HOT TOPICS IN MARINE BIOLOGY 7.1

Endless Microbial DNA Sequences, Most Beautiful

The concentration of bacteria in the ocean surface waters varies from about 10^4 to 10^6 cells per milliliter. Given the potential rate of cell division of bacteria on the order of hours and the variability of nutrients in the ocean, this is strange. Why so constant? Some have suggested that some regulatory mechanism exists, such as predation by protists, but this is not known. Does the constancy of numbers reveal anything about the constancy of the ocean?

When we look at bacteria and other microorganisms, we immediately notice that there is a tremendous diversity of metabolic styles, which makes the constancy of total numbers even more peculiar. If we go to Yellowstone National Park, we discover bacteria that are specially adapted to live in boiling water and very high concentrations of metals and sulfur. We find bacteria with similar adaptations in the deep sea near volcanically active deep ocean ridges (see Chapter 16). Other bacteria are photosynthetic, with varying biochemical mechanisms. Cyanobacteria are especially abundant in the surface ocean, using nitrogen fixation to convert dissolved nitrogen gas to ammonia and eventually to proteins. One newly discovered group of bacteria may use a light-harvesting protein simply to gain energy and not for photosynthesis in the conventional sense. Other bacteria in the same small blob of ocean water are strict saphrophytes, using the breakdown products of cells directly as food. One of the most exciting discoveries in recent years was the discovery of coexistence in the same water of different species* of the cyanobacterium *Prochlorococcus* with very different abilities to grow under different light conditions. This discovery was done with aid of PCR, discussed on pages 182–183.

The numerical dullness may, therefore, be very misleading. We seem to have a very diverse array of microbial organisms in the sea with profoundly different metabolic lifestyles, chemical ecologies, and ecological dependencies. This diversity ranges from distantly related groups with very different biologies to closely related bacterial strains that are adapted and grow in response to different conditions.

But how diverse? Well, every time we look we find more. Almost every decade, new microorganisms are discovered to be of great abundance and importance in the ocean. On page 182 we discuss the discovery by means of PCR of groups of Archaea in open Pacific waters, which had never been discovered before. In the past couple of decades, there has been great focus in the water column on cyanobacteria and viruses, both of which are very small, difficult to investigate with conventional techniques, and even difficult to identify. Yet they are very abundant and important in ecological processes. That is why molecular methods have become so important and why we discussed some useful methods in this chapter.

But how can we really get an idea of the amount of microbial diversity in the sea? The methods we use today are really designed to identify and count what we already know are present. PCR-based techniques have produced evidence for a striking diversity of approximately 1,500 nonautotrophic bacterial species, most of which cannot be cultured in the lab. But even this is clearly not an approximation of the total diversity. You can't do immunology unless you have antibodies. You can't make antibodies unless you can already isolate the organism or a close relative in culture. The same goes for many molecular probe and sequencing techniques. Molecular probes require past knowledge of what already exists. We can only search for what we already know about. This is like looking under the lamppost for your lost keys. But what if your keys are out there in the dark? Not only are the number of bacterial species undersampled, but we also are probably not sampling many gene types for which we have no PCR-based primers or species whose existence we do not even know about in the first place.

One solution to this problem is audacious, and it has been tackled by a large team of molecular biologists. J. Craig Venter, a pioneer in DNA sequencing, converted his yacht *Sorcerer II* into an ocean-going laboratory (see Box Figure 7.1), designed to sample seawater and use modern DNA technology to sequence *all* of the microorganismal DNA from seawater samples. Many of these organisms are completely unknown, so conventional probes are not being used. Instead, they are employing a method known as shotgun sequencing, which breaks up DNA into many shorter stretches without any specificity, using high-throughput sequencing methods, and then using sophisticated DNA sequence matching and alignment programs to link sequences together into genes and even whole genomes. DNA was divided into small fragments, which were cloned and sequenced. These so-called *reads* were then aggregated into longer stretches, called *contigs*, which were connected by overlapping identical sequences. A statistical continuation of this process produced longer and longer sequences. Many of the sequences could be identified by statistical comparisons with the nearly 600 genomes already sequenced. Other techniques are analyzing the DNA sequences to characterize the proteins that might be encoded by some of the DNA sequences.

The journeys begin with sampling runs using a pump, filters, and freezers to store samples. They continue at a laboratory on land where sequencing is performed and end with a large-scale bioinformatics analysis that matches sequences so that they could be strung into larger stretches of DNA, hopefully eventually to reconstruct the true diversity of microbial genomes in the sea. The whole process strains both modern DNA technology and computational methods, which are sophisticated and often crafted specifically for this project.

The first adventure involved a sample of 900 L of seawater from the northwest Sargasso Sea, just southeast of Bermuda. Over a billion nucleotides of total sequence derived from at least 450 distinct species, including many evolutionarily distinct evolutionary lines of bacteria. Since many species were probably missed because of inability to analyze low numbers of divergent sequences, the true number of species was likely much more, maybe in the thousands! Venter and colleagues† also characterized one million previously unknown genes. Also of great interest was a strong variation

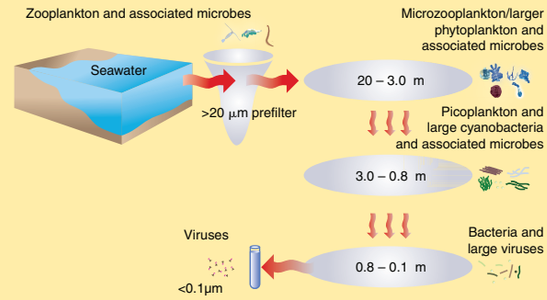
*Assigning a bacterium to one species or another has its problems, since there is no direct analogue to animal or higher plant species.

†See Rusch and others, 2007, in Further Reading, Hot Topics in Marine Biology.

▶

HOT TOPICS IN MARINE BIOLOGY 7.1 CONT

(a)



(b)

BOX FIG. 7.1 (a) The *Sorcerer II*. (b) Flowchart of sample fractionation of microorganisms in preparation for DNA sequencing. (Courtesy of the J. Craig Venter Institute.)

among individual sample sites in the apparent relative abundance of different dominant bacterial species. In total, we suddenly could see through a window at a level of organismal diversity that was literally beyond imagination using other techniques or perspectives. Their sample, as large as it was, managed to recover the entire approximate genome of only a couple of species. True complete sequences of the genomes of the many detected rare species may perhaps require thousands of times more sequence. The sampling issue is also difficult. What if that 900 L represented mainly the bacteria from a single decomposing larvacean house? It will be difficult to connect these studies to ecological processes for some time. Finally, we must remember that this study focuses on bacteria, which have much smaller genomes than the larger phytoplankton. These also are known to be undersampled with regard to biodiversity.

Since 2005, the Venter team has sampled a far vaster part of the world ocean, following the course laid by previous expeditions such as the H.M.S. *Challenger* (see Chapter 1). The journey led them from the North Atlantic through the Panama Canal and into the Pacific, spanning 9,000 km. Two recent studies reported a vast diversity of species, within and between oceans. A study of the diversity of sequences coding for proteins revealed a vast diversity of proteins and also differences from terrestrial species. Over 1.2 million new genes were identified. One of the most interesting results was the discovery of a wide diversity of genes for proteorhododopsin, a light-harvesting protein found in bacteria and closely related to another light-harvesting pigment named bacteriorhodopsin. The protein functions as a proton pump and probably provides energy to bacterial cells. This protein had been previously discovered[‡] in marine bacteria, but the known diversity in the ocean was greatly increased in this study. The widespread presence of proteorhodopsin suggests a major source of biological energy gathering in open-ocean, nutrient-poor environments.

It is somewhat arbitrary to use DNA sequence difference to demonstrate species identity, especially in bacteria, but a convention among previously studied species suggests that a greater than 3% difference in sequence indicates a species difference. Using this criterion, the investigators inferred the presence of 811 distinct bacterial strains or species in their samples. Comparisons with published DNA databases suggested that over half of these strains were not previously discovered. Differences between ocean areas were discovered, but it is not clear if this represents local evolution or just sampling of ecologically different sites.

In the past few decades, there have been great advances in the general understanding of the evolution of bacteria. A broad diversity of types has been discovered with broad capabilities to gather energy by means of light harvesting, use of sulfur and other elements for chemosynthesis, and heterotrophic uptake. The use of molecular clocks demonstrates that many of these lifestyles are very ancient, one billion years and much more. Nearly all of this information has come from molecular studies of bacteria by an emerging and very productive group of marine molecular bacteriologists. Whitman and colleagues[§] termed these organisms the "unseen majority" because they are ubiquitous, extremely abundant and important ecologically, but unseen by conventional techniques. One of the most exciting attacks on this problem has been a massive probing for microbial organisms by Edward DeLong and colleagues,[‖] who used molecular probes to investigate the depth distributions of a wide variety of microbial groups in the North Pacific Gyre. They found extensive variation in the depth distribution

[‡] See Beja and others, 2001, in Further Reading, Hot Topics in Marine Biology.

[§] See Whitman and others, 1998, in Further Reading, Hot Topics in Marine Biology.

[‖] See DeLong and others, 2006, in Further Reading, Hot Topics in Marine Biology.

HOT TOPICS IN MARINE BIOLOGY 7.1 CONT

of microbes, according to variations in metabolic adaptations and taxonomic groups. Their survey sets a framework of variation of microbial groups that now allows us to pose hypotheses about dominance of different means of processing carbon, associations of microbes with viruses, and many other types of microbial activities. For example, sequences associated with photosynthesis, as expected, were recovered with great frequency in shallow waters. DeLong and colleagues report many such depth-related variations, which can be interpreted because of the presence of extensive DNA libraries against which functional gene comparisons can be made. This type of approach differs from Venter's "shotgun" in that it targets specific known groups and gene types that have already been discovered, but these groups are clearly abundant and important in processing materials in the ocean. Previous work by Bess Ward and colleagues[¶] has produced a number of insights by focusing on nitrogen-metabolizing bacteria. The targeted functional approach is in contrast to the shotgun approach that promises some day to find everything!

Now, Venter's group has increased the stakes. We now potentially have a massive sampling tool to take ocean water and describe the total diversity, the extent of coexistence of different adapted bacterial species, and even to come across wholly new groups of important species in the economy of the ocean. The already exciting field of marine molecular biology has become even more so.

[¶] For example, See Ward and O'Mullan, 2002, in Further Reading, Hot Topics in Marine Biology.

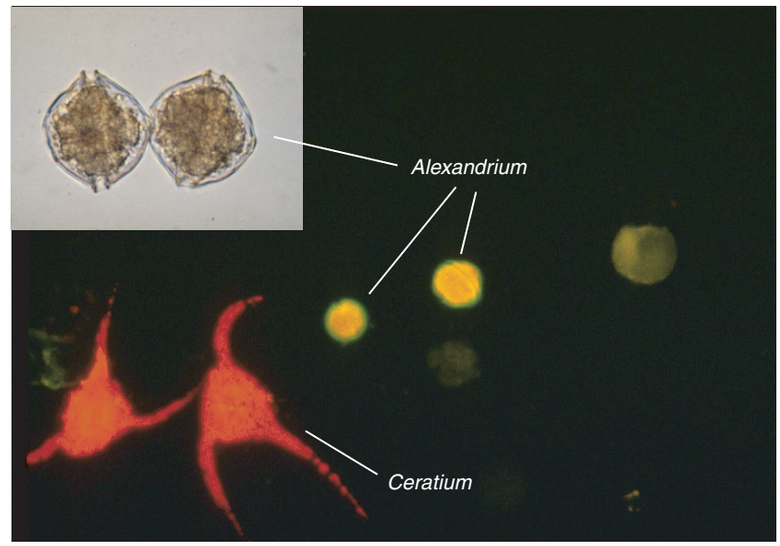


FIG. 7.19 Immunofluorescence staining: cells of the red-tide dinoflagellate *Alexandrium tamarense* stained yellow with an antibody probe to cell surface proteins, along with some *Ceratium* cells that are autofluorescing red under ultraviolet light. Light microscopic image of *A. tamerense* in upper left. (Photograph by D. M. Anderson.)

light for easy identification. First, as is done in regular immunological methods, an **antibody** is produced that responds specifically to an antigen, which may be an extract of cells of a phytoplankton species in culture or even a purified protein from that species. These antibodies are usually generated against antigens from the cell surface of the species to be probed in the plankton. Then, the fluorescent label is added, which binds to the antibody synthesized upon exposure to the phytoplankton species under study (**Figure 7.19**). When the antibody reacts with an appropriate antigen in the plankton sample, the fluorescent dye is released and fluoresces under ultraviolet light. This technique can be used to identify species of nannophytoplankton that would be otherwise difficult or impossible to identify with light microscopy. Another related immunological technique, **monoclonal antibodies**, is far more specific to a single protein and has also been used successfully to identify smaller phytoplankton.[6]

[6] See Caron and others, 2003, in Further Reading, Molecular Methods and Microorganismal Diversity.