doi.org/10.1038/s41592-022-01568-2>

References

1. Dominguez-Huerta, G. et al. *Science* **376**, 1202–1208 (2022).
2. Zayed, A. A. et al. *Science* **376**, 156–162 (2022).
3. Gregory, A. C. et al. *Cell* **177**, 1109–1123.e14 (2019).
4. Sunagawa, S. et al. *Nat. Rev. Microbiol.* **18**, 428–445 (2020).
5. Brown, C. T. et al. *Nature* **523**, 208–211 (2015).
6. Edgar, R. C. et al. *Nature* **602**, 142–147 (2022).
7. Kaneko, H. et al. *iScience* **1**, 1002002 (2021).
8. Zhang, Y.-Z., Shi, M. & Holmes, E. C. *Cell* **172**, 1168–1172 (2018).
9. Shi, M. et al. *Nature* **540**, 539–543 (2016).