An Ocean of Discovery: Biodiversity Beyond the Census of Marine Life

Paul V. R. Snelgrove

Department of Ocean Sciences and Biology Department, Memorial University of Newfoundland, St. John’s NL, Canada

Abstract

The 70% of Earth’s surface covered by oceans supports significant biological diversity and immense untapped potential for marine bioproducts. The recently completed International Census of Marine Life (2000–2010) invested heavily in evaluating the diversity, abundance, and distribution of life in the ocean but concluded that at least 50% and potentially >90% of marine species remain undescribed by science. Despite this potential, and numerous successes spanning pharmaceuticals, nutraceuticals, anti-foulants and adhesives, biofuels, biocatalysts (enzymes), and cosmetics, and the enormous intertidal zone to the deep plains [3]. When the CoML ended in 2010, researchers from over 80 countries around the world to catalyze discoveries on the diversity, distribution, and abundance of life in the ocean [2]. The program included 17 projects that spanned the large pool of potential “raw material” for marine bioproducts, namely the wide diversity of ocean life. The international Census of Marine Life (CoML) that ended in 2010 created a whole new level of interest and recognition of marine biodiversity for a wide range of ocean stakeholders. This 10-year program united some 2700 researchers from over 80 countries around the world to catalyze discoveries on the diversity, distribution, and abundance of life in the ocean [2]. The program included 17 projects that spanned from microbes to marine mammals and from the intertidal zone to the deep plains [3]. When the program officially ended in 2010, participants estimated that they had found some 5000 new species [2]. Noting typical lag times between the collection of new species and their formal description in the scientific literature of seven years and that some novel species may sit in specimen jars unrecognized for a century or more [4], the full range of CoML species discovery may not be fully known for another decade or more. Although none of the CoML projects considered bioactive compounds and the scientific description of those estimated new species will probably take decades to complete, the technologies and new information on biodiversity that continue to emerge from that program and related research illustrate the potential for discovery of new environments, new species, and new processes. These findings also point to great potential for marine natural product discovery that could surpass all other environments on Earth.
For example, the oceans house 34 of 36 known animal phyla in contrast to the 17 found on land; 15 phyla occur exclusively in the oceans. From metabolically active microbes deep in the ocean seafloor crust [8] to seabirds that soar hundreds of meters above the ocean [9] to the fishes and invertebrates that migrate thousands of meters through the deep ocean on a daily basis [10], our recognition of the diversity of life and range of specialized habitats in the oceans continues to expand every year. New technologies for observing, sampling, and recognizing ocean life [2, 11], many developed or enhanced during the CoML, all point towards an ocean of discovery. For example, the development of new genetic tools offers a workaround to the problem that scientists can cultivate fewer than 1% of marine microbes; these tools suggest immense microbial diversity that could reach a billion different types [12]. The diverse adaptations of marine organisms to high pressure, high and low temperatures, salt, and tolerance to environments with elevated hydrogen sulphide, methane efflux, and low oxygen all point to an immense opportunity for bioprospecting. Importantly, opportunities could emerge that utilize genetic as well as species diversity to address a wide range of societal needs.

**An Ocean of Opportunity – Marine Bioprospecting to Date ▼**

Pharmaceuticals, nutraceuticals, anti-foulants and adhesives, biofuels, biocatalysts (enzymes), and cosmetics all utilize products from marine organisms. Of these categories, medically approved pharmaceuticals represent the least productive line of marine bioprospecting to date, with the first drug approved for sale in 2004—a chronic pain product produced from a marine snail neurotoxin. Nucleosides identified from marine sponges had attracted great promise as far back as the 1950s and inspired synthetic U.S. Food and Drug Administration (FDA) approved pharmaceuticals such as cytarabine decades ago [13], but the low success rate in the direct production of marine pharmaceuticals that followed has been particularly discouraging. Drug approval typically requires an 8- to 15-year period from discovery to market, with an estimated typical development price tag of US$900 million and hundreds of thousands of failures for every successful commercial product [14]. Although the FDA has approved only seven drugs derived from marine organisms [15], the success rate relative to the 28,000 registered marine products compares favorably with terrestrial products. Furthermore, numerous other products currently in clinical trial in tandem with the potential application of marine natural products for inhibiting pain [1], different hallmarks of cancer [16], asthma [17], and herpes simplex treatment [18] points to major research potential and interest.

In contrast, marine organisms have already yielded a wide range of nutraceuticals beneficial to human health [19], with significant growth in the European and Asian nutraceutical industries totaling US$1.5 billion in 2009 and expected growth to US$180 billion by 2017 [19]. These products span from vitamins and minerals and food supplements, such as highly popular fish omega-3 oils, to compounds with cancer prevention, anti-inflammatory, antioxidant, and antimicrobial activities [20] to restorative cosmetics, the latter often derived from marine algae. Cosmetics in-clude materials developed for application to human bodies for cleansing, protection, or altering appearance without affecting body structure or function [21]. Potential cosmetics from marine organisms span from sunscreen compounds derived from algal symbionts in invertebrates [22] and aquatic vertebrates [23] to marine algal-derived anti-photoaging agents and skin whiteners in Asia [24]. A wide range of marine algal products has also been developed to treat multiple skin conditions [24].

On the one hand, the banning of many biocidal compounds from marine applications in tandem with a need to combat fouling organisms has catalyzed successful efforts to develop naturally derived antifoulants [25], including many from marine algae. On the other hand, the bio-adhesives industry capitalizes on marine organisms such as barnacles and mussels that produce strong natural glues that adhere under water. The capacity of sea stars to quickly attach and detach from hard surfaces points to intriguing possibilities for temporary adhesives [26].

Food and energy production from marine organisms also shows great promise. The utility of transgenic technology to enhance aquaculture has been recognized for decades; for example, the transfer of antifreeze proteins from cod to salmon can extend aquaculture production geographically and seasonally [27] and improve growth rates. Indeed, Atlantic salmon recently became the first FDA approved genetically modified animal. Marine algae, and diatoms specifically, show great potential for biofuel production, noting the capacity of some species to produce oils and carbohydrates from photosynthesis that may yield different biofuels [28]. Bioluminescence proteins extracted from marine organisms have also been widely used in the biotechnology industry and in research [29]. Enzymes produced by specialized marine organisms spanning from microbes to algae to animals remain stable under specific environmental conditions (high pressure, extreme hot or cold temperatures, high salt, and hydrogen sulphide and methane rich as well as oxygen poor water) and can act on proteins, lipids, carbohydrates, and other compounds. The utility of these compounds remains largely unknown, with most of the current focus on inventory [30].

**An Ocean of Opportunity – Biodiversity Technologies ▼**

Technologies developed over the last decade have rapidly advanced science capacity for new species discovery, and include tools that expedite the discovery of new habitats, others that facilitate specimen collection, and others that expedite or clarify taxonomic descriptions. Some of these capacities were developed or extended within the CoML, whereas other techniques emerged in parallel.

Dramatic improvements in acoustic tools, such as multibeam acoustics and associated analytical software capabilities that use sound waves to map seabed, greatly enhance the capacity for habitat discovery [31]. The huge areas of ocean habitat on Earth limit the proportion of well-quantified habitats to date, however, seabed mapping has matured in its capacity to differentiate broadly different habitat types and produce highly detailed physical maps [31]. Habitat offers the best predictor of species distributions [32, 33] and thus the likelihood of identifying promising locations for species discovery. Ironically, coastal depths can be especially challenging to map from surface ships because vessels required for multibeam equipment cannot efficiently navigate the shallowest waters without significant risk. Although optical-based remote sensing methods such as LiDAR (Light Detection and Ranging) can image wide swaths of shallow seabeds under clear waters [34], water actually attenuates light quickly, typically 100 s of m or less. This opacity limits optical mapping tools, including SCUBA, to a small percent of ocean depths. Noting the
ephemeral, highly patchy, and discrete nature of some marine habitats such as cold seeps and hydrothermal vents [35, 36], the expansion of exploration of ocean biota from ship-based sampling to submersibles [37] and, more recently, remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) [11] greatly increases the potential for habitat discovery, particularly in light of major advances in digital imaging and analysis. Submersibles and ROVs also provide excellent platforms for specimen collection when equipped to collect delicate biota from highly specialized locations, as illustrated by the initial collections of novel chemo- symbionts from hydrothermal vents [38]. Nonetheless, only nine nations currently operate deep diving submersibles [39] and the various nets, grabs, and bottles used for the last century continue to dominate the tools used in the collection of marine specimens, including new species [2].

New technologies enable habitat discovery through means other than habitat surveys. New and improved environmental sensors that measure a wide range of chemical constituents, including hydrogen sulphide, oxygen, temperature, methane, pH, salinity, and other variables [40], can help in discovering unique or specialized environments. Improved stability of such sensors, combined with ever-increasing miniaturization, enable the use on sampling platforms such as ocean gliders [41] and ARGO floats [42] that drift across large swaths of ocean collecting data, AUVs [43], and underwater cabled observatories [44]. Collectively, these tools provide opportunities to identify novel environments to catalyze new discovery.

The Taxonomic Impediment to Species Discovery

As time-consuming and expensive as the collection of specimens can be, the process of documenting biodiversity itself requires complex taxonomic scrutiny, and even then taxonomic confusion may arise from morphologically similar species [45], misspellings, and misidentification. The World Register of Marine Species (WoRMS), a partner of the CoML program, is cataloging all valid scientific names for marine species, with the current count at just under 230 000 of the ~245 000 known species [46]; this same project documented 56 different taxonomic names for a single species already minimize ambiguities and accelerate species discovery rates. Genetic barcoding, which first referred to animal identification using the mitochondrial cytochrome c oxidase I (COI) gene [47], has expanded to utilize other genetic identifiers [48, 49], and now provides unambiguous species identification spanning from microbes to whales. Species identification using genetic barcoding requires comparison of genetic material from specimens to library sequences of similar genetic material. Therefore, the current lack of a complete library of known and unknown marine species constrains genetic taxonomy and will continue to do so until every species has been barcoded, a very long-term prospect given the many unknown species and current rates of species discovery (see below). Moreover, barcoding may produce errors just as serious as those that have plagued morphological taxonomy, and points to the need for collaborative approaches between researchers using genetic and morphological-based taxonomy. Nonetheless, barcoding provides an unambiguous tool for differentiating among specimens within the library, and the convergence of barcoding with morphological taxonomy offers huge potential in cataloging ocean life.

High-throughput mass sequencing encompasses several different genetic approaches, including environmental DNA sampling that can be likened to “biota in a blender”; in essence, analysis of gene sequences within a sample of water, sediment, or even biota within organisms, produces estimates of the numbers of species or operational taxonomic units (microbiologists in particular use OTU to refer to genetically distinct life forms that may or may not conform to different species definitions). Mass sequencing can therefore identify the number of different life forms in a sample, even those forms undescribed taxonomically. Indeed, concerns about the rapid loss of coral reef diversity prompted CoML scientists to propose the mass sequencing of reef biota, even if simply to produce a list of genetic identifiers of unknown species [50]. At least the level of biodiversity loss might be quantifiable even in poorly known areas. Genetic techniques also allow for the identification of specific gene expressions, such as toxic strains of dinoflagellates. Indeed, the Environmental Sample Processor was developed to collect in situ samples and test for the presence of a particular gene in microbial samples or in larval invertebrates [51]. When deployed on ocean observatories that relay data to shore [44], this strategy could provide real-time information on the presence and activity of species of interest.

For some of these new technologies, the application for bioprospecting may not be immediately obvious. Certainly any technologies that facilitate specimen collection can expedite the discovery of new bioproducts, and mass-sequencing tools can help to identify biodiversity hotspots that could be targeted for exploration. Similarly, improved sensors and platforms offer the promise of the discovery of new environments or at least new habitat locations with potentially novel species, such as the first discovery of hydrothermal vents in the Southern Ocean recently [52].

Novel Environments

The propagation of technologies that enable new habitat discovery offers a two-pronged benefit for marine bioprospecting. New habitats often support novel species and thus bioproduct potential, but perhaps more importantly, by human standards, the extremes of temperature, chemistry, and pressure that characterize many of these newly discovered environments require specialized biota with specific enzymes and physiological adaptations. The examples below illustrate new environments, most discovered in the last few decades. Not surprisingly, most novel habitat discoveries occur in deep and remote waters [11]. In some of these environments, such as hydrothermal vents, new species discoveries show no sign of slowing down [4], even though the first reports of unique hydrothermal vent fauna appeared almost 40 years ago [38].

Ice biota offers a shallow-water example of novel habitats highlighted by the CoML. Scientists first reported cold and ice-adapted organisms and ice enzymes long ago [53], and transgenic technology has already improved aquaculture growth rates of cold water fishes [54], however, knowledge of entire ice-associated communities that span from microbes to fishes and marine mammals only recently came into sharp focus [55, 56]. From a bioprospecting perspective, the low diversity of ice biota offers modest promise for new bioproducts with only a few species per major taxonomic group [53], but the extreme conditions in which these species live (temperatures <−10 °C and salinities >100 ppt) suggest potentially unusual physiological adaptations. Polar environments more generally support organisms capable
of anabolic and catabolic processes at cold temperatures that could offer significant potential for new habitat discovery, particularly given the large areas of undersampled Arctic and Antarctic environments such as those under multiyear ice [53, 57] or remote deep fjords.

Deep-water coral patches were first reported in the 1800s [58], but first recognized as distinct habitats in the early 20th century [59]. Habitat models now predict likely locations for corals on seamounts [60] and are based on environmental characteristics [33, 61]. We now know that deep-water corals [62] and sponges provide a distinct habitat for a variety of invertebrates and microbes, and bioprospecting has already begun to target sponge microbes [63] and a wide range of natural products from corals in general [6].

The discovery of hydrothermal vents in 1978 changed our view of life on Earth and illustrated the potential for the discovery of fundamentally different ecosystems, in this case an ecosystem fueled not by sunlight but by chemicals such as hydrogen sulphide and methane emitted from superheated seawater within the seafloor [38]. The capacity of some vent organisms to tolerate normally toxic hydrogen sulphide, with little to no oxygen, and seawater heated well above ambient levels [64] forced the rewriting of textbooks describing metabolic pathways, enzyme characteristics, adaptation, and even basic ecology. The complex biogeography of hydrothermal vents [65] and high levels of endemism and taxonomic distinctness [66] add to the expectation of discovery. Research cruises in recent years have discovered new vent habitats in the Southern [52], Indian [67], and Arctic [68] Oceans, with expectations of more vent habitats to come [36].

The discovery of seeps and their chemosynthetic communities [69] some 30 years ago was thought to represent a habitat extension of hydrothermal vents, but we now recognize seeps, which may include methane, natural gas, or hydrocarbon emissions, as quite distinct habitats with little species overlap with vents [66] and high rates of endemism [70]. Although we know that seeps, and the mud volcanoes that characterize some seep environments, occur widely in all oceans [71], new discoveries continue, such as the first report of an Antarctic seep just a decade ago [72]. Oxygen minimum zones and anoxic basins exclude many or most life forms, respectively [73, 74], but the taxa that can tolerate such conditions necessarily display specialized adaptations and physiologies to cope with low oxygen. Anoxic basins, such as those at 3200–3600 m in the deep Mediterranean, exclude most life forms other than specialized bacteria and Archaea [74]. But DNA evidence points to the presence of living Protista [75]; metabolically active species from the animal phylum Loricifera with specialized organelles and symbionts were recently reported from the anoxic L’Atalante Basin [76].

The last decade also saw recognition of the “deep biosphere,” where “deep” refers to depths of more than 1000 m below seafloor sediments deep in the Earth’s crust in an environment thought devoid of life until just two decades ago [8]. The presence of metabolically active microbes in these environments points to capacities to utilize extremely limited resources and/or a greater prevalence of symbioses than previously recognized [11]. These findings suggest that our current expectations of adaptations and life forms underestimate the novelty of nature and physiology, and the associated potential for discovery of fundamentally novel bioproducts and processes in the global oceans. But new species discovery hardly requires new environments, noting that even well-studied ocean regions such as the Gulf of Maine continue to yield new species [77]. In addition to discrete, hard-to-locate habitats such as those described above, the sheer volume of the ocean and the massive volume of biological habitats it encompasses point to a discouragingly large potential for species discovery. Indeed, science has sampled only a few percent of the global oceans [4, 78] and although specific habitats such as coral reefs have yielded the most natural products to date, the reality is that the 95% of the global biosphere encompassed by the oceans offers great potential for species and habitat discovery.

Who – Known and Unknown Diversity

So where does this plethora of new tools leave us in terms of completely documenting ocean diversity? The short answer is that we probably have a very long way to go. Despite the investment of many millions of U.S. dollars in the CoML, many unknowns remain and the program illustrated the unknown as effectively as the newly known for species discovery. Marine scientists currently describe about 1650 new species per year [79], but available taxonomists primarily constrain that number [80] more than the availability of novel species. The technologies described above can facilitate, but not overcome, this “taxonomic impediment” [81]. Even marine experts cannot agree on the correct order of magnitude of unknown species, with total species projections beyond 10 million [7, 82] to as little as 300,000 [83], with median total marine species projections of 1 million [5] to 2 million [84]. Importantly, these estimates do not include microbes, whose diversity could number over a billion [12]. A simple calculation based on a prediction of 2 million species and a stable annual description rate of 1650 new species per year suggests that we might expect to finish inventorying marine animals in about 1000 years. This gap is particularly interesting in light of the rapid increase in microbial patents in recent years, catalyzed by the genetics revolution [85]. Although a complete marine species inventory looks highly unlikely in any of our lifetimes, or perhaps even those of our great, great, great grandchildren, we have sufficient knowledge to ask where new discoveries are most likely and which groups of organisms are poorly versus well described. Although our knowledge of marine biodiversity decreases with the size of organisms, new species include some large-sized taxa such as the 3-kg lobster discovered off the coast of Africa [86]. Even widespread species can elude scientists, such as Prochlorococcus, first reported in 1988 [87] and now recognized as the most abundant photosynthetic organism on Earth. The rates of new species description vary hugely among animal phyla, with a strong bias towards crustaceans, molluscs, and fishes [79]. Not surprisingly, projections of unknowns based on surveys of taxonomic experts for different major taxa [5] do not map well onto taxonomic effort in that fishes are relatively well known (<20% unknown), whereas the least known major taxa include nematodes (>81% unknown), flatworms (>67% unknown), sponges (>67% unknown), and molluscs (>64% unknown) [5]. Several taxonomic orders of crustaceans also have projected unknowns that exceed 60%. This bias reflects several factors [79], including the number of practicing taxonomists working on a given group, the contributions of amateur taxonomists (significant for molluscs and crustaceans), and the challenge in collecting and analyzing some of the smaller groups, noting the inverse relationship between size and taxonomic completion [88]. Intriguingly, the excellent mollusc taxonomist Philippe Bouchez commented that “the taxa for which [taxonomic extrapolations] are fairly accurate (fishes, echinoderms, decapods) con-
tribute little to the global numbers, whereas for the taxa contributing much (parasites, nematodes) the estimates are vague" [79].

**Whence – The Geography of Species Discovery**

The geography of discovery reflects a combination of patterns of total biodiversity based on current knowledge (i.e., the available species pool) and the degree to which different habitats have been sampled (i.e., the total area of that habitat and proportion of those habitats sampled). The seafloor generally supports a far greater diversity of animals than the water column above it, noting that zooplankton account for only an estimated 7000 species [89] of the – 245,000 described marine species [46] and spanning 15 of the 34 phyla known to occur in the ocean. But the story for microbes remains murkier, given the discovery of 1800 genomic microbial species and 1.2 million new genes in 200 L of Sargasso Sea surface water [90]. The recent discovery of deep-living microbes in the ocean crust [8] in tandem with the extraordinary number of rare species in the microbial realm [12] makes geographic prediction of unknowns for microbes extremely difficult. A recent examination of fishes provides some insight into the habitat types that offer the highest potential for new discovery. Fishes represent the greatest certainty in terms of knowns and unknowns because they are the best-known animal group taxonomically other than marine mammals. Examination of species discovery curves indicates 1000–4000 undiscovered marine fishes, with the largest gaps in the deep sea [91] and significant potential in tropical coastal ecosystems [92].

Considering the state of knowledge of biodiversity patterns more broadly, coastal biodiversity appears highest in the tropical Indo-Pacific, decreasing towards the poles for better known groups (fishes, corals, cephalopods), with some notable exceptions (seals, seagrasses) [93]. Temperate biodiversity peaks appear more typical for oceanic species such as tunas, squids, cetaceans, and euphausiids. Decreasing diversity with increasing latitude has also been reported for deep-sea fauna [94] and for coastal bi-valve molluscs [95], and linked to temperature [93] and energy input [94,95]. Most sampling has focussed on northern hemisphere temperate latitudes, decreasing in effort with depth [78] and decreasing sizes of organisms [88].

But total habitat area and sampling effort add further complication; coral reef habitats cover a relatively small proportion of ocean environments (< 0.2%) but harbor significant undescribed biodiversity because they are so poorly sampled [50,92]. To illustrate this point, sampling of brachyuran crabs from just 22 coral heads from the remote Pacific Northern Line Islands yielded 30% of the total brachyurans recorded from all of European seas, which are among the best studied oceans in the world and cover an immensely larger area than 22 coral heads [50]. Coral reef biologists have generated projections of ~1 million [96] to 2.5 million [97] species from coral reefs alone based on different extrapolation methods.

Deep-sea environments offer a particularly fertile environment for species discovery ranging from invertebrates [98] to fishes [91]. Despite general declines in biomass and abundance related to strong food limitation, deep-sea benthic ecosystems generally support many species per number of individuals relative to most environments [58]. Of particular note, deep-sea ecosystems typically harbor many rare species [7], which creates a challenge in trying to fully sample a given region, but rarity also creates opportunities for bioprospecting. Deep-sea environments comprise some 90% of the global biosphere by volume and > 55% of Earth’s surface, and science has sampled just a few percent of its immense content [2]. Indeed, the cumulative total of deep seafloor sampled to date for sedimentary fauna covers just a few square meters for the smallest animals (meiofauna, or animals 44 – 300 microns) [82] and a few football fields for slightly larger organisms (macrofauna, or animals > 300 microns) [4]. The water column of the deep ocean also offers a massive volume of habitat that may represent the least sampled habitat on Earth [78]. Thus, the potential for species discovery in the deep ocean remains almost unlimited for the foreseeable future.

Reduced diversity typically characterizes special deep-sea environments such as hydrothermal vents, cold seeps, upwelling regions, and trenches [98], however, as noted earlier, the organisms that utilize these environments often display unique physiological adaptations to cope with the unusual conditions. In some cases, these adaptations result in organisms distinct from other habitats at higher taxonomic levels (e.g. family, order). This high “taxonomic distinctness” may offer particular promise for bioprospectors because organisms may differ fundamentally from any other known relatives [66]. One other aspect of these special environments merits comment. The relatively discrete, discontinuity of many of these habitats differs from most ocean environments that are generally perceived as “open” with a free exchange of individuals. The relative isolation of these special environments can promote genetic and even species divergence from similar habitats elsewhere in the deep sea, explaining the distinct biogeography reported for hydrothermal vents and seeps [65] as well as trenches [99], for example.

Despite reports of generally decreased diversity at higher latitudes [93,100], polar ecosystems may not be as species depauperate as previously assumed. Comparison of sampling effort and species reported for Canadian waters indicates more species in the Arctic than the Atlantic, despite few samples collected [101]. Thus, polar systems also offer potential for discovery, particularly under multiyear ice habitats that remain largely unsampled [53,57].

Rarity adds one further complication regarding where to look for bioactive compounds. Ecologists have long known that some marine species may be quite common and others quite rare; indeed, several studies note the frequency of rare species, whether referring to microbes [12] or deep-sea invertebrates [7]. Rarity may refer to low numbers of individuals of a species in a region or total numbers of regions in which a species is found, but either characteristic adds an additional challenge for bioprospectors in terms of “missed opportunities” in the form of never sampling a given rare species. Bioactive compounds have been isolated from rare as well as common species. For example, the sponge *Lissodendoryx* sp., known from a single New Zealand peninsula, produces a strong antitumor compound, whereas the bryozoan *Bu- gula neritina*, the source of an cytostatic compound, spans a wide range of tropical and temperate coastal environments [102]. Acknowledging the challenge of rare species and not knowing whether a new natural product will come from a common or rare species adds a challenge to the search for new products.
Whence? The Geography of Bioactive Compounds

So where do new marine natural product discoveries occur? A recent analysis of marine invertebrate product discoveries since the 1990s reported more than 60% of discoveries from the Northern Hemisphere, and ~55% from the tropics [6]: the Pacific accounted for ~63%, with ~20% from the Atlantic and ~13% from the Pacific. Not surprisingly, almost half of new marine invertebrate natural products came from Asian countries, reflecting high levels of natural product research activity in Japan, Taiwan, and China. But this same analysis shows apparent shifts in effort over time, with increasing proportions of discoveries from the tropics and decreasing contributions from temperate environments. Hotspots of natural product discovery also shifted spatially, from Japan and Australia in the 1990s to Japan, China, and Polynesia-Micronesia since 2000 [103]. In the context of Large Marine Ecosystems [104], the Kuroshio and South China Sea lead the way for marine bioproducts, followed by the Caribbean and East China Seas, although this pattern partly reflects some countries registering multiple closely related products.

To some extent, these patterns reflect a shift in effort from waters of developed nations with a strong science infrastructure (North America, Mediterranean, Japan) to waters where science capacity has expanded rapidly (China and, to a lesser extent, Polynesia-Micronesia), but they also reflect increasing recognition of global patterns of diversity. Therefore, the unknowns of biodiversity map loosely onto patterns of new invertebrate natural products but a clear bias towards accessible shallow waters.

Who – The Taxonomy of Natural Product Discovery

Even in the last decades, biodiversity discoveries led taxonomists to completely redraw the Tree of Life that describes relationships among life on Earth. In 1990, a three-domain system comprised of Bacteria, Archaea (superficially similar to bacteria but structurally different), and Eukarya (fungi, plants, animals, single-celled protists) replaced the five kingdom system that dominated textbooks for generations [105]. This realignment created controversy; for example, some key similarities between Archaea and Eukarya point to a two-domain system [106]. While this realignment debate may appear esoteric from a natural products point of view, it illustrates that at the very highest level of biological diversity – the domain – taxonomists continue to realign relationships based on morphologies, physiologies, and, of course, genetics. These realignments follow from documentation of fundamentally different biological capacities in different organisms. For example, in the last decade or so, we learned that Archaea, a group previously thought limited to extreme environments, may dominate microbial biomass in the deep ocean [107,108]. We also first learned about anammox, a novel metabolic pathway in marine sediments that oxidizes ammonium, in 2001 [109]. In short, we know metabolites produced by marine microbes exhibit a wide range of bioactive properties [110], and evaluation of the marine microbial world and its potential for natural products has really just begun.

Marine animals have already yielded a wide range of natural products; in decreasing numbers of compounds to date, these include terpenoids, alkaloids, aliphatics, steroids, carbohydrates, amino acids and peptides, and others [6]. Although the complex approval process from natural product discovery to commercial product has contributed to relatively modest numbers of marine-derived natural products on the market [111], the wide range of marine organisms from which natural products have been discovered points to tremendous potential.

A recent analysis of taxonomic distribution of natural product discovery across groups of animals [6] demonstrates particular promise in some phyla. The sponges (Porifera) alone account for almost half of new natural products since 1990, producing a wide range of bioactive compounds [111], with corals and jellies (Cnidaria) adding another 30%. Seastars and relatives (Echinodermata), vertebrates (Chordata), and molluscs (Mollusca) add more than 5% each, with additional discoveries in polychaetes (Annelida), moss animals (Bryozoa), flatworms (Platyhelminthes), acorn worms and relatives (Hemichordata), lamp shells (Brachiopoda), and crustaceans (Arthropoda).

Interestingly, new natural products to date suggest contrasting geographic sources for new marine bioproducts in different taxa [6]. Since 1990, Porifera and Cnidaria accounted for most temperate (>67%) and tropical (>87%) new marine natural products, whereas Echinodermata dominates polar natural products (>40%). Of all taxa, the order Alcyonacea (soft corals and fans) has contributed more new natural products than any other, irrespective of geography (19.6% of polar, 16.5% of temperate, and 34.6% of tropicals).

The taxonomic focus of marine natural product discovery has also shifted over the last decade. A leading focus on Porifera that began to take off in the early 1980s peaked in discovery of marine natural products around 2000 and has remained a leading source of products [112]. Cnidarians, though not an early focus of product development, have attracted steadily increasing interest that has levelled off since 2013 and now rival sponges as a taxonomic source of products. Around 2000, interest in Ascomycota (Fungi) and Bacteria began to rise sharply and will likely continue to increase and surpass Porifera and Cnidaria in their proportional contribution to marine natural products as the natural product powerhouses.

The Irony of Marine Biodiversity and Marine Bioproduct Research Needs

The remarkable advances in genetics and biochemistry in recent decades juxtapose significant practical deterrents to marine natural product research: (1) As demonstrated during the CoMML, taxonomic effort on new species discovery focuses on larger, charismatic taxa such as fishes, molluscs, and crustaceans rather than taxonomic groups that show the greatest promise for marine products (Fig. 1). (2) A geographic disconnect in centers of scientific capacity in taxonomy and biological diversity hotspots constrains new species discovery, with the bulk of morphologically and genetically based taxonomy in temperate developed countries and the bulk of diversity hotspots in tropical environments of developing nations. (3) A geographic disconnect between scientific capacity for marine bioproduct development and the most promising taxa and environments limits new product discovery. (4) The retirement of many practicing taxonomists without replacements reflect a low societal priority for marine taxonomic research, which hinders rates of new species descriptions [79,80]. In comparing the taxa that show the most promise of new species with those that show the greatest promise of yielding marine natural products, several striking gaps emerge. Among invertebrates, the sponges and Cnidaria, and to a lesser extent the Echi-
Clearly differ from the rate lines of evidence support the assertion of a mismatch. The number of species yielding new products (Mollusca, Crustacea, Chordata, and other metazoans), undiscovered species (Porifera ~ 67% unknown, Cnidaria 20–40% unknown, Echinodermata 25–45% unknown) [5], the least known taxa (e.g., Mollusca, Nematoda, Platyhelminthes, Protista, some groups of Crustacea, and multiple lesser known phyla) have not figured greatly in marine bioprospecting [2]. Two separate lines of evidence support the assertion of a mismatch. The “star performers” for new products (Porifera, algae, Cnidaria) clearly differ from the “star performers” for new species discoveries (Mollusca, Crustacea, Chordata, and other metazoans), whether considering the number of species yielding new products or products per phylum. This mismatch extends into the projected unknowns of diversity, suggesting no relationship between bioprospecting focus and unknown species diversity (Fig. 2). Whether this mismatch between unknown and untapped species reflects a bias in where bioprospectors have looked or a real paucity of potential products in some taxa is difficult to assess, but both factors likely contribute. Geographically, new product discovery and biodiversity hotspots partially coincide, but with clear differences likely driven by science capacity and research investment. Not surprisingly, the majority of product discoveries have come from the Northern Hemisphere, and particularly from industrialized countries with a strong science capacity. In the case of some groups of organisms such as seaweeds, the success in Japan, China, Taiwan, and Korea corresponds to a region of strong research investment [113] and high levels of known macroalgal diversity and endemism [114]. Large numbers of bioproduct discoveries coincide with some invertebrate diversity hotspots such as Indonesia and China, but other biodiversity hotspots such as coastal Africa, South America, the Red Sea, the Caribbean, and the South Pacific produced modest numbers of marine natural products from 2000–2009 [6], clearly reflecting undersampling for both biodiversity and for natural product potential.

Fig. 1 Comparison of taxonomy of species discovery to taxonomy of natural products. Data on annual rates of species discovery are based on published values [79] from 2002–2003 and assume modest year-to-year variation in the discovery rate. Published values of compounds per phylum as of 2015 [112].

Conclusions and Priorities

As a marine ecologist bringing a biodiversity perspective on the future of marine bioprospecting, I believe the potential looks both compelling and exciting, but points to a need to manage expectations. The great promise of marine bioproducts built expectations of massive payoffs that, in the vast majority of cases, will never occur. Marine ecologists themselves helped fuel this expectation in arguing for conservation and sustainable practices for future economic benefits [115], but this enthusiasm has raised (legitimate) concerns about biopiracy. These concerns catalyze protectionism that limits potential discovery in developing nations that may lack the capacity or resources to develop natural products. In the absence of the benefit of sharing agreements outlined in the Nagoya Protocol, this concern will seriously limit future discovery and marine bioprospecting, noting the taxonomic and geographic distribution of many past success stories. Many opportunities remain for the discovery of marine bioproducts, but the spatial mismatch between science capacity, hotspots for biodiversity and bioproducts, and resource access and development adds a great challenge. Without question, marine resources offer promise for many developing nations with some of the richest biodiversity globally in their jurisdictional waters [113]. Relatively few documented global extinctions in the ocean [116] juxtapose many local and regional extirpations (and associated genetic losses) [117] and suggest urgency in stemming the loss of biodiversity [118, 119] and marine biopродuct potential that threatens “future options”.

For most species that have yielded bioproducts to date, we lack anywhere near sufficient data to apply the sorts of management plans developed for fisheries, for example, to ensure species sustainability. Indeed, for future options and marine bioprospecting, we rarely know which species to protect. The strong call for marine protected areas (MPAs) from marine scientists [120, 121] offers the most promising strategy for maintaining future options, particularly if efforts include known biodiversity hotspots [122]. By protecting habitats, MPAs offer at least minimal protection for unknown as well as known species. MPAs may, in some instances,
encompass a diversity of permitted uses, and marine bioprospect testing and development often require modest numbers of individuals compared to activities such as commercial fisheries or aquaculture. Marine bioprospectors could, therefore, add a contrasting and potentially effective voice for ocean conservation while simultaneously preserving future options.

**Conflict of Interest**

The author declares no conflict of interest.

**Acknowledgements**

The author thanks Biopros: The 7th International Conference on Marine Bioprospecting and Kjersti Lie Gabrielsen for the speaking invitation that made him think about the ideas presented here, and to Deniz Tasdemir for the invitation to put them in writing. He also thanks Philippe Archambault and Kim Juniper for discussions that helped in framing this review, and two anonymous reviewers who provided insightful comments on an earlier draft. Finally, he thanks the Consortium for Ocean Leadership and Kristen Yarincik for their hospitality during the sabbatical that supported the writing of the manuscript.

**References**
