

Review

Marine Metagenomics: New Tools for the Study and Exploitation of Marine Microbial Metabolism

Jonathan Kennedy¹, Burkhardt Flemer¹, Stephen A. Jackson¹, David P. H. Lejon¹, John P. Morrissey^{1,2}, Fergal O’Gara^{1,2,3} and Alan D. W. Dobson^{1,2,*}

¹ Marine Biotechnology Centre, Environmental Research Institute, University College Cork, Cork, Ireland; E-Mails: jonathan.kennedy@ucc.ie (J.K.); bflemer@web.de (B.F.); s.a.jackson@student.ucc.ie (S.A.J.); d.lejon@ucc.ie (D.P.H.L.); j.morrissey@ucc.ie (J.P.M.); f.ogara@ucc.ie (F.O.)

² Department of Microbiology, University College Cork, Cork, Ireland

³ BIOMERIT Research Centre, University College Cork, Cork, Ireland

* Author to whom correspondence should be addressed; E-Mail: a.dobson@ucc.ie; Tel.: +353-21-4901946; Fax: +353-21-4901932.

Received: 20 January 2010; in revised form: 4 February 2010 / Accepted: 12 March 2010 /

Published: 15 March 2010

Abstract: The marine environment is extremely diverse, with huge variations in pressure and temperature. Nevertheless, life, especially microbial life, thrives throughout the marine biosphere and microbes have adapted to all the divergent environments present. Large scale DNA sequence based approaches have recently been used to investigate the marine environment and these studies have revealed that the oceans harbor unprecedented microbial diversity. Novel gene families with representatives only within such metagenomic datasets represent a large proportion of the ocean metagenome. The presence of so many new gene families from these uncultured and highly diverse microbial populations represents a challenge for the understanding of and exploitation of the biology and biochemistry of the ocean environment. The application of new metagenomic and single cell genomics tools offers new ways to explore the complete metabolic diversity of the marine biome.

Keywords: metagenome; marine; microorganism; enzyme

1. Introduction

The marine environment is by far the largest habitat on Earth, with oceans covering ~70% of the earth's surface. Within this environment, habitats range from tropical sunlit surface waters to ocean trenches 11,000 m deep with pressures exceeding 100 MPa. While salinity is fairly uniform at 3.5%, temperatures range from sea ice in the polar regions to temperature in excess of 100 °C at deep sea hydrothermal vents. Microorganisms are able to survive and grow throughout this environment, either surviving as free-living organisms or in association with other organisms. The marine environment is dominated by microorganisms from the three domains of life, namely Archaea, Bacteria and Eukarya.

Microorganisms are believed to be responsible for up to 98% of marine primary productivity [1], playing key roles in marine food webs and in carbon and energy cycles. Microorganisms dominate the oceans in terms of total biomass and metabolic activity; with pelagic photosynthetic microorganisms producing the majority of organic matter present within the oceans, while at the other end of the food web, sediment microorganisms degrade dead marine macro organisms and marine snow. The different environmental challenges within the marine environment have led to different microbes adopting different survival and growth strategies. For example, the photosynthetic pelagic bacterium *Prochlorococcus marinus* MED4 has adopted a minimalist approach, with conditions within its environmental niche varying little, the bacterium has reduced its genome size to 1.66 Mbp to gain a competitive advantage [2]; while *Pelagibacter ubique* from the SAR11 clade of marine alphaproteobacteria has no transposons, extrachromosomal elements, pseudogenes or introns and constitutes the smallest genome known for a free-living microorganism, encoding the smallest number of predicted open reading frames [3]. Adaptation to the challenges of the deep sea, where nutrient sources, e.g., from marine detritus, can be unpredictable; has led to greater adaptability in both signaling pathways involved in the detection of nutrients and greater metabolic capabilities in utilizing them. Likewise, microorganisms from the extremes within the marine environment have adapted to the extremes of temperature and pressure, resulting in novel biochemistry which is able to operate at high or low temperatures and pressures, perhaps uniquely suited for many industrial processes.

Life in the deep-sea can also occur entirely independently of sunlight driven photosynthesis. At hydrothermal vents, for example, chemolithoautotrophic bacteria are the primary producers for these unusual ecosystems, with archaea isolated from these environments able to grow at 121 °C. Photosynthetic bacteria entirely dependent on the geothermal radiation from these vents have also been isolated [4]. Cold seeps, in which hydrocarbon rich fluids seep from the ocean floor, are also hotspots of diversity, in which the primary producers are chemosynthetic microorganisms [5]. These discoveries at the extremes of ocean life have led to a significant extension of the known biosphere.

Thus, the ability of microbial life to thrive throughout the entire marine environment, from deep ocean vents to surface sea ice, is entirely due to the diverse and adaptable biochemistry encoded within their genetic resources. In this review, we highlight recent discoveries in the study of enzymes from marine microbial sources, and advances in the technologies available to study the vast genetic resources of marine microbiota.

2. Molecular Approaches to Study Marine Microbial Biodiversity

It is now widely accepted that <1% of microbes can be isolated using traditional culturing techniques, leaving the vast majority of microorganisms and their biochemical pathways inaccessible. The emergence of culture-independent and metagenomic techniques has however provided us with the tools to determine the full extent of the uncultured microbial diversity and allowed access to the biochemical pathways within these uncultured microorganisms. Metagenomics is defined as the study of the pooled genetic complement of a given environmental sample, and analyses can be either sequence driven or function driven [6]. To date, sequence-based metagenomic analyses of marine microbiota have attempted to describe; ‘who is there?’; ‘what are they doing?’; ‘who is doing what?’ and ‘what evolutionary processes determine these parameters?’ Recent conceptual and technological advances have allowed the increased use of sequence guided metagenomic investigations of the marine environment. Recent advances in high-throughput sequencing technology and lower cost sequencing technologies have made random shotgun sequencing of environmental DNA economically feasible. The majority of metagenomic investigations to date have employed whole metagenome shotgun sequencing approaches for the cloning and sequencing of microbial DNA from marine environments. This involves the generation of small-insert DNA clone libraries, and their subsequent analysis using Sanger dideoxy sequencing; generating sequences that can then be used to query databases, whereby phylogeny can be inferred or putative functional genes identified. Where putative protein encoding sequences are detected proximal to phylogenetic marker genes on a single cloned fragment, function can be linked to taxonomy. This approach can give read lengths ranging from 600 to 900 bp in length, which can be extended through the entire fosmid clone. The contigs generated using this approach give a fuller context for any given gene detected, thus giving a greater probability of reconstructing the metabolic pathways of individual members within a particular marine microbial consortium.

454 pyrosequencing, involving a novel sequencing-by-synthesis technology, is also now gaining popularity as a cost effective methodology that can be employed for metagenomic analysis of microbial communities. While the benefits of the technology both from a sequence data generation and cost standpoint (25 to 40 Mb of DNA sequence per run at an accuracy of ~98% [7]) are obvious, the limitations of lower read lengths of between 200 to 300 bp can be problematic [8]. Read lengths achievable using pyrosequencing are currently around 450 bp, and this is likely to continue to improve, which should improve the amount of useful data generated from pyrosequencing of metagenomic samples.

The shorter contigs generated with this method, often do not enable the assignment of function, and reconstruction of the metabolic pathways present in individual microbes is only possible for the simplest of consortia. However notwithstanding this, these random shotgun sequencing and 454 pyrosequencing approaches require no prior sequence knowledge for data generation, and thus there is a good possibility that novel gene sequences can be discovered.

The assignment of taxonomic status to these sequences requires the use of recently developed software such as Metagenomic And rDNA Taxonomy Assigner (MARTA), or Genomic Signature based Taxonomic Classifier GSTaxClassifier. MARTA, which is available at <http://bergelson.uchicago.edu/software/marta>, allows the classification of taxonomic status using percent-identity and thresholds; employing DNA sequence data from 5,000 typed bacterial species,

which allows results generated with MARTA to be compared to those from the phylogenetic algorithm CARMA and the ribosomal database project RDP-II Classifier [9]. GSTaxClassifier, which is available at <http://helix2.biotech.ufl.edu:26878/metagenomics/>, is a program that allows the evaluation of genomic signatures between metagenomic query sequences and reference genome databases [10]. The assignment of function typically requires profile-to-sequence searches, whereby predicted protein sequences are compared to sequence alignment profiles from the protein families PFAM [11], COGs [12] and TIGRFAM [13], using RPS-BLAST [14]. PFAMs facilitates both the identification and annotation of protein domains, TIGRFAM includes models for both domain and full-length proteins, while COGs also allows the annotation of full-length proteins. InterPro, an integrated protein database, is also widely used for the classification and annotation of proteins and is hosted at <http://www.ebi.ac.uk/interpro/>. Thus, a number of bioinformatic tools now exist for data mining of metagenomic sequences, based not only on primary sequence homology but also on predicted protein structures. It can therefore be anticipated in the near future that searches of sequence-based metagenome databases in combination with bioinformatic tools, some of which will be discussed later in this review, will be equally as powerful an approach as functional-based methods in the mining of novel genes encoding enzymes from marine microbes.

2.1. Who is there?

Diversity and the abundance of free-living marine microorganisms have been described using 16S rRNA based analyses and total metagenomic analyses from a variety of different marine environments such as in ocean surface waters [15,16], mesopelagic waters [17], the deep sea [1,18,19], water columns [20] and sea subfloor sediments [21]. The microbial diversity of marine invertebrate-associated organisms has also been investigated including marine corals [22,23] and marine sponges [24–26]. A brief summary of the results obtained will now be described.

In studies on surface waters of the Sargasso Sea [16], different phylogenetic biomarker genes were used to elucidate the composition of the bacterial community, resulting in the identification of nine major bacterial phyla (*Proteobacteria*, *Actinobacteria*, *Cyanobacteria*, *Firmicutes*, *Bacteroidetes*, *Chloroflexi*, *Spirochaetes*, *Fusobacteria* and *Deinococcus-Thermus*) and two archaeal phyla (*Crenarchaeota* and *Euryarchaeota*). Over 1800 bacterial species, including 148 new phylotypes, were detected, with the vast majority of all sequences clustering within the α - and γ -divisions of the phylum *Proteobacteria*.

In studies on mesopelagic waters involving analyses of the Bermuda Atlantic Time-Series (BATS), the *SAR202* clade of Green Non Sulphur (GNS) bacteria were shown to be prevalent at the interface of the photic and euphotic zones and rare in the immediate sub-surface [17]. Studies on the deep ocean, involving analysis of a metagenomic fosmid library created from seawater sampled at a depth of 3,000 m in the Ionian Sea, revealed sequences related to *Acidobacteria*, *Chloroflexi*, *Planctomycetes* and α - and γ -*Proteobacteria*, while 28 archaeal sequences related to *Crenarchaeota symbosium* were obtained [18]. Other work investigating a 16S clone library from a depth of 500 m in the Antarctic Ocean revealed a novel group of δ -*Proteobacteria* with low homology to previously known organisms [27].

Studies on the microbial diversity in depth specific profiles of water columns have also been undertaken, such as the comparative analysis of datasets from the Greenland Sea, the Ionian Sea and the Sargasso Sea surface water (5 m) [19]. This study identified that—despite differences between sample sites: Sargasso Sea (5 m), Greenland Sea 50m (cold) and Ionian Sea 50 m (warm), and differences in the methodologies used—clear vertically stratified community differences were evident and certain similarities between samples were obvious, in particular the widespread presence of DNA sequences that shared homology with genes from *Pelagibacter ubique*. Studies on the microbial community structure from sub-seafloor sediments in the Peru Margin have reported a consortium dominated by *Crenarchaeota*, with *Euryarchaeota* and *Chloroflexi* also abundant [21].

Microbial communities form specific intimate and incidental associations with many marine invertebrates and the microbiota associated with several sponge species in particular have been widely investigated [24,25,28–31]; particularly since microbial endosymbionts of marine sponges are thought to be a rich source of marine natural products [32]. A good recent example is the study from the Wagner group, which employed 16S rRNA tag pyrosequencing to investigate the bacterial diversity associated with three sponge species and the surrounding seawater. Over 250,000 sequences were generated, which aligned to 23 bacterial phyla, demonstrating a much higher diversity than had previously been reported [26]. Metagenomic analyses have also been reported from corals [22,23,33], with fungal sequences accounting for 38% of the classified sequences, and the remainder of the identified sequences aligning to bacteria (7%), archaea (1%), eukaryotic viruses (2%) and phage (3%). The bacterial sequences identified were most similar to *Proteobacteria*, *Firmicutes*, *Cyanobacteria* and *Actinobacteria*.

2.2. Resources for marine metagenomic analysis

With the advent of the “omics” era, more and more metagenomic and metatranscriptomic datasets are being generated from marine environments. Analysis resources and tools have been developed in an attempt to maximize the analysis of data that can be obtained from these vast datasets. For example, for marine-derived data, a dedicated web based resource, the Community CyberInfrastructure for Advanced Marine Microbial Ecology Research and Analysis (CAMERA <http://camera.calit2.net/>), has been developed [34]. The CAMERA database hosts marine metagenomic datasets and completed marine microbial genome sequences as well as sample metadata such as sample site GPS location, depth, temperature, salinity, *etc.* Additional tools available for metagenome analysis include: the Integrated Microbial Genome/Microbiome (IMG/M) <http://img.jgi.doe.gov/cgi-bin/m/main.cgi> [35], which provides tools for analyzing the functional capability of microbial communities based on their metagenome sequence by allowing comparative analysis of genome and metagenome sequences available on the JGI-Integrated Microbial Genomes system; MetaBioME [36], which is a resource specifically aimed at novel enzyme discovery from metagenomic data; and Orphelia, which is a tool for discovering ORFs in short metagenomic sequences of unknown phylogeny [37]. Recently, the Metagenomics RAST server hosted at <http://metagenomics.nmpdr.org/> has been launched, which provides a fully-automated service for annotating metagenome samples. Not only does it provide annotation of sequence fragments, but it also provides a means for comparing phylogenetic classifications and metabolic reconstructions of metagenomes [38]. Other resources include JANE

hosted at <http://jane.bioapps.biozentrum.uni-wuerzburg.de>, which facilitates the mapping of prokaryotic Expressed Sequence Tags (ESTs) or variable sequence reads from DNA sequence and transcriptomic data to related template genomes [39]. WebCARMA hosted at <http://webcarma.cebitec.uni-bielefeld.de> has recently been developed to facilitate the functional and taxonomic classification of unassembled metagenomic reads as short as 35 bp [40]. MEGAN, a standalone downloadable computer programme for the analysis of metagenomic data [41], has also recently been extended to allow the comparative analysis of multiple datasets, specifically allowing visual and statistical comparison of metagenomes [42]; and can be obtained from <http://www-ab.informatik.uni-tuebingen.de/software/megan>. As the major drawback to metagenomic data analysis is the lack of reference sequences and genomes, the increased data and tools available should gradually minimize this problem. Also, the inclusion of comprehensive sampling metadata will assist in contextualizing sequence data for improved inferences. Thus, the rapidly developing field of sequence based metagenomic analysis of marine microbiota promises much for the future of marine biotechnology.

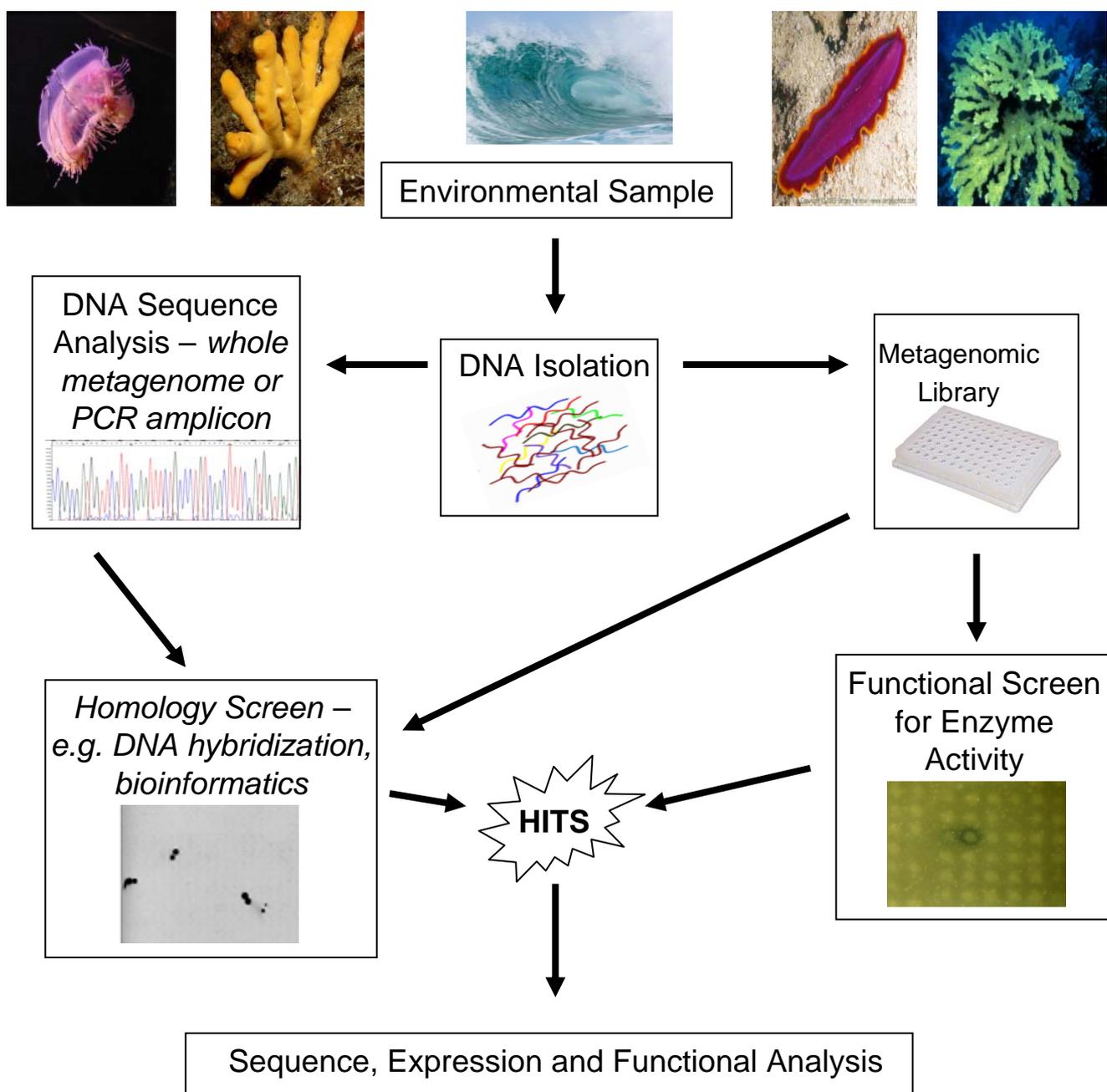
3. Functional Metagenomic Based Approaches

While sequence based metagenomic approaches rely on comparing sequence data, which is obtained with sequences deposited in databases, functional metagenomics focuses on screening DNA library clones directly for a phenotype, *i.e.*, the genes are recognizable by their function rather than by their sequence. The power of such an approach is that it does not require the genes of interest to be recognizable by sequence analysis, ensuring that this approach has the potential to directly identify entirely new classes of genes for both known and indeed novel functions [6,43]. Another advantage of such an approach is that while sequence based approaches can result in the incorrect annotation of sequences with weak similarities to biochemically characterized gene products or of sequences similar to gene products with multiple functions, the results from a functional metagenomics approach are unambiguous.

The basic strategy for functional metagenomics is outlined in Figure 1. An environmental sample, e.g., seawater or marine invertebrate is collected and then total community DNA is extracted from the sample. The isolated DNA is then used to generate a metagenomic library using a suitable cloning vector. This library is then transferred to a suitable host strain, usually *Escherichia coli*, and individual clones can then be screened for the presence of enzymatic or other bioactivities encoded by the environmental DNA fragment. When coupled with a robust and high-throughput screen, this method is an extremely effective way of isolating novel enzymes from otherwise inaccessible microbes. Examples of enzymes isolated from marine sources using a functional metagenomics approach are outlined in Table 1; these include esterases, lipases, chitinases, amylases and amidases. While the functional based approach has been successfully applied to terrestrial environments, especially soil, with the discovery of new genes for antibiotics, antibiotic resistance and industrial enzymes [44–46], the full potential of marine functional metagenomics has to date remained largely unrealized. It is estimated that about 1×10^6 bacteria are present in 1 mL of seawater, with far greater concentrations in marine snow [47–49]. However, only about 0.0001–0.1% are believed to be culturable [50], making a culture independent approach to harvest the metabolic potential of these organisms even more

attractive. In addition, data from the Global Ocean Sampling (GOS) expedition indicates that despite current large-scale sequencing efforts the rate of discovery of new protein families from the marine environment is linear, implying that marine microorganisms will continue to be a source of novel enzymes in the foreseeable future [51]. As many of these gene products are entirely novel, their activity cannot be inferred from comparison to known protein databases, thus a functional metagenomics approach has the ability to identify novel genes on the basis of phenotypes which lend themselves to high throughput screens.

Figure 1. Enzyme discovery from metagenomes: functional and sequence-based approaches.



As can be seen in Table 1, functional metagenomic based approaches have led to the discovery of many new enzymes from marine sources. However, despite the undoubted promise of functional

metagenomics for the discovery of new enzymes, the approach is limited to some extent by the ability of metagenomic clones to produce active enzymes. Many functional metagenomic approaches rely on the use of *E. coli* as a host for the expression of metagenome encoded proteins. While quite a large number of genes derived from *Enterobacteriaceae* will be readily expressed in the most common *E. coli* host, many genes from more distantly related organisms may not be expressed due to the promoter regions of these genes not being recognized by the *E. coli* transcriptional machinery or be expressed at low levels due to differences in codon usage, for example. Even where transcription and translation of foreign genes results in efficient protein expression, additional problems can arise when proteins need to be post-translationally modified or exported for activity. For these reasons, the availability of suitable heterologous expression hosts remains a barrier to extracting the maximum information from functional metagenomic analyses [6].

On the other hand, due to advanced screening methodologies and the use of robotic instrumentation, it is now possible to screen large clone libraries for functional activities in a high throughput fashion, in relatively short timescales [52]. The development of new hosts for heterologous expression will also contribute to advances in function-based metagenomic approaches. For example, *Streptomyces lividans*, *Pseudomonas putida* and *Rhizobium leguminosarum* have all been assessed for their potential as expression hosts for metagenomic libraries [53–55]. Other approaches have looked at ways of increasing the levels of functional expression of foreign genes in *E. coli*. Issues such as restricted expression due to codon usage have been addressed by supplementing *E. coli* with plasmids containing additional tRNA genes [56]. Inadequate folding of target proteins can be addressed through co-expression of chaperone proteins [56]. Indeed, the expression of the chaperonin 60 gene and the cochaperonin 10 gene from the psychrophilic bacterium *Oleispira antarctica* RB8T can effectively lower the minimum temperature for growth of *E. coli*, allowing the functional expression of cold-adapted proteins [57].

Coupling classical functional screens of metagenomic libraries with innovative approaches such as substrate-induced gene expression screening (SIGEX), pre-amplification inverse-PCR (PAI-PCR) and metagenomic DNA shuffling can also further improve function-based metagenomics. SIGEX was initially developed to detect metagenomic clones that express desired catabolic genes induced by the presence of suitable substrates using green fluorescent protein (GFP) as a reporter gene [58]. Constitutively expressing clones are first removed by fluorescence activated cell sorting (FACS), then the expression of desired genes is induced by the addition of substrates and positive clones are again sorted by FACS, separated on agar plates and characterized [59]. With SIGEX, genes possibly involved in benzoate and catechol degradation have been discovered in a groundwater metagenomic library [58]. Despite the limitations of SIGEX, which can only detect genes that are correctly induced in a heterologous host, the method is extremely high throughput since large number of clones can be screened in relatively short timescales [60].

Inverse PCR (I-PCR) is a method that can be used for cloning the upstream and downstream flanking regions of known sequences [61], and has been employed in the amplification of known gene families. However, I-PCR has found little utility when applied to metagenomes, presumably due to the potential increased complexity and low copy numbers of target sequences. Pre-amplification of the template DNA, using phi29 DNA polymerase coupled with inverse PCR (PAI-PCR), has, however, been shown to facilitate the isolation of novel genes from small amounts of environmental DNA [61].

Using PAI-PCR, DNA glycosyl hydrolase genes have been identified, with a 10^5 -fold increase in sensitivity compared to I-PCR without pre-amplification [61]. Thus, this approach could also be useful for marine DNA-samples where the amount of extracted DNA is low. Finally, metagenomic DNA shuffling for its part simulates lateral gene transfer and gene rearrangements in microbial genomes. This approach has been used to generate novel lindane degrading genes in terrestrial systems [62]. The advantages of this method are that the diversity of the metagenome can actually be increased, and by applying stringent selection steps enzymes particularly suited to industrial applications can be generated. Thus, such an approach could also be useful to find novel enzyme encoding genes in the marine environment.

Table 1. Marine enzymes discovered from Microbial and Metagenomic sources.

Activity	Source	Habitat	Reference
Esterase	Metagenome	Deep-sea sediment	[97]
	Metagenome	Deep-sea basin	[72]
	Metagenome	Surface seawater	[69]
	Metagenome	Arctic sediment	[98]
	<i>Vibrio</i> sp.	Sea Hare Eggs	[99]
	<i>Pseudoalteromonas haloplanktis</i>	Antarctic Seawater	[71]
Lipase	Metagenome	Tidal Flat	[74]
	Metagenome	Deep Sea sediment	[75]
	Metagenome	Baltic Sea sediment	[67]
	<i>Pseudoalteromonas haloplanktis</i> TAC125	Antarctic Seawater	[73]
	<i>Aureobasidium pullulans</i> HN2.3	Sea saltern	[100]
Cellulase	<i>Pseudoalteromonas</i> sp. DY3	Deep-sea sediment	[77]
	<i>Pseudoalteromonas haloplanktis</i>	Antarctic Seawater	[78]
	<i>Teredinibacter turnerae</i> T7902T	Shipworm	[79]
	<i>Marinobacter</i> sp. MSI032.	Marine sponge	[76]
Chitinase	Metagenome	Estuary	[80]
	<i>Arthrobacter</i> sp. TAD20	Antarctic ice	[101]
	<i>Rhodothermus marinus</i>	Marine hot spring	[81]
Amidase	Metagenome	Marine sediments / sludges	[102]
Amylase	<i>Aureobasidium pullulans</i> N13d	Deep-sea sediment	[103]
	Metagenome	Deep sea hydrothermal vent	[66]
Phytase	<i>Kodomaea ohmeri</i> BG3	Fish gut	[104]
Protease	<i>Pseudomonas</i> strain DYA	Deep-sea sediment	[105]
	Marine bacterium	Antarctic Seawater	[106]
	<i>Aeropyrum pernix</i> K1	Coastal solfataric vent	[107]
Alkane hydroxylase	Metagenome	Hydrocarbon seep	[82]
	Metagenome	Deep sea sediment	[68]
Xylanase	<i>Pseudoalteromonas haloplanktis</i>	Antarctic Seawater	[108]

4. Marine Microbes and Their Potential for Functional Metagenomics

4.1. Biotechnological uses of marine enzymes

The potential for novel biochemistry in the marine environment has not been overlooked, especially in the field of specialty and diagnostic enzymes. The luminescent properties of the jellyfish *Aequorea victoria* led to the characterization of the green fluorescent protein (GFP), which now has widespread uses in molecular biology as a reporter protein [63], and has been utilized in studies involving bacteria, fungi, fish and mammalian cells. Similarly, the protein aequorin, also from the *Aequorea* jellyfish, has found use as a biosensor for Ca^{2+} signaling as it emits a blue light in the presence of Ca^{2+} ions [64]. Microbial light emitting systems have also been discovered, the most well known of which is the bioluminescent bacterium *Vibrio fischeri*, which is found within squid light organs. Luciferase, the enzyme responsible for this light emission, has also found widespread use as a reporter system [65].

For research and other purposes, enzymes isolated from microbial sources at the extremes of temperature have proven to be particularly useful. DNA polymerases isolated from thermophilic marine eubacteria (Tth polymerase) and archaea (VENT® polymerase) have proved useful alternatives to Taq DNA polymerase, while thermolabile phosphatases have also found utility in the research laboratory. In the bulk enzyme market, a heat and acid stable α -amylase, Valley Ultra-Thin, discovered from deep sea hydrothermal vent archaea, has been developed by Diversa/Verenium to facilitate the processing of corn into ethanol [66].

4.2. Novel enzyme discoveries

The potential of marine microorganisms for the discovery of new enzymes has led to numerous new discoveries in recent years (Table 1). These discoveries have been made from both microorganisms that have been isolated from the marine environment and from metagenomic libraries. The spectrum of marine-derived genes coding for enzymes with unique abilities includes lipases [67], cellulases, proteases, alkane hydroxylase genes [68], esterases with high tolerances for salt and organic solvents [69], and metalloproteases with a high temperature optima [70].

Esterases have potential commercial value for use in industrial biotransformations. Cold active esterases have been isolated from *P. haloplanktis* [71], and from metagenomic libraries, esterases with preferences for high pressure and salt [72] and with high tolerance for organic solvents [69] have been discovered. Lipases have wide ranging applications in the food, detergent, and pharmaceutical industries, and marine microorganisms have proven to be a source of novel lipases with new families of lipase [73,74] and cold active lipases [75] discovered from isolated microorganisms and metagenomes. Cellulases, a class of enzyme under intense study for their role in the generation of biofuels from renewable cellulosic substrates, have been discovered from various marine microorganisms including an alkaline cellulase, stable to pH 12, from a bacteria isolated from a sponge [76], cold active cellulases [77,78] and cellulases from bacteria isolated from a shipworm [79]. A shipworm metagenome project is also currently being pursued by the U.S. Joint Genome Institute to better understand the role of the shipworm microbiota in the degradation of wood. Chitin is also a major marine carbon source and chitinases have been discovered in metagenomic libraries [80] and a thermostable chitinase has been purified from a marine Bacterium [81]. Hydrocarbon cold seeps and

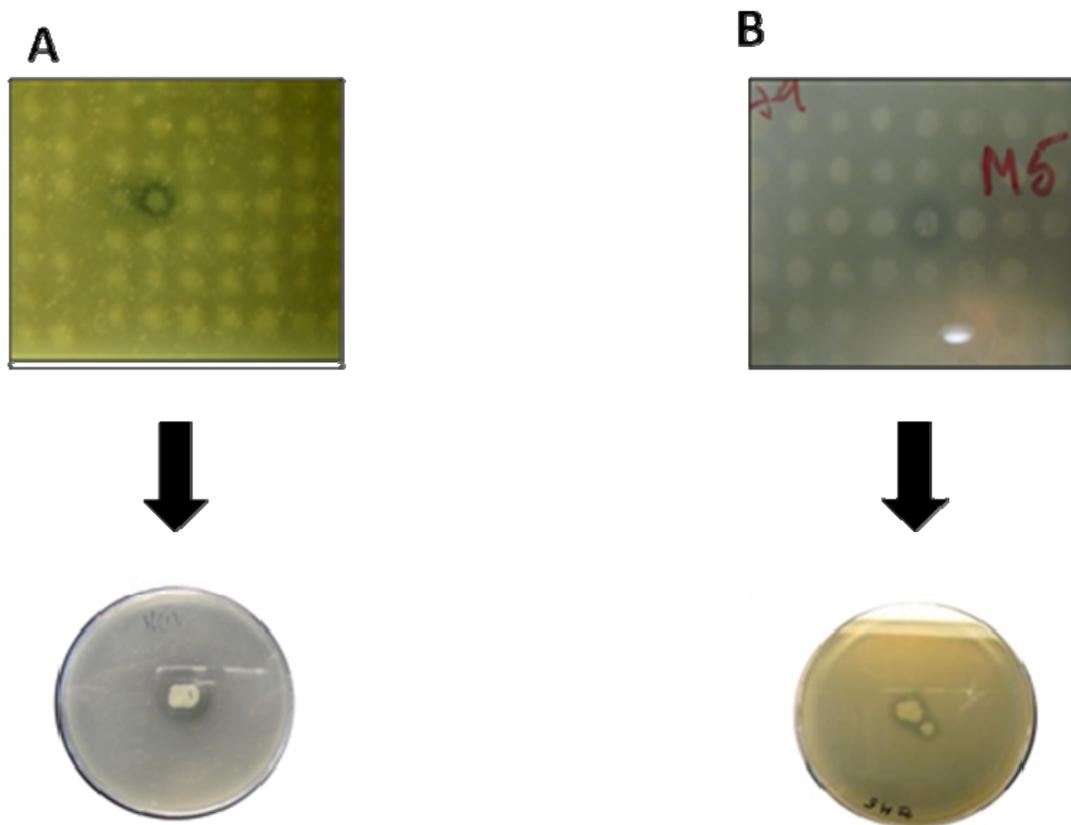
oil spills have led to the adaptation of certain marine microbiota to utilize alkanes as carbon sources. Such bacteria and their enzymes have a potential role in bioremediation and oil processing. Alkane hydroxylase genes encoding enzymes with specificity for alkanes of varying lengths have been isolated from deep sea and cold seep metagenomic libraries [68,82].

4.3. Phenotypic screens

Robust and sensitive screens are essential for the discovery of novel biocatalysts. As previously mentioned, sequence-based approaches rely on homologies to known protein families; a drawback that does not apply to functional screens [83]. The most commonly reported type of functional screen for the discovery of enzymes from metagenomic libraries is based on colony phenotype changes on agar plates. These assays typically rely on the utilization of a substrate resulting in the appearance of a zone of clearing around positive clones.

For example, our group has detected both protease and lipase activity in clones from a *Haliclona simulans* sponge metagenomic library by replicating the library onto Luria-Bertani containing skimmed milk [84], or in the case of lipase activity, supplemented with tributyrin [85]. Clear halos around these colonies indicate either protease or lipase activity in each case (Figure 2). Chromogenic substrate analogs can also be employed to improve the sensitivity of these assays, with positive clones causing a loss or appearance of color. One major advantage of using agar plate assays is that they are highly scalable and with integrated robotics many thousands of clones can be screened per day. A major disadvantage, however, is that agar plate assays often require the active transport of either the substrate or the enzyme to detect activity, and it is widely believed that many potentially novel enzymes are in many cases simply not detected as they fail to come in contact with the substrate. To circumvent this problem, cell extracts have been successfully employed to identify novel genes involved in aromatic compound degradation [46]. Additional screening approaches that can be employed include phenotypic screens where the introduction of a new catabolic pathway into a host strain allows it to subsequently grow on a previously inaccessible substrate. A variation on this approach is the use of an engineered host that lacks a particular activity, and the reintroduction of this activity via a metagenomic clone, thereby providing a straightforward method by which to screen very large numbers of clones in a metagenomic library for novel activities. In addition, the utilization of engineered host strains can allow screens to pass beyond the simple growth / no growth assay based systems. For example, engineered strains in which the induction of a specific pathway is linked to a reporter gene can be used to isolate metagenomic clones that induce particular cellular pathways. This strategy has been used to identify clones affecting signal transduction from a moth gut metagenome [86]. Thus, the use of sensitive biosensor host strains such as these provides much promise for the discovery of genes encoding novel properties from marine metagenomic libraries.

Figure 2. Examples of positive lipase (A) and protease activities (B) from a *Haliclona simulans* metagenomic library.



5. Novel Approaches

New approaches, which are likely to have a major impact on the discovery of novel enzymes from marine microorganisms that are difficult to culture, include single cell genomics, metatranscriptomics and metaproteomics. While the metagenomics approaches described above are useful for exploiting the biochemistry of microbial communities, they are however unable to access the metabolic capabilities associated with specific microorganisms within the consortia; given that they largely rely more on a “take all” approach. While functional metagenomic based approaches have proven to be successful for the discovery of novel enzymes, it is quite likely that many novel enzymes from rare microbes in complex communities are poorly represented in metagenomic libraries. One solution to this problem is the use of multiple displacement amplification (MDA) for whole genome amplification from single cells, which will allow the study of the entire biochemical potential of single uncultured microbes from complex microbial communities [87,88]. In this approach, microbial cells are first sorted using FACS and collected. Then, MDA is used to amplify the entire genome from the sorted cells. The amplified genomic DNA can then be analyzed to infer phylogeny, from 16S rRNA sequence, and used as a template for anchor-based analysis such as PCR, to isolate known families of genes, or can be directly sequenced to access the full metabolic potential of the microorganism. This approach has great potential for the discovery of novel enzymes from marine microbes, as it affords easier access to rare microbiota and should also enable the matching of phylogenetically characterized

amplified genomes to suitable expression hosts, thereby increasing the chances of discovering novel biochemistry.

While the standard metagenomic approach of accessing the entire genetic material associated with an environment may well indicate what genes are present within a marine environment, it does not, however, offer clear evidence regarding which genes are active within this environment. To overcome this particular problem, metatranscriptomic based approaches have been employed to study marine microbial populations, in which only transcriptionally active genes are accessed [89]. Applying this approach to the study of ocean waters has resulted in the detection of genes known to be involved in key aspects of the metabolism of open ocean microbial species, and in addition many unique transcripts, presumably also of biochemical importance have been detected [90,91]. Metatranscriptomic analyses can be used to determine the enzymes in an environment that are actively involved in a particular biochemical pathway [92]. An alternative strategy to accessing the metagenome is to directly analyze the proteins within the marine environment using metaproteomics, for review see [93,94]. This has proven useful in providing new physiological data such as the identification of proteins that are essential for stress and starvation resistance in culturable marine bacteria. A recent terrestrial based approach has used the binding properties of specific proteins to isolate them from a particular environment, for example, bacterial cellulose-binding proteins from sheep rumen extracts were isolated and identified in this way [95]. Similarly, an array of dye-linked substrates has recently been used to identify enzymes with specific activities from microbial extracts and from metagenomic libraries [96]. It is clear that by applying novel screening tools such as these to samples from marine ecosystems that novel and metabolically relevant enzymes can be discovered.

Acknowledgements

This project was supported by the Beaufort Marine Research Award under the Sea Change Strategy and the Strategy for Science Technology and Innovation (2006–2013), with the support of the Marine Institute, funded under the Marine Research Sub-Programme of the National Development Plan 2007–2013. DPHL is in receipt of Marie Curie Transfer of Knowledge Host Fellowships; [grant no. MTKD-CT-2006-042062].

References and Notes

1. Sogin, M.L.; Morrison, H.G.; Huber, J.A.; Mark Welch, D.; Huse, S.M.; Neal, P.R.; Arrieta, J.M.; Herndl, G.J. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 12115–12120.
2. Dufresne, A.; Garczarek, L.; Partensky, F. Accelerated evolution associated with genome reduction in a free-living prokaryote. *Genome Biol.* **2005**, *6*, R14.
3. Giovannoni, S.J.; Tripp, H.J.; Givan, S.; Podar, M.; Vergin, K.L.; Baptista, D.; Bibbs, L.; Eads, J.; Richardson, T.H.; Noordewier, M.; Rappe, M.S.; Short, J.M.; Carrington, J.C.; Mathur, E.J. Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **2005**, *309*, 1242–1245.
4. Beatty, J.T.; Overmann, J.; Lince, M.T.; Manske, A.K.; Lang, A.S.; Blankenship, R.E.; Van Dover, C.L.; Martinson, T.A.; Plumley, F.G. An obligately photosynthetic bacterial anaerobe from a deep-sea hydrothermal vent. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9306–9310.

5. Jorgensen, B.B.; Boetius, A. Feast and famine--microbial life in the deep-sea bed. *Nat. Rev. Microbiol.* **2007**, *5*, 770–781.
6. Handelsman, J. Metagenomics: application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 669–685.
7. Edwards, R.A.; Rodriguez-Brito, B.; Wegley, L.; Haynes, M.; Breitbart, M.; Peterson, D.M.; Saar, M.O.; Alexander, S.; Alexander, E.C., Jr.; Rohwer, F. Using pyrosequencing to shed light on deep mine microbial ecology. *BMC Genomics* **2006**, *7*, 57.
8. Wommack, K.E.; Bhavsar, J.; Ravel, J. Metagenomics: read length matters. *Appl. Environ. Microbiol.* **2008**, *74*, 1453–1463.
9. Horton, M.; Bodenhausen, N.; Bergelson, J. MARTA: A suite of Java-based tools for assigning taxonomic status to DNA sequences. *Bioinformatics* **2009**, *6*, 568–569.
10. Yu, F.; Sun, Y.; Liu, L.; Farmerie, W. GSTaxClassifier: a genomic signature based taxonomic classifier for metagenomic data analysis. *Bioinformatics* **2010**, *4*, 46–49.
11. Finn, R.D.; Tate, J.; Mistry, J.; Coghill, P.C.; Sammut, S.J.; Hotz, H.R.; Ceric, G.; Forslund, K.; Eddy, S.R.; Sonnhammer, E.L.; Bateman, A. The Pfam protein families database. *Nucleic Acids Res.* **2008**, *36*, D281–D288.
12. Tatusov, R.L.; Koonin, E.V.; Lipman, D.J. A genomic perspective on protein families. *Science* **1997**, *278*, 631–637.
13. Selengut, J.D.; Haft, D.H.; Davidsen, T.; Ganapathy, A.; Gwinn-Giglio, M.; Nelson, W.C.; Richter, A.R.; White, O. TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. *Nucleic Acids Res.* **2007**, *35*, D260–D264.
14. Markowitz, V.M.; Ivanova, N.; Palaniappan, K.; Szeto, E.; Korzeniewski, F.; Lykidis, A.; Anderson, I.; Mavromatis, K.; Kunin, V.; Garcia Martin, H.; Dubchak, I.; Hugenholtz, P.; Kyrpides, N.C. An experimental metagenome data management and analysis system. *Bioinformatics* **2006**, *22*, e359–e367.
15. Rusch, D.B.; Halpern, A.L.; Sutton, G.; Heidelberg, K.B.; Williamson, S.; Yooseph, S.; Wu, D.; Eisen, J.A.; Hoffman, J.M.; Remington, K.; Beeson, K.; Tran, B.; Smith, H.; Baden-Tillson, H.; Stewart, C.; Thorpe, J.; Freeman, J.; Andrews-Pfannkoch, C.; Venter, J.E.; Li, K.; Kravitz, S.; Heidelberg, J.F.; Utterback, T.; Rogers, Y.H.; Falcon, L.I.; Souza, V.; Bonilla-Rosso, G.; Eguarte, L.E.; Karl, D.M.; Sathyendranath, S.; Platt, T.; Bermingham, E.; Gallardo, V.; Tamayo-Castillo, G.; Ferrari, M.R.; Strausberg, R.L.; Nealson, K.; Friedman, R.; Frazier, M.; Venter, J.C. The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* **2007**, *5*, e77.
16. Venter, J.C.; Remington, K.; Heidelberg, J.F.; Halpern, A.L.; Rusch, D.; Eisen, J.A.; Wu, D.; Paulsen, I.; Nelson, K.E.; Nelson, W.; Fouts, D.E.; Levy, S.; Knap, A.H.; Lomas, M.W.; Nealson, K.; White, O.; Peterson, J.; Hoffman, J.; Parsons, R.; Baden-Tillson, H.; Pfannkoch, C.; Rogers, Y.H.; Smith, H.O. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **2004**, *304*, 66–74.
17. Giovannoni, S.J.; Rappe, M.S.; Vergin, K.L.; Adair, N.L. 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the Green Non-Sulfur bacteria. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 7979–7984.

18. Martin-Cuadrado, A.B.; Lopez-Garcia, P.; Alba, J.C.; Moreira, D.; Monticelli, L.; Strittmatter, A.; Gottschalk, G.; Rodriguez-Valera, F. Metagenomics of the deep Mediterranean, a warm bathypelagic habitat. *PLoS One* **2007**, *2*, e914.
19. Zaballos, M.; Lopez-Lopez, A.; Ovreas, L.; Bartual, S.G.; D'Auria, G.; Alba, J.C.; Legault, B.; Pushker, R.; Daae, F.L.; Rodriguez-Valera, F. Comparison of prokaryotic diversity at offshore oceanic locations reveals a different microbiota in the Mediterranean Sea. *FEMS Microbiol. Ecol.* **2006**, *56*, 389–405.
20. DeLong, E.F.; Preston, C.M.; Mincer, T.; Rich, V.; Hallam, S.J.; Frigaard, N.U.; Martinez, A.; Sullivan, M.B.; Edwards, R.; Brito, B.R.; Chisholm, S.W.; Karl, D.M. Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **2006**, *311*, 496–503.
21. Biddle, J.F.; Fitz-Gibbon, S.; Schuster, S.C.; Brenchley, J.E.; House, C.H. Metagenomic signatures of the Peru Margin subseafloor biosphere show a genetically distinct environment. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10583–10588.
22. Cooney, R.P.; Pantos, O.; Le Tissier, M.D.; Barer, M.R.; O'Donnell, A.G.; Bythell, J.C. Characterization of the bacterial consortium associated with black band disease in coral using molecular microbiological techniques. *Environ. Microbiol.* **2002**, *4*, 401–413.
23. Frias-Lopez, J.; Klaus, J.S.; Bonheyo, G.T.; Fouke, B.W. Bacterial community associated with black band disease in corals. *Appl. Environ. Microbiol.* **2004**, *70*, 5955–5962.
24. Kennedy, J.; Codling, C.E.; Jones, B.V.; Dobson, A.D.; Marchesi, J.R. Diversity of microbes associated with the marine sponge, *Haliclona simulans*, isolated from Irish waters and identification of polyketide synthase genes from the sponge metagenome. *Environ. Microbiol.* **2008**, *10*, 1888–1902.
25. Lafi, F.F.; Fuerst, J.A.; Fieseler, L.; Engels, C.; Goh, W.W.; Hentschel, U. Widespread distribution of poribacteria in demospongiae. *Appl. Environ. Microbiol.* **2009**, *75*, 5695–5699.
26. Webster, N.S.; Taylor, M.W.; Behnam, F.; Lucker, S.; Rattei, T.; Whalan, S.; Horn, M.; Wagner, M. Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ. Microbiol.* **2009**, 29 September, [Epub ahead of print].
27. Moreira, D.; Rodriguez-Valera, F.; Lopez-Garcia, P. Metagenomic analysis of mesopelagic Antarctic plankton reveals a novel deltaproteobacterial group. *Microbiology* **2006**, *152*, 505–517.
28. Fieseler, L.; Horn, M.; Wagner, M.; Hentschel, U. Discovery of the novel candidate phylum "Poribacteria" in marine sponges. *Appl. Environ. Microbiol.* **2004**, *70*, 3724–3732.
29. Holmes, B.; Blanch, H. Genus-specific associations of marine sponges with group 1 crenarchaeotes. *Mar. Biol.* **2007**, *150*, 759–771.
30. Mohamed, N.M.; Cicirelli, E.M.; Kan, J.; Chen, F.; Fuqua, C.; Hill, R.T. Diversity and quorum-sensing signal production of Proteobacteria associated with marine sponges. *Environ. Microbiol.* **2008**, *10*, 75–86.
31. Pape, T.; Hoffmann, F.; Queric, N.-V.; von Juterzenka, K.; Reitner, J.; Michaelis, W. Dense populations of Archaea associated with the demosponge *Tentorium semisuberites* Schmidt, 1870 from Arctic deep-waters. *Polar Biol.* **2006**, *29*, 662–667.
32. Cragg, G.M.; Grothaus, P.G.; Newman, D.J. Impact of natural products on developing new anti-cancer agents. *Chem. Rev.* **2009**, *109*, 3012–3043.

33. Wegley, L.; Edwards, R.; Rodriguez-Brito, B.; Liu, H.; Rohwer, F. Metagenomic analysis of the microbial community associated with the coral *Porites astreoides*. *Environ. Microbiol.* **2007**, *9*, 2707–2719.
34. Seshardi, R.; Kravitz, S.A.; Smarr, L.; Gilna, P.; Frazier, M. CAMERA: A Community Resource for Metagenomics. *PLoS Biol.* **2007**, *5*, e75.
35. Markowitz, V.M.; Ivanova, N.N.; Szeto, E.; Palaniappan, K.; Chu, K.; Dalevi, D.; Chen, I.M.; Grechkin, Y.; Dubchak, I.; Anderson, I.; Lykidis, A.; Mavromatis, K.; Hugenholtz, P.; Kyrpides, N.C. IMG/M: a data management and analysis system for metagenomes. *Nucleic Acids Res.* **2008**, *36*, D534–D538.
36. Sharma, V.K.; Kumar, N.; Prakash, T.; Taylor, T.D. MetaBioME: a database to explore commercially useful enzymes in metagenomic datasets. *Nucleic Acids Res.* **2009**, *38*, D468–D472.
37. Hoff, K.J.; Lingner, T.; Meinicke, P.; Tech, M. Orphelia: Predicting genes in metagenomic sequencing reads. *Nucleic Acids Res.* **2009**, *37*, W101–W105.
38. Meyer, F.; Paarmann, D.; D'Souza, M.; Olson, R.; Glass, E.M.; Kubal, M.; Paczian, T.; Rodriguez, A.; Stevens, R.; Wilke, A.; Wilkening, J.; Edwards, R.A. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinf.* **2008**, *9*, 386.
39. Liang, C.; Schmid, A.; Lopez-Sanchez, M.J.; Moya, A.; Gross, R.; Bernhardt, J.; Dandekar, T. JANE: efficient mapping of prokaryotic ESTs and variable length sequence reads on related template genomes. *BMC Bioinf.* **2009**, *10*, 391.
40. Gerlach, W.; Junemann, S.; Tille, F.; Goesmann, A.; Stoye, J. WebCARMA: a web application for the functional and taxonomic classification of unassembled metagenomic reads. *BMC Bioinf.* **2009**, *10*, 430.
41. Huson, D.H.; Auch, A.F.; Qi, J.; Schuster, S.C. MEGAN analysis of metagenomic data. *Genome Res.* **2007**, *17*, 377–386.
42. Mitra, S.; Klar, B.; Huson, D.H. Visual and statistical comparison of metagenomes. *Bioinformatics* **2009**, *25*, 1849–1855.
43. Kennedy, J.; Marchesi, J.R.; Dobson, A.D. Metagenomic approaches to exploit the biotechnological potential of the microbial consortia of marine sponges. *Appl. Microbiol. Biotechnol.* **2007**, *75*, 11–20.
44. D'Costa, V.M.; Griffiths, E.; Wright, G.D. Expanding the soil antibiotic resistome: exploring environmental diversity. *Curr. Opin. Microbiol.* **2007**, *10*, 481–489.
45. Lammle, K.; Zipper, H.; Breuer, M.; Hauer, B.; Buta, C.; Brunner, H.; Rupp, S. Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. *J. Biotechnol.* **2007**, *127*, 575–592.
46. Suenaga, H.; Ohnuki, T.; Miyazaki, K. Functional screening of a metagenomic library for genes involved in microbial degradation of aromatic compounds. *Environ. Microbiol.* **2007**, *9*, 2289–2297.
47. Azam, F. OCEANOGRAPHY: Microbial Control of Oceanic Carbon Flux: The Plot Thickens. *Science* **1998**, *280*, 694–696.
48. Azam, F.; Long, R.A. Sea snow microcosms. *Nature* **2001**, *414*, 497–498.

49. Whitman, W.B.; Coleman, D.C.; Wiebe, W.J. Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6578–6583.
50. Amann, R.L.; Ludwig, W.; Schleifer, K.H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **1995**, *59*, 143–169.
51. Yoosuf, S.; Sutton, G.; Rusch, D.B.; Halpern, A.L.; Williamson, S.J.; Remington, K.; Eisen, J.A.; Heidelberg, K.B.; Manning, G.; Li, W.; Jaroszewski, L.; Cieplak, P.; Miller, C.S.; Li, H.; Mashiyama, S.T.; Joachimiak, M.P.; van Belle, C.; Chandonia, J.M.; Soergel, D.A.; Zhai, Y.; Natarajan, K.; Lee, S.; Raphael, B.J.; Bafna, V.; Friedman, R.; Brenner, S.E.; Godzik, A.; Eisenberg, D.; Dixon, J.E.; Taylor, S.S.; Strausberg, R.L.; Frazier, M.; Venter, J.C. The Sorcerer II Global Ocean Sampling expedition: expanding the universe of protein families. *PLoS Biol.* **2007**, *5*, e16.
52. Kennedy, J.; Marchesi, J.R.; Dobson, A.D. Marine metagenomics: strategies for the discovery of novel enzymes with biotechnological applications from marine environments. *Microb. Cell Fact.* **2008**, *7*, 27.
53. Martinez, A.; Kolvek, S.J.; Yip, C.L.; Hopke, J.; Brown, K.A.; MacNeil, I.A.; Osburne, M.S. Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. *Appl. Environ. Microbiol.* **2004**, *70*, 2452–2463.
54. Wexler, M.; Bond, P.L.; Richardson, D.J.; Johnston, A.W. A wide host-range metagenomic library from a waste water treatment plant yields a novel alcohol/aldehyde dehydrogenase. *Environ. Microbiol.* **2005**, *7*, 1917–1926.
55. Wang, G.Y.; Graziani, E.; Waters, B.; Pan, W.; Li, X.; McDermott, J.; Meurer, G.; Saxena, G.; Andersen, R.J.; Davies, J. Novel natural products from soil DNA libraries in a streptomycete host. *Org. Lett.* **2000**, *2*, 2401–2404.
56. Makrides, S.C. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol. Rev.* **1996**, *60*, 512–538.
57. Ferrer, M.; Chernikova, T.N.; Yakimov, M.M.; Golyshin, P.N.; Timmis, K.N. Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nat. Biotechnol.* **2003**, *21*, 1266–1267.
58. Uchiyama, T.; Abe, T.; Ikemura, T.; Watanabe, K. Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. *Nat. Biotechnol.* **2005**, *23*, 88–93.
59. Yun, J.; Ryu, S. Screening for novel enzymes from metagenome and SIGEX, as a way to improve it. *Microb. Cell Fact.* **2005**, *4*, 8.
60. Uchiyama, T.; Watanabe, K. Substrate-induced gene expression (SIGEX) screening of metagenome libraries. *Nat. Protoc.* **2008**, *3*, 1202–1212.
61. Yamada, K.; Terahara, T.; Kurata, S.; Yokomaku, T.; Tsuneda, S.; Harayama, S. Retrieval of entire genes from environmental DNA by inverse PCR with pre-amplification of target genes using primers containing locked nucleic acids. *Environ. Microbiol.* **2008**, *10*, 978–987.
62. Boubakri, H.; Beuf, M.; Simonet, P.; Vogel, T.M. Development of metagenomic DNA shuffling for the construction of a xenobiotic gene. *Gene* **2006**, *375*, 87–94.
63. Chalfie, M.; Tu, Y.; Euskirchen, G.; Ward, W.W.; Prasher, D.C. Green fluorescent protein as a marker for gene expression. *Science* **1994**, *263*, 802–805.

64. Shimomura, O.; Johnson, F.H.; Saiga, Y. Microdetermination of Calcium by Aequorin Luminescence. *Science* **1963**, *140*, 1339–1340.
65. Engebrecht, J.; Simon, M.; Silverman, M. Measuring gene expression with light. *Science* **1985**, *227*, 1345–1347.
66. Mathur, E.J.; Toledo, G.; Green, B.D.; Podar, M.; Richardson, T.H.; Kulwiec, M.; Chang, H.W. A biodiversity-based approach to development of performance enzymes. *Ind. Biotechnol.* **2006**, *1*, 283–287.
67. Hardeman, F.; Sjolting, S. Metagenomic approach for the isolation of a novel low-temperature-active lipase from uncultured bacteria of marine sediment. *FEMS Microbiol. Ecol.* **2007**, *59*, 524–534.
68. Xu, M.; Xiao, X.; Wang, F. Isolation and characterization of alkane hydroxylases from a metagenomic library of Pacific deep-sea sediment. *Extremophiles* **2008**, *12*, 255–262.
69. Chu, X.; He, H.; Guo, C.; Sun, B. Identification of two novel esterases from a marine metagenomic library derived from South China Sea. *Appl. Microbiol. Biotechnol.* **2008**, *80*, 615–625.
70. Lee, D.G.; Jeon, J.H.; Jang, M.K.; Kim, N.Y.; Lee, J.H.; Lee, J.H.; Kim, S.J.; Kim, G.D.; Lee, S.H. Screening and characterization of a novel fibrinolytic metalloprotease from a metagenomic library. *Biotechnol. Lett.* **2007**, *29*, 465–472.
71. Aurilia, V.; Parracino, A.; D'Auria, S. Microbial carbohydrate esterases in cold adapted environments. *Gene* **2008**, *410*, 234–240.
72. Ferrer, M.; Golyshina, O.V.; Chernikova, T.N.; Khachane, A.N.; Martins Dos Santos, V.A.; Yakimov, M.M.; Timmis, K.N.; Golyshin, P.N. Microbial enzymes mined from the Urania deep-sea hypersaline anoxic basin. *Chem. Biol.* **2005**, *12*, 895–904.
73. de Pascale, D.; Cusano, A.M.; Autore, F.; Parrilli, E.; di Prisco, G.; Marino, G.; Tutino, M.L. The cold-active Lip1 lipase from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125 is a member of a new bacterial lipolytic enzyme family. *Extremophiles* **2008**, *12*, 311–323.
74. Lee, M.H.; Lee, C.H.; Oh, T.K.; Song, J.K.; Yoon, J.H. Isolation and characterization of a novel lipase from a metagenomic library of tidal flat sediments: evidence for a new family of bacterial lipases. *Appl. Environ. Microbiol.* **2006**, *72*, 7406–7409.
75. Jeon, J.H.; Kim, J.T.; Kim, Y.J.; Kim, H.K.; Lee, H.S.; Kang, S.G.; Kim, S.J.; Lee, J.H. Cloning and characterization of a new cold-active lipase from a deep-sea sediment metagenome. *Appl. Microbiol. Biotechnol.* **2009**, *81*, 865–874.
76. Shanmughapriya, S.; Kiran, G.S.; Selvin, J.; Thomas, T.A.; Rani, C. Optimization, purification, and characterization of extracellular mesophilic alkaline cellulase from sponge-associated *Marinobacter* sp. MSI032. *Appl. Biochem. Biotechnol.* **2009**, doi:10.1007/s12010-009-8747-0.
77. Zeng, R.; Xiong, P.; Wen, J. Characterization and gene cloning of a cold-active cellulase from a deep-sea psychrotrophic bacterium *Pseudoalteromonas* sp. DY3. *Extremophiles* **2006**, *10*, 79–82.
78. Garsoux, G.; Lamotte, J.; Gerday, C.; Feller, G. Kinetic and structural optimization to catalysis at low temperatures in a psychrophilic cellulase from the Antarctic bacterium *Pseudoalteromonas haloplanktis*. *Biochem. J.* **2004**, *384*, 247–253.

79. Ekborg, N.A.; Morrill, W.; Burgoyne, A.M.; Li, L.; Distel, D.L. CelAB, a multifunctional cellulase encoded by *Teredinibacter turnerae* T7902T, a culturable symbiont isolated from the wood-boring marine bivalve *Lyrodus pedicellatus*. *Appl. Environ. Microbiol.* **2007**, *73*, 7785–7788.
80. Cottrell, M.T.; Moore, J.A.; Kirchman, D.L. Chitinases from uncultured marine microorganisms. *Appl. Environ. Microbiol.* **1999**, *65*, 2553–2557.
81. Hobel, C.F.; Hreggvidsson, G.O.; Marteinson, V.T.; Bahrani-Mougeot, F.; Einarsson, J.M.; Kristjansson, J.K. Cloning, expression, and characterization of a highly thermostable family 18 chitinase from *Rhodothermus marinus*. *Extremophiles* **2005**, *9*, 53–64.
82. Wasmund, K.; Burns, K.A.; Kurtboke, D.I.; Bourne, D.G. Novel alkane hydroxylase gene (alkB) diversity in sediments associated with hydrocarbon seeps in the Timor Sea, Australia. *Appl. Environ. Microbiol.* **2009**, *75*, 7391–7398.
83. Ferrer, M.; Beloqui, A.; Timmis, K.N.; Golyshin, P.N. Metagenomics for mining new genetic resources of microbial communities. *J. Mol. Microbiol. Biotechnol.* **2009**, *16*, 109–123.
84. Jones, B.V.; Sun, F.; Marchesi, J.R. Using skimmed milk agar to functionally screen a gut metagenomic library for proteases may lead to false positives. *Lett. Appl. Microbiol.* **2007**, *45*, 418–420.
85. Lee, S.W.; Won, K.; Lim, H.K.; Kim, J.C.; Choi, G.J.; Cho, K.Y. Screening for novel lipolytic enzymes from uncultured soil microorganisms. *Appl. Microbiol. Biotechnol.* **2004**, *65*, 720–726.
86. Guan, C.; Ju, J.; Borlee, B.R.; Williamson, L.L.; Shen, B.; Raffa, K.F.; Handelsman, J. Signal mimics derived from a metagenomic analysis of the gypsy moth gut microbiota. *Appl. Environ. Microbiol.* **2007**, *73*, 3669–3676.
87. Stepanauskas, R.; Sieracki, M.E. Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 9052–9057.
88. Woyke, T.; Xie, G.; Copeland, A.; Gonzalez, J.M.; Han, C.; Kiss, H.; Saw, J.H.; Senin, P.; Yang, C.; Chatterji, S.; Cheng, J.F.; Eisen, J.A.; Sieracki, M.E.; Stepanauskas, R. Assembling the marine metagenome, one cell at a time. *PLoS One* **2009**, *4*, e5299.
89. Gilbert, J.A.; Field, D.; Swift, P.; Newbold, L.; Oliver, A.; Smyth, T.; Somerfield, P.J.; Huse, S.; Joint, I. The seasonal structure of microbial communities in the Western English Channel. *Environ. Microbiol.* **2009**, *11*, 3132–3139.
90. Frias-Lopez, J.; Shi, Y.; Tyson, G.W.; Coleman, M.L.; Schuster, S.C.; Chisholm, S.W.; Delong, E.F. Microbial community gene expression in ocean surface waters. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3805–3810.
91. Gilbert, J.A.; Field, D.; Huang, Y.; Edwards, R.; Li, W.; Gilna, P.; Joint, I. Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. *PLoS One* **2008**, *3*, e3042.
92. Gilbert, J.A.; Thomas, S.; Cooley, N.A.; Kulakova, A.; Field, D.; Booth, T.; McGrath, J.W.; Quinn, J.P.; Joint, I. Potential for phosphonoacetate utilization by marine bacteria in temperate coastal waters. *Environ. Microbiol.* **2009**, *11*, 111–125.
93. Wilmes, P.; Bond, P.L. The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. *Environ. Microbiol.* **2004**, *6*, 911–920.

94. Schweder, T.; Markert, S.; Hecker, M. Proteomics of marine bacteria. *Electrophoresis* **2008**, *29*, 2603–2616.
95. Toyoda, A.; Iio, W.; Mitsumori, M.; Minato, H. Isolation and identification of cellulose-binding proteins from sheep rumen contents. *Appl. Environ. Microbiol.* **2009**, *75*, 1667–1673.
96. Beloqui, A.; Guazzaroni, M.E.; Pazos, F.; Vieites, J.M.; Godoy, M.; Golyshina, O.V.; Chernikova, T.N.; Waliczek, A.; Silva-Rocha, R.; Al-Ramahi, Y.; La Cono, V.; Mendez, C.; Salas, J.A.; Solano, R.; Yakimov, M.M.; Timmis, K.N.; Golyshin, P.N.; Ferrer, M. Reactome array: forging a link between metabolome and genome. *Science* **2009**, *326*, 252–257.
97. Park, H.J.; Jeon, J.H.; Kang, S.G.; Lee, J.H.; Lee, S.A.; Kim, H.K. Functional expression and refolding of new alkaline esterase, EM2L8 from deep-sea sediment metagenome. *Protein Expr. Purif.* **2007**, *52*, 340–347.
98. Jeon, J.H.; Kim, J.T.; Kang, S.G.; Lee, J.H.; Kim, S.J. Characterization and its potential application of two esterases derived from the arctic sediment metagenome. *Mar. Biotechnol. (NY)* **2009**, *11*, 307–316.
99. Park, S.Y.; Kim, J.T.; Kang, S.G.; Woo, J.H.; Lee, J.H.; Choi, H.T.; Kim, S.J. A new esterase showing similarity to putative diene lactone hydrolase from a strict marine bacterium, *Vibrio* sp. GMD509. *Appl. Microbiol. Biotechnol.* **2007**, *77*, 107–115.
100. Liu, Z.; Li, X.; Chi, Z.; Wang, L.; Li, J.; Wang, X. Cloning, characterization and expression of the extracellular lipase gene from *Aureobasidium pullulans* HN2-3 isolated from sea saltern. *Antonie Van Leeuwenhoek* **2008**, *94*, 245–255.
101. Lonhienne, T.; Mavromatis, K.; Vorgias, C.E.; Buchon, L.; Gerday, C.; Bouriotis, V. Cloning, sequences, and characterization of two chitinase genes from the Antarctic *Arthrobacter* sp. strain TAD20: isolation and partial characterization of the enzymes. *J. Bacteriol.* **2001**, *183*, 1773–1779.
102. Gabor, E.M.; de Vries, E.J.; Janssen, D.B. Construction, characterization, and use of small-insert gene banks of DNA isolated from soil and enrichment cultures for the recovery of novel amidases. *Environ. Microbiol.* **2004**, *6*, 948–958.
103. Li, H.; Chi, Z.; Wang, X.; Duan, X.; Ma, L.; Gao, L. Purification and characterization of extracellular amylase from the marine yeast *Aureobasidium pullulans* N13d and its raw potato starch digestion. *Enzyme Microb. Technol.* **2007**, *40*, 1006–1012.
104. Li, X.; Chi, Z.; Liu, Z.; Li, J.; Wang, X.; Hirimuthugoda, N.Y. Purification and characterization of extracellular phytase from a marine yeast *Kodamaea ohmeri* BG3. *Mar. Biotechnol. (NY)* **2008**, *10*, 190–197.
105. Zeng, R.; Zhang, R.; Zhao, J.; Lin, N. Cold-active serine alkaline protease from the psychrophilic bacterium *Pseudomonas* strain DY-A: enzyme purification and characterization. *Extremophiles* **2003**, *7*, 335–337.
106. Acevedo, J.P.; Reyes, F.; Parra, L.P.; Salazar, O.; Andrews, B.A.; Asenjo, J.A. Cloning of complete genes for novel hydrolytic enzymes from Antarctic sea water bacteria by use of an improved genome walking technique. *J. Biotechnol.* **2008**, *133*, 277–286.
107. Chavez Crocker, P.; Sako, Y.; Uchida, A. Purification and characterization of an intracellular heat-stable proteinase (pernilase) from the marine hyperthermophilic archaeon *Aeropyrum pernix* K1. *Extremophiles* **1999**, *3*, 3–9.

108. Collins, T.; Meuwis, M.A.; Stals, I.; Claeysens, M.; Feller, G.; Gerday, C. A novel family 8 xylanase, functional and physicochemical characterization. *J. Biol. Chem.* **2002**, *277*, 35133–35139.

© 2010 by the authors; licensee Molecular Diversity Preservation International, Basel, Switzerland. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).