Marine Genomics 17 (2014) 1–8

Contents lists available at ScienceDirect

Marine Genomics

CIESM 2013

Review

Comparative analysis of deep-sea bacterioplankton OMICS revealed the occurrence of habitat-specific genomic attributes

Francesco Smedile [1], Enzo Messina [1], Violetta La Cono, Michail M. Yakimov *

Institute for Coastal Marine Environment, CNR, Spianata S.Raineri 86, 98122 Messina, Italy

ARTICLE INFO

Article history:
Received 4 April 2014
Received in revised form 23 May 2014
Accepted 3 June 2014
Available online 14 June 2014

Keywords:
Deep-sea bacterioplankton
Next-generation sequencing
Metagenomics
Metatranscriptomics
Chemolithoautotrophy

ABSTRACT

Bathyal aphotic ocean represents the largest biotope on our planet, which sustains highly diverse but low-density microbial communities, with yet untapped genomic attributes, potentially useful for discovery of new biomolecules, industrial enzymes and pathways. In the last two decades, culture-independent approaches of high-throughput sequencing have provided new insights into structure and function of marine bakterioplankton, leading to unprecedented opportunities to accurately characterize microbial communities and their interactions with the environments. In the present review we focused on the analysis of relatively few deep-sea OMICS studies, completed thus far, to find the specific genomic patterns determining the lifeway and adaptation mechanisms of prokaryotes thriving in the dark deep ocean below the depth of 1000 m. Phylogenomic and omic studies provided clear evidence that the bathyal microbial communities are distinct from the epipelagic counterparts and, along with generally larger genomes, possess their own habitat-specific genomic attributes. The high abundance in the deep ocean OMICS of the systems for environmental sensing, signal transduction and metabolic versatility as compared to the epipelagic counterparts is thought to enable the deep-sea bacterioplankton to rapidly adapt to changing environmental conditions associated with resource scarcity and high diversity of energy and carbon substrates in the bathyal biotopes. Together with a versatile heterotrophy, mixotrophy and anaplerosis are thought to enable the deep-sea bacterioplankton to cope with these environmental conditions.

Contents

1. Introduction

The oceans cover more than 70% of the Earth' surface and have an average depth of 3800 m and pressure of 38 MPa. As generally accepted, the life forms thriving under pressures higher than 10 MPa (corresponding to oceanic water below 1000 m) are collectively defined as deep ocean biosphere (Jannasch and Taylor, 1984). Because of the difficulty to access the high-pressure biotopes compared to superficial extreme environments, much less is known about the indigenous deep-sea microbial communities and their physiological adaptation mechanisms to withstand the high hydrostatic pressure. Considering that deep ocean and other high-pressurized biotopes identified thus far are the most voluminous in nature, the deep biosphere could therefore represent a largely "unseen and untapped majority" of life on Earth,

* Corresponding author. Tel.: +39 090 6015437; fax: +39 090 609007.
E-mail address: michail.yakimov@iamc.cnr.it (M.M. Yakimov).
[1] These authors contributed equally to the article.

http://dx.doi.org/10.1016/j.margen.2014.06.001
1874-7787/

perhaps constituting up to 30% of the total living biomass on our Planet, or even more (Whitman et al., 1998).

Layers bathy (1000–4000 m depth), abyssopelagic (4000–6000 m depth) and hadal environments (>6000 m depth), form a physically uniform entity characterized by slightly elevated amount of inorganic nitrogen, absence of solar radiation, high hydrostatic pressure (up to 110 MPa), extreme oligotrophy and low temperature (2–3 °C), except for the warm Mediterranean, Red, and Sulu Seas (Bartlett, 1992; Tamburini et al., 2013). The food web was supposed to be exclusively dependent on the flux of sinking particulate organic carbon (POC) and semi-labile dissolved organic carbon (DOC) produced by phytoplankton and other epipelagic heterotrophic organisms (Reinthaler et al., 2005; Hansell et al., 2009). Such environmental setting of the deep ocean sustained a highly diverse but low-density microbial community of predominantly heterotrophic lifeway (Fuhrman et al., 1992; Witte et al., 2003; Sogin et al., 2006). However, earlier studies have shown the unique features of deep-sea organic carbon (Aluwihare et al., 2002). Besides the POC and DOC fluxes mentioned above, two new additional sources of deep-sea C_{org} were identified. The first one is related to the discovery of buoyant detrital POC possessing the highly nutritive values, whereas the second one is related with *de novo* synthesis of deep-sea C_{org} driven by new groups of ammonia-oxidizing chemolithoautotrophic and bicarbonate-assimilating heterotrophic prokaryotes (Karner et al., 2001; Francis et al., 2005; Konneke et al., 2005; Alonso-Sáez et al., 2010; La Cono et al., 2010; Reinthaler et al., 2010; Yakimov et al., 2011; Buckles et al., 2013).

Our current knowledge about the life at high hydrostatic pressure is derived from studies of deep-sea microorganisms that possess adaptations for growth at pressures roughly in the 10–130 MPa range (Bartlett, 2002; Lauro and Bartlett, 2008; Oger and Jebbar, 2010). Unfortunately, despite significant progress in development of new isolation and cultivation approaches (Rappé et al., 2002; Zengler et al., 2002; Stingl et al., 2007; Woyke et al., 2009), more than 99% of marine prokaryotes still resist cultivation (Fuhrman et al., 1992; Rappé and Giovannoni, 2003). In ecologically cohesive microbial populations it relates mostly to occurrence of established network and tight inter-individual interactions, rather than incomprehension of growth and energy source requirements (Overbeek et al., 2005; Woyke et al., 2006; McInerney et al., 2009; Knittel and Boetius, 2009; Klatt et al., 2011). Thus, bearing in mind that at least 6.5×10^{28} cells are thriving in the deep biosphere (Whitman et al., 1998), this microbial realm represents the largest untapped resource of organisms with yet unknown potential for useful biomolecules and industrial enzymes (Simon and Daniel, 2011).

To get an access to this untouched resource and to avoid the aforementioned cultivation limitations, direct sequencing of environmental (metagenomic) DNA has been proposed for assessment of taxonomic composition and functional characterization of complex microbial communities (von Mering et al., 2007). First used by Handelsman et al. (1998), the term "metagenomic" literally means the study of genetic materials, mainly DNA, directly derived from a microbial community without culturing, whereas metatranscriptomics and metatraproteomics refer to analysis of either environmental pool of gene transcripts or proteins obtained from natural microbial assemblages, respectively (Wilmes and Bond, 2004; Shi et al., 2011). Since 2004, when a landmark work was published describing the application of whole metagenome shotgun sequencing to study the microbial populations in the Sargasso Sea (Venter et al., 2004), an ever-increasing number of marine OMICS papers were published (Fig. 1). Currently, the combination of the complementary OMICS approaches has been considered being the most efficient methodological tool for understanding of the functioning of complex microbial communities (Tringe et al., 2005; DeLong et al., 2006; Biddle et al., 2008; Simon and Daniel, 2011; Meersman et al., 2013; Picard and Daniel, 2013).

In the present review we focused on comparative analysis of the relatively few deep-sea OMICS-based studies, completed thus far (Fig. 1) to elucidate specific genomic attributes determining the lifeway and adaptation mechanisms of prokaryotes thriving in the pelagic portion of the deep, dark biosphere. Peculiar biodiversity hotspots in the deep sea, such as whale falls, hydrothermal vents, cold seeps and deep hypersaline anoxic basins were not considered. Discussed accomplishments came from the analysis of common bathy-, abysso- and hadopelagic environments performed in North Pacific Ocean (ALOHA station Hawaii Subtropical Gyre, 2000 m and 4000 m depth) (DeLong et al., 2006; Konstantinidis et al., 2009; Wu et al., 2013), Central Atlantic Ocean (Puerto Rico Trench, 6000 m depth) (Eloe et al., 2011a), and Eastern Mediterranean (Marmara Sea, 1000 m depth; KM3 station 3010 m depth; Matapan–Vavilov Deep canyon, 4908 m depth) (Martín-Cuadrado et al., 2007; Quaiser et al., 2011; Smedile et al., 2013).

2. Results and discussion

2.1. Overview of metagenomic resources

As we discussed above, the culture-independent approaches have provided new insights into the structure and function of complex microbial communities. The exponentially decreasing costs of data production in high-throughput new generation sequencing (NGS) resulted in broadly available tools for microbial genomic and transcriptional surveys. Here we briefly review the technological and computational OMICS-based tools, approaches, analysis methods and public data that are already available to comprehensively and accurately characterize microbial communities and their interactions with environments. Before the advent of NGS, metagenomic shotgun sequencing was implemented with the classical Sanger technology, characterized by longer read length and low error rate, but quite expensive (high cost per bp) and complex to manage (operator needs to manipulate one sequence a time). This complexity has been overcome with the introduction on the market of 454/Roche and, subsequently, Illumina/Solexa platforms (Mardis, 2008; Metzker, 2009). Both can produce many sequences per run (actually over 500 Mb for 454 and dozens of Gb for Illumina), adequate for a metagenomic analysis. Several new sequencing technologies have been developed in recent years (we mention: SOLiD, PacBio, Helicos, Ion Torrent, and, for completeness, to Oxford Nanopore, Halcyon, Polonator, IBM DNA Transistor, ZS Genetics, US Genomics, Complete Genomics), but Illumina is actually the most popular, probably because of its relatively low costs per bp of reads, in particular for recent Illumina MiSeq models, compared to other platforms (Kalorama Information report, 2013; Barba et al., 2014).

In order to perform a correct metagenomic analysis, once output data from NGS platform are provided, the first step is how to treat these raw sequences. Bioinformatics helps to solve this problem mainly in two ways: by online analysis tools (via web servers), as MG-RAST (Meyer et al., 2008), IMG/M (Markowitz et al., 2012, 2014), CAMERA (Sun et al., 2011), and APIS (Badger et al., 2006); by offline analysis tools, generally creating a collection of software programs linked by script languages to form a pipeline. A typical metagenomic analysis pipeline can be divided into the following functional blocks: an initial part, often called "quality check & filtering," where raw sequences are trimmed (or removed) from various contaminants, vectors, ambiguous, unknown or low quality reads. Once this process has completed, a chimera check is applied to remaining sequences. Chimeras are artificial recombinants between two or more parental sequences that are normally formed when prematurely terminated fragments re-anneal to other template DNA during PCR amplification (Bradley and Hillis, 1997). However, detecting chimeras is still challenging, because NGS platforms generate shorter reads of sequences, making them hard to differentiate the source of parents with insufficient taxonomic information (Kim et al., 2013). After this first common part is completed, the analysis continues in different ways depending on the type of analysis sought (taxonomic or functional). In the first case, sequences are passed to a taxonomic binning tool, by composition-based classification or by

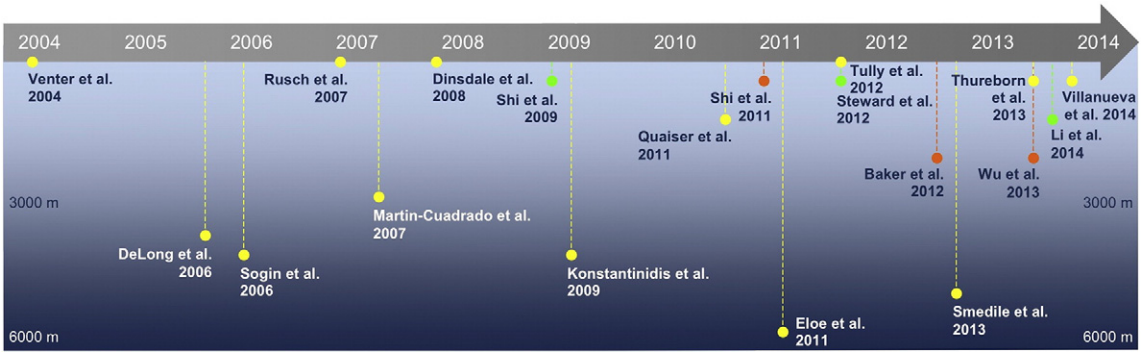

Fig. 1. Overview (timeline vs. depth) of appearance of marine metaomic articles since the pioneered work of Venter et al. (2004), made on Sargasso Sea metagenome. Metagenomic articles are in yellow, metatranscriptomic ones are in green, the articles containing the contemporary performed metagenomic and metatranscriptomic analyses are in orange. Figure inspired by Fig. 1 in the review of Hugenholtz and Tyson (2008).

similarity-based classification. The former includes tools as PhyloPythia (McHardy et al., 2007), and Naive Bayes Classification (Rosen et al., 2011). The latter is principally a BLAST-based classification, and includes tools as MEGAN (Huson et al., 2011), which allows rRNA or protein classification, depending on the database used for analysis: RDP (Cole et al., 2014), Greengenes (DeSantis et al., 2006), or Silva (Quast et al., 2013) for ribosomal RNA; NCBI nr, Pfam (Punta et al., 2012), or SEED (Overbeek et al., 2005) for proteins. In case of functional analysis, sequences are first passed to one of the many assembly programs developed, such as MetaVelvet (Namiki et al., 2012), Newbler (Margulies et al., 2005), or SOAPdenovo (Luo et al., 2012), then to gene prediction tools, such as Glimmer-MG (Kelley et al., 2012), or MetaGene (Noguchi et al., 2006), and finally, to a functional category analysis facilitated by specialized block of tools: the Clusters of Orthologous Groups of proteins (COGs) database (Tatusov et al., 1997), protein family (Pfam, TIGRFAMs) (Haft et al., 2003), gene ontology (GO) (Ashburner et al., 2000), and metabolic pathways and subsystems analysis (KEGG) (Kanehisa and Goto, 2000; Kanehisa et al., 2014).

To compare the metagenomic results obtained from different studies and to attain the statistically acceptable analysis, it is necessary to consider several aspects. The strategy of metagenomic libraries construction, e.g. sequencing of large fosmid inserts vs. random whole genome sequencing (WGS), is among the most important issues. In first case, the total DNA is sheared into large fragments, which further are cloned using suitable vector and sequenced. Due to the large number of clones obtained, it is possible to perform an initial PCR screening looking for particular gene(s) (e.g., 16S rRNA gene) or to randomly sequence the subsamples of pooled clones/fosmids. In WGS analysis, the total DNA is physically broken into much smaller fragments (depending of NGS techniques chosen), directly amplified and then sequenced, typically providing a much larger number of reads, compared to the fosmid sequencing approach. Another aspect to consider is related to the parameters chosen during various steps of analysis (trimming, annotation, etc.), because even small deviations in the values could remarkably change the results of whole analysis. Additionally, the choice toward one or another procedure usually depends on main expectations of obtained outcomes. For example, to overcome the problems associated with miss-assembly, the fosmid sequencing should be used while looking at complex and large multi-domain enzymes (multi-subunit proteins encoded by long operons, polyketide and non-ribosomal peptide synthases, etc.), whereas the WGS should be a preferable procedure to have an overview on the structure and function of complex and highly diverse microbial communities. It should be also noted that, both mentioned techniques have their own limitations. For example, the assembly-facilitated fosmid sequencing approach demonstrated clone library-derived biases compared to direct 454 pyrosequencing of the same environmental (Ghai et al., 2010). Taking into account all these circumstances related to comparative analysis of differentially-produced metadata, many authors and program-developers have suggested the re-elaboration of raw data by using the same programs, filters and parameters (Meyer et al., 2008; Quaiser et al., 2011; Smedile et al., 2013).

2.2. Phylogenomic analysis of prokaryotic communities in the bathyal ocean

As we mentioned above, recent breakthrough in NGS technology has significantly improved and promoted studies on microbial biodiversity. Noteworthy, the taxonomic diversity profiling of 16S rRNA genes found in deep-sea metagenomes was in general comparable with those obtained via conventional 16S rRNA gene-based clone library approach, notwithstanding biases that could characterized amplicon approach (DeLong et al., 2006).

As generally accepted, the PCR-based phylogenetic biodiversity studies can be affected by different sources of errors and lead to a different yield of amplification of the original amplicons: primers misalignment, different secondary structure of region to be amplified, initial dis-equilibrated proportions of target amplicons, etc. (Sogin et al., 2006). Moreover, many prokaryotic groups hold multiple and, at times, heterologous 16S gene copies in the genome. We were aware, that these factors could influence the interpretation of the analysis and lead to an overestimation of abundant species at the expense of rare species, greatly complicating results interpretation. As the main trend detected, members of the Domain *Bacteria* tend to dominate the deep-sea microbial metagenomes at the expenses of *Archaea* (>70% of the total 16 rRNA genes) (Konstantinidis et al., 2009; Eloe et al., 2011a; Smedile et al., 2013). This finding corroborated with taxonomic binning of whole metatranscriptomic data indicating that *Bacteria* are absolutely predominant in the prokaryotic communities in the deep-sea water samples (Wu et al., 2013), whereas *Archaea* were typically found at higher levels in either photic layer or superficial marine sediments rather than in deep pelagic compartments (Quaiser et al., 2011; Schippers et al., 2012).

By adopting the massively 16S rRNA gene parallel tag sequencing data and available bathyal metagenomes, the phylogenetic analysis of the deep-sea microbial communities generally indicated the overwhelming dominance of reads, related to the *Gamma* and *Alphaproteobacteria* classes (≥40% of the total SSU and LSU ribosomal genes), although some differences were found between their distributions with depth in geographically distinct sites. Namely, the *Alphaproteobacteria* members were highly abundant in deep-sea samples collected at ALOHA station (4000 m depth, North Pacific Gyre, [Delong et al., 2006]) and in fosmid terminal sequence analysis-deduced biodiversity at KM3 station (3,000 m depth, the Eastern Mediterranean Sea, [Martín-Cuadrado et al., 2007]), whereas the predominance of the *Gammaproteobacteria* over the *Alphaproteobacteria* was observed in deep water masses of the North Atlantic (4100 m depth, [Sogin et al., 2006; Wu et al., 2013]), the Greenland Sea and the Ionian

Sea (2000 m and 3000 m depth, respectively [Zaballos et al., 2006]). The 16S rRNA fragments, recruited from the abyssal water masses (4908 m depth) at the Matapan–Vavilov Deep station (the Eastern Mediterranean Sea), were placed almost exclusively within the *Gammaproteobacteria* (Smedile et al., 2013). Observed discrepancy may be explained either by the different environmental settings of the sampling sites or by the technical distinctions in the sampling and sample processing. In particular, it might be referred to the sample prefiltration, which was applied in some studies to avoid the interference by DNA from eukaryotic organisms, and which at the same time could eliminate significant amount of POC and, as a consequence, the *Gammaproteobacteria*, which are in perceptible amounts described as marine particle-associated organisms (DeLong et al., 1993; Lauro and Bartlett, 2008).

As for the other eubacterial taxa, members of the phyla *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Deferribacteres*, and *Planctomycetes* generally constituted the remaining groups of deep-sea bacterioplankton. *Actinobacteria* and *Firmicutes* increased their representation in the deep-sea prokaryotic community compared to surface water, probably due to the adaptive advantage of these organisms to thrive under extreme oligotrophic conditions typically reigning in the deep ocean (Gärtner et al., 2011; Wu et al., 2013). Together, anaerobic members of *Bacteroidia*, *Deferribacteres* and *Planctomycetes*, generally absent in oxygenated superficial seawater, made up a significant portion of deep-sea bacterioplankton. This is consistent with previous statement that the preferable lifeway of deep-sea bacterioplankton is likely associated with large marine snow particles that may contain anoxic microniches (Karl et al., 1984; Shanks and Reeder, 1993; Stocker et al., 2008; Swan et al., 2011). However, this finding does not neglect the role of other factors, such as generally low bioavailability of oxygen and malfunction of membrane-based aerobic cytochrome respiratory system under extremely high hydrostatic pressure of the deep sea (Bartlett, 2002; Lauro et al., 2004). Hence, the presence of anaerobic organisms in oxygenated deep-sea water is likely the result of a combination of factors (Wu et al., 2013; Villanueva et al., 2014).

Conspicuous numbers of taxonomically significant reads related to the phylum *Cyanobacteria*, generally dominating the superficial photic layer, was detected in dark deep-sea metagenomes. Simultaneous metagenomic and metatranscriptomic analysis, performed with bakterioplankton of meso- and bathypelagic realms of North Pacific Ocean, demonstrated the occurrence of ribosome-containing photosynthetic bacteria belonging to the genera *Cyanobium*, *Cyanothece*, *Synechococcus*, *Anabaena*, *Cylindrospermopsis*, *Nodularia*, *Nostoc*, *Arthrospira*, *Lyngbya*, *Microcoleus* and *Oscillatoria* (Wu et al., 2013). Although metatranscriptomic analysis revealed that only few of them were metabolically active in the deep (Wu et al., 2013). Due to the fact, that cell aggregation phenomenon is frequently observed in many *Cyanobacteria* species (Koblížek et al., 2001), metabolically active *Cyanobacteria* could be dragged down by marine particles (Amacher et al., 2013), although further study is necessary to provide more information about the presence of metabolically active *Cyanobacteria* in dark, deep-sea environments (Wu et al., 2013).

As we mentioned above, bacteria are absolutely predominant (>70%) over archaea in all available deep-sea OMICS datasets. Given that unlike eubacteria, the marine archaeal cells harbor in their genomes only one 16S rRNA gene copy in average, the reduced number of organisms taken together with the overabundance of 16S rRNA genes of bacterial origin, may explain the low number of archaeal 16S rRNA gene in the deep-sea metagenomes (Konstantinidis et al., 2009; Eloe et al., 2011a; Smedile et al., 2013). Noteworthy, this trend seems more evident in the abyssal and hadal samples, where only 1.5% of archaeal ribosomal genes were found in Puerto Rico Trench at 6000 m depth (Eloe et al., 2011a) and none of archaeal 16S rRNA signatures were detected in Matapan–Vavilov Deep abyssal metagenome (4908 m depth, Smedile et al., 2013). Based on taxonomic binning, whole genomic content of these two deepmost metagenomes, the archaeal fraction was mainly represented by members of the *Euryarchaeota* (*Halobacteria*, *Methanomicrobia* and *Thermococci*), rather than by *Thaumarchaeota* phylotypes. Recent studies on meso- and bathypelagic Atlantic metagenomes suggested that archaeal abundances were correlated with the particular water masses and that in the deep North Atlantic the *Thaumarchaeota* display a latitudinal decline in abundances toward the equator (Varela et al., 2008; Galand et al., 2009a, 2009b). At a first sight, these findings are opposed to the previous surveys, obtained via conventional 16S rRNA gene-based clone library approach. Majority of these reports, dealing with shallower compartments (<4000 m depth), have demonstrated the absolute dominance of the *Thaumarchaeota* over the *Euryarchaeota* (La Cono et al., 2009; Eloe et al., 2011a;). Moreover, compared with the superficial photic layer of water column, the members of Marine Group I of the *Thaumarchaeota* in the dark ocean were typically increasing in abundance with depth (Karner et al., 2001; Herndl et al., 2005; Wuchter et al., 2006). Same trend was reported in the PCR-derived 16S rRNA diversity studies performed in the meso- and bathypleagic Mediterranean Sea (La Cono et al., 2009, 2010; Martín-Cuadrado et al., 2009; Eloe et al., 2011b; Yakimov et al., 2011). These reports indicated an almost exclusively prevalence of Group I Marine *Thaumarchaeota* in pelagic archaeal communities inhabiting aphotic water column to a depth of 3500 m. Given the differences observed between these results, it is clear that neither metagenome-based, nor PCR-derived 16S rRNA-based analysis is sufficient to fully characterize deep-sea archaeal communities and to gain the information on their *in situ* metabolic activities. For this aim, contemporaneous more comprehensive OMICS approaches, including metatranscriptomics and metaproteomics, must be applied (Wu et al., 2013).

2.3. Habitat-specific genomic attributes of bathyal microbial communities

Since the relatively few OMICS studies have been carried out so far in deep pelagic ocean below 1000 m depth, we do not possess a sufficient knowledge for proper understanding of the selective pressures and metabolic pathways prevailing there as yet. It is well known, that more than a half of all deduced protein sequences in the deep-sea metagenomes were annotated as hypothetical or conserved hypothetical proteins (Xu and Ma, 2007; Lauro and Bartlett, 2008; Wang et al., 2008; Konstantinidis et al., 2009; Ferrer et al., 2012). Nevertheless, the comparative analyses of epi- and bathypelagic metagenomes revealed striking similarity in general outcomes, suggesting that, regardless to geographic location of the studied sites, the differences between the dark deep-sea and photic epipelagic biotopes are consistent and reflect the distinct lifestyle of organisms living there (Konstantinidis et al., 2009). So far as the deep ocean microbial communities are distinct from those thriving in epipelagic water column, their genomic attributes should be also biotope-specific (Eloe et al., 2011a; Colwell and D'Hondt, 2013; Meersman et al., 2013). As a generic finding observed in bathyal genomic data, many deep-sea microorganisms, having a similar metabolism to their surface counterparts, generally possessed larger coding regions among clade orthologs, larger intergenic regions and larger estimated average genome size (Konstantinidis et al., 2009; Meersman et al., 2013; Thrash et al., 2014). Using the method for genome size estimation (Raes et al., 2007), Konstantinidis et al. (2009) predicted a substantial increase of 1.35 ± 0.25-fold in genome size in bathy- versus epipelagic, shallow-water-based microbial genomes. Analysis of deep-sea genomic attributes highlighted remarkably higher proportions of genetic loci for regulatory and signal transduction, transcriptional regulation, diverse transport and metabolic pathways, and mobile and transposable elements.

Analyses of deepmost abyssal and hadal metagenomes led to very similar observations, i.e. a high abundance of genes for functional categories of environmental sensing, signal transduction, transcription, transport and use various carbon sources including recalcitrant organic compound and biopolymers (Eloe et al., 2011a; Smedile et al., 2013).

More specifically, the COGs overrepresented in deepmost metagenomes fell within the category T (signal transduction mechanisms): FOG: PAS/PAC domain proteins (COG2202), FOG: CheY-like receiver protein (COG0784), signal transduction histidine kinases (COG0642), and FOG: GGDEF, EAL, and GAF domain proteins (COG2199, COG5001, COG2200, and COG2203). The PAS/PAC domain-containing proteins may serve as examples of the expanded signal transduction capabilities. They are located in the cytosol and functioned as internal sensors of redox potential and oxygen availability (Taylor and Zhulin, 1999). The genes encoding histidine kinase are important for chemotaxis and quorum sensing (Wolanin et al., 2002). The category K (transcription and transcriptional regulation) is also overrepresented in very deep metagenomes. Example of this category is an increased occurrence of the DNA-directed RNA polymerase specialized sigma subunit 24. As it was shown for the moderate piezophile *Photobacterium profundum* SS9 and piezotolerant *Shewanella piezotolerans* WP3, this enzyme serves as an alternative RNA polymerase sigma factor, which plays a significant role in outer membrane protein synthesis, heavy metal resistance and growth at low-temperature and high pressure (Chi and Bartlett, 1995; Vezzi et al., 2005). These trends are in agreement with the statement that an adaptation to increased pressure implies a remarkable ability to control the transcriptional or post-transcriptional activity (Konstantinidis et al., 2009). DNA repair capacity and cell wall/membrane/envelope biogenesis (Heusipp et al., 2007; Shi et al., 2012) seem to be two other characteristics required for survival in the deep ocean, likely due to the elevated requirements under high pressure for cell wall maintenance and DNA integrity (Nakai et al., 2011; Wu et al., 2013). The subsequent system includes genes or gene clusters encoding for nutrient acquisition, energy production and polyunsaturated fatty acid (PUFA) synthesis (Eloe et al., 2008; Campanaro et al., 2008; Wang et al., 2008). Due to energetic constraints occurring in the oligotrophic and generally iron-depleted deep ocean, the deep-sea bacterioplankton apparently prefer to own the sophisticated iron acquisition systems, rather than the energetically expensive *de novo* siderophore biosynthesis systems (Cordero et al., 2012). In fact, a comparison of epipelagic and deep-sea metagenomes showed an emergence at higher depths of Fe^{3+}-siderophores transport/acquisition systems (COG1120) at the expenses of siderophore synthases (Smedile et al., 2013; Li et al., 2014).

Along with the consideration of the over-represented genes, at least two categories of genes should be discussed due to their dramatic under-representation in the deep-sea (meta)genomes. We mean the genes associated with light-driven processes and with repair of light-induced damages, including photosynthesis, rhodopsin photoproteins, DNA photolyase and catalase genes. All of them were found in members of epipelagic bacterioplankton and largely absent from dark deep-oceanic environments (Swan et al., 2011; Meersman et al., 2013; Luo et al., 2014).

2.4. Particle-associated lifestyle and oxygen availability in the dark ocean

The cultivation of deep-sea bacteria (moderate piezophile *P. profundum* SS9 and piezotolerant *S. piezotolerans* WP3) under high hydrostatic pressure significantly increased the expression of genes involved in chemotaxis and flagellar motility systems, which are considered as a manifestation of bacterial adaptation to the deep-sea extreme oligotrophy (Eloe et al., 2008; Wang et al., 2008; Campanaro et al., 2005). The ability to perceive the trails left by the POC and the power to swim actively toward this nutrient-rich marine snow particles should be an attractive way to overcome both organic and nutrient poorness of deep pelagic water masses (Azam and Long, 2001; Kiørboe and Jackson, 2001). Thus, the POC-associated lifestyle seems a preferable behavior for the deep ocean bacterioplankton to maintain a reasonable metabolic activity under these extreme conditions (Eloe et al., 2011a; Meersman et al., 2013; Smedile et al., 2013). This statement could explain the high quantity of subsystems and categories involved in heavy metal resistance and detoxification, such as mercuric reductase, Co/Zn/Cd efflux system components typical for deep-sea genomes and metagenomes. The elevated presence of mercury resistance genes in the genome of bathypelagic isolate *Alteromonas macleodii* AltDE, as opposed to its shallow counterpart, ecotype ATCC 27126, is thought to help the bacterium to withstand high concentration of heavy metals and trace elements, generally adsorbed on the surface of marine particles (Ivars-Martinez et al., 2008; Smedile et al., 2013). Accordingly to the nature of marine particles and to the POC-associated lifestyle, the deep-sea metagenomes are generally enriched in genes encoding diverse sets of enzymes involved in degradation of biopolymers and recalcitrant compounds (Eloe et al., 2011a; Meersman et al., 2013; Smedile et al., 2013). Enzymes such as alpha(beta)-glucosidases, leucine aminopeptidases, arylsulfatases, alkaline phosphatases generally required to initiate the re-mineralization of various high molecular weight organic compounds like glycoproteins, hetero- and lipopolysaccharides that are principal components of POM (Arnosti, 2002; Aristegui et al., 2009). Recently it was demonstrated that both oxygen-depleted and anaerobic microniches could be found within large organic marine particle aggregates (Ivars-Martinez et al., 2008) and that active sulphur cycle, involving sulfate reduction and sulphide oxidation, could occur even in well-oxygenated dark water columns (Swan et al., 2011). Taking into consideration also the fact, that compared to atmospheric pressure, some redox reactions (e.g., nitrate and nitrite reduction) yield more energy under low-temperature/high-pressure conditions (Fang et al., 2010), it is logical to hypothesize that the deep ocean bacterioplankton may contain an appreciable amount of both microaerophilic and strict anaerobic organisms. Consistently with this assumption, the phylogenomic analysis of the deep-sea metagenomic data, mentioned above, has demonstrated the presence of several taxa of obligate anaerobes. Moreover, the genomic signatures of metabolic pathways, such as anaerobic ammonium oxidation, sulfate and nitrate reduction, autotrophic CO_2 fixation, coupled to the oxidation of reduced sulfur compounds, were detected in this database (Swan et al., 2011; Smedile et al., 2013; Wu et al., 2013).

2.5. Bicarbonate assimilation in the dark ocean

As it well established, the metabolic activities of microbial populations rely on resources availability and their variability. Understanding carbon and nutrient cycling in the deep ocean is central to understand the functioning of whole ecosystem. In addition to heterotrophic microbial activity, postulated as predominant metabolism in the dark deep ocean (Witte et al., 2003), different evidences of chemolithoautotrophy (e.g., inorganic carbon assimilation) have been identified in deep-sea metagenomic studies. This process, termed the dark ocean primary production, is very significant as *de novo* synthesis of deep-sea C_{org} and may be of the same order of magnitude as the dark ocean's heterotrophic production (Reinthaler et al., 2010; Yakimov et al., 2011). The predominant type of chemolithoautotrophy in the dark ocean is generally assumed be conducted by ammonia-oxidizing members of Marine Group I of *Thaumarchaeota* via 3-hydroxypropionate/4-hydroxybutyrate cycle (Venter et al., 2004; Herndl et al., 2005; Hallam et al., 2006a, 2006b; Ingalls et al., 2006; Wuchter et al., 2006; Berg et al., 2007; Yakimov et al., 2007). However, in some ammonia-depleted deep-sea regions archaeal nitrification may be insufficient to support the measured inorganic carbon fixation rates (Herndl et al., 2005; Reinthaler et al., 2010). Moreover, there are recent evidences of active bicarbonate assimilation in archaea-impoverished bathyal environments (Smedile et al., 2013). Thus, as yet unidentified microbial lineages and energy sources may be responsible for a significant fraction of carbon fixation in the dark ocean. As we mentioned above, several phylogenomic studies revealed a slight increase of *Planctomycetes* in deep-sea. These organisms are capable of anaerobic ammonium oxidation coupled with assimilation of inorganic carbon, which may be another metabolic pathway supporting primary production in dark bathyal environment (Kuypers et al., 2005; Wu et al., 2013; Villanueva

et al., 2014). The analysis of 502 mesopelagic single amplified genomes demonstrated that certain lineages of indigenous and abundant in the dark ocean *Proteobacteria* have the potential for autotrophic fixation of bicarbonate, fuelled by the oxidation of reduced sulphur compounds (Swan et al., 2011). Additionally, some of the dark ocean prokaryotes may be methano- and methylotrophs, using reduced C_1-compounds as energy and carbon sources for growth (Tavormina et al., 2010).

Besides the true autotrophy, recent studies suggested that other metabolic processes could contribute to the total assimilation of inorganic carbon in the dark ocean (Martín-Cuadrado et al., 2009; Alonso-Sáez et al., 2010; Swan et al., 2011; Yakimov et al., in press (accepted)). For example, Cox genes encoding different subunits of the carbon monoxide dehydrogenase (CoxL/CoxM/CoxS) involved in aerobic CO oxidation were detected in remarkable concentration in several deep-sea metagenomes all over the world (Martín-Cuadrado et al., 2007; DeLong and Béjà, 2010; Quaiser et al., 2011; Smedile et al., 2013). Oxidation of carbon monoxide, initially thought to be exclusive of carboxydotrophic autotrophs, is now discovered in a plethora of marine bacteria and is likely involved in the heterotrophic fixation of inorganic carbon as an alternative or supplementary energy source (Martín-Cuadrado et al., 2007, 2009). Additionally, the heterotrophic bacteria can contribute to assimilation of CO_2 via carboxylation reactions (known as anaplerotic pathways) performed by various carboxylases (Romanenko, 1964). In some marine ecosystems these anaplerotic reactions are likely responsible for about 1%–8% of the total heterotrophic bacterioplankton biomass production (Alonso-Sáez et al., 2010; Reinthaler et al., 2010). These previously unrecognized metabolic types of dark ocean bacterioplankton may play an important role in global biogeochemical cycles, and their activities may in part reconcile current discrepancies in the dark ocean's carbon budget.

2.6. Conclusions and future perspectives

Since the first publication on deep-sea marine metagenomics-based studies, OMICS approaches have strikingly increased our knowledge of abundance, diversity and gene content of marine microbes promoting our understanding on the functioning of the dark ocean ecosystems (Ferrer et al., 2012). In order to unveil veritable metabolic activities of the microorganisms and to monitor how those activities are changing in response to environmental forces or biotic interactions, metatranscriptomics was juxtaposed to metagenomics (Mitra et al., 2011; Shi et al., 2012). Despite that it was proved to be efficacious in deciphering of metabolic potential and operative pathways of deep-sea microbial communities, these OMICS tools can do little against the insufficient functional characterization and non-validated annotations of proteins in databases and obviously against the paucity of suitable reference genomes (Ferrer et al., 2008; Woyke et al., 2009). Recent breakthrough in metaproteomics techniques (reviewed in this special issue by Hartmann et al.), applied to marine biotopes and dealt directly with the pool of environmental proteins, gave a major boost to the identification and annotation of novel proteins of unknown functions (Wang et al., 2014). Thanks to the new NGS techniques, such as Illumina and single molecule real time sequencing applied for the analyses of prokaryotic single amplified genomes, the deficiency of suitable reference genomes was almost resolved (Woyke et al., 2009; Schadt et al., 2010; Shin et al., 2013; Thrash et al., 2014). Although not without its own drawbacks, this approach has expanded our knowledge on physiology of uncultured deep-sea microbes and became a good alternative to solve the problems related with elevated costs of large insert libraries analysis and with WGS assembling of complex microbial communities (Lasken, 2012; Blainey, 2013; Lloyd et al., 2013). It is worth also to notice the alternative approach to deal with untapped genomic resources related to construction of (meta)genomic expression libraries and their further screening for targeted activity. Thus, the application of this approach could generally help to overcome the relying exclusively on sequence homology searches for enzyme detection. Combining of OMICS techniques together with this approach, is obviously the method of choice to capture the greatest untapped enzyme diversity present in the dark ocean, allowing the functional attribution of protein domains of unknown function, the discovery of new functionalities of known enzymes, new sequence classes and sometimes new activities (Ferrer et al., 2008; Wang et al., 2014).

Acknowledgments

This paper was produced thanks to financial support of CNR in the frame of EU Projects (FP7-KBBE-2009-2B-226977 MAMBA and EC FP7-OCEAN.2011-2-287589 MICROB3) and in the frame of Italian National project FIRB 2008 EXPLODIVE RBFR08AWP6_002.

References

Alonso-Sáez, L., Galand, P.E., Casamayor, E.O., Pedrós-Alió, C., Bertilsson, S., 2010. High bicarbonate assimilation in the dark by Arctic bacteria. ISME J. 4, 1581–1590.
Aluwihare, L.I., Repeta, D.J., Chen, R.F., 2002. Chemical composition and cycling of dissolved organic matter in the Mid-Atlantic Bight. Deep Sea Res. Part II 49, 4421–4437.
Amacher, J., Neuer, S., Lomas, M., 2013. DNA-based molecular fingerprinting of eukaryotic protists and cyanobacteria contributing to sinking particle flux at the Bermuda Atlantic time-series study. Deep Sea Res. Part II 93, 71–83.
Aristegui, J., Gasol, J.M., Duarte, C.M., Herndl, G.J., 2009. Microbial oceanography of the dark ocean's pelagic realm. Limnol. Oceanogr. 54, 1501–1529.
Arnosti, C., 2002. Microbial extracellular enzymes and their role in dissolved organic matter cycling. In: Findlay, S., Sinsabaugh, R.L. (Eds.), Aquatic Ecosystems: Interactivity of Dissolved Organic Matter. Academic Press, San Diego, CA, USA, pp. 316–342.
Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al., 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25, 25–29.
Azam, F., Long, R.A., 2001. Sea snow microcosms. Nature 414, 497–498.
Badger, J.H., Hoover, T.R., Brun, Y.V., Weiner, R.M., Laub, M.T., Alexandre, G., Mrázek, J., Ren, Q., Paulsen, I.T., Nelson, K.E., et al., 2006. Comparative genomic evidence for a close relationship between the dimorphic prosthecate bacteria *Hyphomonas neptunium* and *Caulobacter crescentus*. J. Bacteriol. 188, 6841–6850.
Barba, M., Czosnek, H., Hadidi, A., 2014. Historical perspective, development and applications of next-generation sequencing in plant virology. Viruses 6, 106–136.
Bartlett, D.H., 1992. Microbial life at high pressures. Sci. Prog. 76, 479–496.
Bartlett, D.H., 2002. Pressure effects on in vivo microbial processes. Biochim. Biophys. Acta 1595, 367–381.
Berg, I.A., Kockelkorn, D., Buckel, W., Fuchs, G., 2007. A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. Science 318, 1782–1786.
Biddle, J.F., Fitz-Gibbon, S., Schuster, S.C., Brenchley, J.E., House, C.H., 2008. Metagenomic signatures of the Peru Margin subseafloor biosphere show a genetically distinct environment. Proc. Natl. Acad. Sci. USA 105, 10583–10588.
Blainey, P.C., 2013. The future is now: single-cell genomics of bacteria and archaea. FEMS Microbiol. Rev. 37, 407–427.
Bradley, R.D., Hillis, D.M., 1997. Recombinant DNA sequences generated by PCR amplification. Mol. Biol. Evol. 14, 592–593.
Buckles, L.K., Villanueva, L., Weijers, J.W., Verschuren, D., Damsté, J.S., 2013. Linking isoprenoidal GDGT membrane lipid distributions with gene abundances of ammonia-oxidizing *Thaumarchaeota* and uncultured crenarchaeotal groups in the water column of a tropical lake (Lake Challa, East Africa). Environ. Microbiol. 15, 2445–2462.
Campanaro, S., Vezzi, A., Vitulo, N., Lauro, F.M., D'Angelo, M., Simonato, F., Cestaro, A., Malacrida, G., Bertoloni, G., Valle, G., et al., 2005. Laterally transferred elements and high pressure adaptation in *Photobacterium profundum* strains. BMC Genomics 6, 122.
Campanaro, S., Treu, L., Valle, G., 2008. Protein evolution in deep sea bacteria: an analysis of amino acids substitution rates. BMC Evol. Biol. 8, 313.
Chi, E., Bartlett, D.H., 1995. An rpoE-like locus controls outer membrane protein synthesis and growth at cold temperatures and high pressures in the deep-sea bacterium *Photobacterium* sp. strain SS9. Mol. Microbiol. 17, 713–726.
Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., Tiedje, J.M., 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 42, D633–D642.
Colwell, F.S., D'Hondt, S., 2013. Nature and extent of the deep biosphere. Rev. Mineral. Geochem. 75, 547–574.
Cordero, O.X., Wildschutte, H., Kirkup, B., Proehl, S., Ngo, L., Hussain, F., Le Roux, F., Mincer, T., Polz, M.F., 2012. Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. Science 337, 1228–1231.
DeLong, E.F., Béjà, O., 2010. The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. PLoS Biol. 8, e1000359.
DeLong, E.F., Frankel, R.B., Bazylinski, D.A., 1993. Multiple evolutionary origins of magnetotaxis in bacteria. Science 259, 803–806.
DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., Frigaard, N.U., 2006. Community genomics among stratified microbial assemblages in the ocean's interior. Science 311, 496–503.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., Andersen, G.L., 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. 72, 5069–5072.
Eloe, E.A., Lauro, F.M., Vogel, R.F., Bartlett, D.H., 2008. The deep-sea bacterium *Photobacterium profundum* SS9 utilizes separate flagellar systems for swimming and swarming under high-pressure conditions. Appl. Environ. Microbiol. 74, 6298–6305.
Eloe, E.A., Fadrosh, D.W., Novotny, M., Zeigler Allen, L., Kim, M., Lombardo, M.J., Yee-Greenbaum, J., Yooseph, S., Allen, E.E., Lasken, R., et al., 2011a. Going deeper: metagenome of a hadopelagic microbial community. PLoS One 6, e20388.
Eloe, E.A., Shulse, C.N., Fadrosh, D.W., Williamson, S.J., Allen, E.E., Bartlett, D.H., 2011b. Compositional differences in particle-associated and free-living microbial assemblages from an extreme deep-ocean environment. Environ. Microbiol. Rep. 3, 449–458.
Fang, J., Zhang, L., Bazylinski, D.A., 2010. Deep-sea piezosphere and piezophiles: geomicrobiology and biogeochemistry. Trends Microbiol. 18, 413–422.
Ferrer, M., Beloqui, A., Timmis, K.N., Golyshin, P.N., 2008. Metagenomics for mining new genetic resources of microbial communities. J. Mol. Microbiol. Biotechnol. 16, 109–123.
Ferrer, M., Werner, J., Chernikova, T.N., Bargiela, R., Fernández, L., La Cono, V., Waldmann, J., Teeling, H., Golyshina, O.V., Glöckner, F.O., et al., 2012. Unveiling microbial life in the new deep-sea hypersaline Lake Thetis. Part II: a metagenomic study. Environ. Microbiol. 14, 268–281.
Francis, C.A., Roberts, K.J., Beman, J.M., Santoro, A.E., Oakley, B.B., 2005. Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. Proc. Natl. Acad. Sci. U. S. A. 102, 14683–14688.
Fuhrman, J.A., McCallum, K., Davis, A.A., 1992. Novel major archaebacterial group from marine plankton. Nature 356, 148–149.
Galand, P.E., Lovejoy, C., Hamilton, A.K., Ingram, R.G., Pedneault, E., Carmack, E.C., 2009a. Archaeal diversity and a gene for ammonia oxidation are coupled to oceanic circulation. Environ. Microbiol. 11, 971–980.
Galand, P.E., Casamayor, E.O., Kirchman, D.L., Potvin, M., Lovejoy, C., 2009b. Unique archaeal assemblages in the Arctic Ocean unveiled by massively parallel tag sequencing. ISME J. 3, 860–869.
Gärtner, A., Blümel, M., Wiese, J., Imhoff, J.F., 2011. Isolation and characterisation of bacteria from the Eastern Mediterranean deep sea. Antonie Van Leeuwenhoek 100, 421–435.
Ghai, R., Martín-Cuadrado, A.B., Molto, A.G., Heredia, I.G., Cabrera, R., Martin, J., Verdú, M., Deschamps, P., Moreira, D., López-García, P., et al., 2010. Metagenome of the Mediterranean deep chlorophyll maximum studied by direct and fosmid library 454 pyrosequencing. ISME J. 4, 1154–1166.
Haft, D.H., Selengut, J.D., White, O., 2003. The TIGRFAMs database of protein families. Nucleic Acids Res. 31, 371–373.
Hallam, S.J., Mincer, T.J., Schleper, C., Preston, C.M., Roberts, K., Richardson, P.M., DeLong, E.F., 2006a. Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine *Crenarchaeota*. PLoS Biol. 4, e95.
Hallam, S.J., Konstantinidis, K.T., Putnam, N., Schleper, C., Watanabe, Y., Sugahara, J., Preston, C., de la Torre, J., Richardson, P.M., DeLong, E.F., 2006b. Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. Proc. Natl. Acad. Sci. U. S. A. 103, 18296–18301.
Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., Goodman, R.M., 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem. Biol. 5, R245–R249.
Hansell, D.A., Carlson, C.A., Repeta, D.J., Schlitzer, R., 2009. Dissolved organic matter in the ocean: new insights stimulated by a controversy. Oceanography 22, 52–61.
Herndl, G.J., Reinthaler, T., Teira, E., van Aken, H., Veth, C., Pernthaler, A., Pernthaler, J., 2005. Contribution of *Archaea* to total prokaryotic production in the deep Atlantic Ocean. Appl. Environ. Microbiol. 71, 2303–2309.
Heusipp, G., Falker, S., Schmidt, M.A., 2007. DNA adenine methylation and bacterial pathogenesis. Int. J. Med. Microbiol. 297, 1–7.
Hugenholtz, P., Tyson, G.W., 2008. Microbiology – Metagenomics. Nature 455, 481–483.
Huson, D.H., Mitra, S., Weber, N., Ruscheweyh, H., Schuster, S.C., 2011. Integrative analysis of environmental sequences using MEGAN4. Genome Res. 21, 1552–1560.
Ingalls, A.E., Shah, S.R., Hansman, R.L., Aluwihare, L.I., Santos, G.M., Druffel, E.R., Pearson, A., 2006. Quantifying archaeal community autotrophy in the mesopelagic ocean using natural radiocarbon. Proc. Natl. Acad. Sci. U. S. A. 103, 6442–6447.
Ivars-Martinez, E., Martin-Cuadrado, A.B., D'Auria, G., Mira, A., Ferriera, S., Johnson, J., Friedman, R., Rodriguez-Valera, F., 2008. Comparative genomics of two ecotypes of the marine planktonic copiotroph *Alteromonas macleodii* suggests alternative lifestyles associated with different kinds of particulate organic matter. ISME J. 2, 1194–1212.
Jannasch, H.W., Taylor, C.D., 1984. Deep-sea microbiology. Annu. Rev. Microbiol. 38, 487–514.
Kanehisa, M., Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30.
Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M., 2014. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 42, D199–D205.
Karl, D.M., Knauer, G.A., Martin, J.H., Ward, B.B., 1984. Bacterial chemolithotrophy in the ocean is associated with sinking particles. Nature 309, 54–56.
Karner, M.B., DeLong, E.F., Karl, D.M., 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. Nature 409, 507–510.
Kelley, D.R., Liu, B., Delcher, A.L., Pop, M., Salzberg, S.L., 2012. Gene prediction with Glimmer for metagenomic sequences augmented by classification and clustering. Nucleic Acids Res. 40, e9.
Kim, M., Lee, K.H., Yoon, S.W., Kim, B.S., Chun, J., Yi, H., 2013. Analytical tools and databases for metagenomics in the next-generation sequencing era. Genomics Inform. 11, 102–113.
Kiørboe, T., Jackson, G.A., 2001. Marine snow, organic solute plumes, and optimal sensory behaviour of bacteria. Linmol. Oceanogr. 46, 1309–1318.
Klatt, C.G., Wood, J.M., Rusch, D.B., Bateson, M.M., Hamamura, N., Heidelberg, J.F., Grossman, A.R., Bhaya, D., Cohan, F.M., Kühl, M., et al., 2011. Community ecology of hot spring cyanobacterial mats: predominant populations and their functional potential. ISME J. 5, 1262–1278.
Knittel, K., Boetius, A., 2009. Anaerobic oxidation of methane: progress with an unknown process. Annu. Rev. Microbiol. 63, 311–334.
Koblížek, M., Komenda, J., Masojídek, J., Pechar, L., 2001. Cell aggregation of the cyanobacterium *Synechococcus elongatus*: role of the electron transport chain. J. Phycol. 36, 662–668.
Konneke, M., Bernhard, A.E., de la Torre, J.R., Walker, C.B., Waterbury, J.B., Stahl, D.A., 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. Nature 437, 543–546.
Konstantinidis, K.T., Braff, J., Karl, D.M., DeLong, E.F., 2009. Comparative metagenomic analysis of a microbial community residing at a depth of 4,000 meters at station ALOHA in the North Pacific subtropical gyre. Appl. Environ. Microbiol. 75, 5345–5355.
Kuypers, M.M.M., Lavik, G., Woebken, D., Schmid, M., Fuchs, B.M., Amann, R., Jørgensen, B.B., Jetten, M.S.M., 2005. Massive nitrogen loss from the Benguela upwelling system through anaerobic ammonium oxidation. Proc. Natl. Acad. Sci. U. S. A. 102, 6478–6483.
La Cono, V., Tamburini, C., Genovese, L., La Spada, G., Denaro, R., Yakimov, M.M., 2009. Cultivation-independent assessment of the bathypelagic archaeal diversity of Tyrrhenian Sea: comparative study of rDNA and rRNA-derived libraries and influence of sample decompression. Deep Sea Res. Part II 56, 768–773.
La Cono, V., Smedile, F., Ferrer, M., Golyshin, P.N., Giuliano, L., Yakimov, M.M., 2010. Genomic signatures of fifth autotrophic carbon assimilation pathway in bathypelagic Crenarchaeota. Microb. Biotechnol. 3, 595–606.
Lasken, R.S., 2012. Genomic sequencing of uncultured microorganisms from single cells. Nat. Rev. Microbiol. 10, 631–640.
Lauro, F.M., Bartlett, D.H., 2008. Prokaryotic lifestyles in deep sea habitats. Extremophiles 12, 15–25.
Lauro, F.M., Bertoloni, G., Obraztsova, A., Kato, C., Tebo, B.M., Bartlett, D.H., 2004. Pressure effects on *Clostridium* strains isolated from a cold deep-sea environment. Extremophiles 8, 169–173.
Li, M., Toner, B.M., Baker, B.J., Breier, J.A., Sheik, C.S., Dick, G.J., 2014. Microbial iron uptake as a mechanism for dispersing iron from deep-sea hydrothermal vents. Nat. Commun. 5, 3192.
Lloyd, K.G., Schreiber, L., Petersen, D.G., Kjeldsen, K.U., Lever, M.A., Steen, A.D., Stepanauskas, R., Richter, M., Kleindienst, S., Lenk, S., et al., 2013. Predominant archaea in marine sediments degrade detrital proteins. Nature 496, 215–218.
Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., et al., 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience 1, 18.
Luo, H., Tolar, B.B., Swan, B.K., Zhang, C.L., Stepanauskas, R., Ann Moran, M., Hollibaugh, J.T., 2014. Single-cell genomics shedding light on marine *Thaumarchaeota* diversification. ISME J. 8, 732–736.
Mardis, E.R., 2008. The impact of next-generation sequencing technology on genetics. Trends Genet. 24, 133–141.
Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., et al., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376–380.
Markowitz, V.M., Chen, I.M., Chu, K., Szeto, E., Palaniappan, K., Grechkin, Y., Ratner, A., Jacob, B., Pati, A., Huntemann, M., et al., 2012. IMG/M: the integrated metagenome data management and comparative analysis system. Nucleic Acids Res. 40, D123–D129.
Markowitz, V.M., Chen, I.M., Chu, K., Szeto, E., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., Pagani, I., Tringe, S., et al., 2014. IMG/M 4 version of the integrated metagenome comparative analysis system. Nucleic Acids Res. 42, D568–D573.
Martín-Cuadrado, A.B., López-García, P., Alba, J.C., Moreira, D., Monticelli, L., Strittmatter, A., Gottschalk, G., Rodríguez-Valera, F., 2007. Metagenomics of the deep Mediterranean, a warm bathypelagic habitat. PLoS One 2, e914.
Martín-Cuadrado, A.B., Ghai, R., Gonzaga, A., Rodriguez-Valera, F., 2009. CO dehydrogenase genes found in metagenomic fosmid clones from the deep Mediterranean Sea. Appl. Environ. Microbiol. 75, 7436–7444.
McHardy, A.C., Martín, H.G., Tsirigos, A., Hugenholtz, P., Rigoutsos, I., 2007. Accurate phylogenetic classification of variable-length DNA fragments. Nat. Methods 4, 63–72.
McInerney, M.J., Sieber, J.R., Gunsalus, R.P., 2009. Syntrophy in anaerobic global carbon cycles. Curr. Opin. Biotechnol. 20, 623–632.
Meersman, E., Steensels, J., Mathawan, M., Wittocx, P.J., Saels, V., Struyf, N., Bernaert, H., Vrancken, G., Verstrepen, K.J., 2013. Detailed analysis of the microbial population in Malaysian spontaneous cocoa pulp fermentations reveals a core and variable microbiota. PLoS One 8, e81559.
Metzker, M.L., 2009. Sequencing technologies—the next generation. Nat. Rev. Genet. 11, 31–46.
Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., et al., 2008. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinforma. 9, 386.
Mitra, S., Rupek, P., Richter, D.C., Urich, T., Gilbert, J.A., Meyer, F., Wilke, A., Huson, D.H., 2011. Functional analysis of metagenomes and metatranscriptomes using SEED and KEGG. BMC Bioinforma. 12, S21.
Nakai, R., Abe, T., Takeyama, H., Naganuma, T., 2011. Metagenomic analysis of 0.2-μm-passable microorganisms in deep-sea hydrothermal fluid. Mar. Biotechnol. 13, 900–908.
Namiki, T., Hachiya, T., Tanaka, H., Sakakibara, Y., 2012. MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. Nucleic Acids Res. 40, e155.

Next Generation Sequencing Trends (Clinical Applications, Systems Used, Accreditation Status, Regulated Patient Sample Handling and Other Trends). In: Kalorama Information report, 1 July 2013, p. 159 (Pub ID: KLI5088424).
Noguchi, H., Park, J., Takagi, T., 2006. MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. Nucleic Acids Res. 34, 5623–5630.
Oger, P.M., Jebbar, M., 2010. The many ways of coping with pressure. Res. Microbiol. 161, 799–809.
Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.Y., Cohoon, M., de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R., et al., 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res. 33, 5691–5702.
Picard, A., Daniel, I., 2013. Pressure as an environmental parameter for microbial life—a review. Biophys. Chem. 183, 30–41.
Punta, M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., et al., 2012. The Pfam protein families database. Nucleic Acids Res. 40, D290–D301.
Quaiser, A., Zivanovic, Y., Moreira, D., López-García, P., 2011. Comparative metagenomics of bathypelagic plankton and bottom sediment from the Sea of Marmara. ISME J. 5, 285–304.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D590–D596.
Raes, J., Korbel, J.O., Lercher, M.J., von Mering, C., Bork, P., 2007. Prediction of effective genome size in metagenomic samples. Genome Biol. 8, R10.
Rappé, M.S., Giovannoni, S.J., 2003. The uncultured microbial majority. Annu. Rev. Microbiol. 57, 369–394.
Rappé, M.S., Connon, S.A., Vergin, K.L., Giovannoni, S.J., 2002. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. Nature 418, 630–633.
Reinthaler, T., Winter, C., Herndl, G.J., 2005. Relationship between bacterioplankton richness, respiration, and production in the Southern North Sea. Appl. Environ. Microbiol. 71, 2260–2266.
Reinthaler, T., van Aken, H.M., Herndl, G.J., 2010. Major contribution of autotrophy to microbial carbon cycling in the deep North Atlantic's interior. Deep Sea Res. Part II 57, 1572–1580.
Romanenko, V.I., 1964. Heterotrophic assimilation of CO_2 by bacterial flora of water. Mikrobiologiya 33, 610–613.
Rosen, G.L., Reichenberger, E.R., Rosenfeld, A.M., 2011. NBC: the Naive Bayes Classification tool webserver for taxonomic classification of metagenomic reads. Bioinformatics 27, 127–129.
Schadt, E.E., Turner, S., Kasarskis, A., 2010. A window into third-generation sequencing. Hum. Mol. Genet. 19, R227–R240.
Schippers, A., Kock, D., Höft, C., Köweker, G., Siegert, M., 2012. Quantification of microbial communities in subsurface marine sediments of the Black Sea and off Namibia. Front. Microbiol. 3, 16.
Shanks, A.L., Reeder, M.L., 1993. Reducing microzones and sulfide production in marine snow. Mar. Ecol. Prog. Ser. 96, 43–47.
Shi, Y., Tyson, G.W., Eppley, J.M., DeLong, E.F., 2011. Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. ISME J. 5, 999–1013.
Shi, Y., McCarren, J., Delong, E.F., 2012. Transcriptional responses of surface water marine microbial assemblages to deep-sea water amendment. Environ. Microbiol. 14, 191–206.
Shin, S.C., Ahn do, H., Kim, S.J., Lee, H., Oh, T.J., Lee, J.E., Park, H., 2013. Advantages of single-molecule real-time sequencing in high-GC content genomes. PLoS One 8, e68824.
Simon, C., Daniel, R., 2011. Metagenomic analyses: past and future trends. Appl. Environ. Microbiol. 77, 1153–1161.
Smedile, F., Messina, E., La Cono, V., Tsoy, O., Monticelli, L.S., Borghini, M., Giuliano, L., Golyshin, P.N., Mushegian, A., Yakimov, M.M., 2013. Metagenomic analysis of hadopelagic microbial assemblages thriving at the deepest part of Mediterranean Sea, Matapan–Vavilov Deep. Environ. Microbiol. 15, 167–182.
Sogin, M.L., Morrison, H.G., Huber, J.A., Mark Welch, D., Huse, S.M., Neal, P.R., Arrieta, J.M., Herndl, G.J., 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". Proc. Natl. Acad. Sci. U. S. A. 103, 12115–12120.
Stingl, U., Tripp, H.J., Giovannoni, S.J., 2007. Improvements of high-throughput culturing yielded novel SAR11 strains and other abundant marine bacteria from the Oregon coast and the Bermuda Atlantic Time Series study site. ISME J. 1, 361–371.
Stocker, R., Seymour, J.R., Samadani, A., Hunt, D.E., Polz, M.F., 2008. Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches. Proc. Natl. Acad. Sci. U. S. A. 105, 4209–4214.
Sun, S., Chen, J., Li, W., Altintas, I., Lin, A., Peltier, S., Stocks, K., Allen, E.E., Ellisman, M., Grethe, J., et al., 2011. Community cyberinfrastructure for Advanced Microbial Ecology Research and Analysis: the CAMERA resource. Nucleic Acids Res. 39, D546–D551.
Swan, B.K., Martinez-Garcia, M., Preston, C.M., Sczyrba, A., Woyke, T., Lamy, D., Reinthaler, T., Poulton, N.J., Masland, E.D., Gomez, M.L., et al., 2011. Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. Science 333, 1296–1300.
Tamburini, C., Boutrif, M., Garel, M., Colwell, R.R., Deming, J.W., 2013. Prokaryotic responses to hydrostatic pressure in the ocean—a review. Environ. Microbiol. 15, 1262–1274.
Tatusov, R.L., Koonin, E.V., Lipman, D.J., 1997. A genomic perspective on protein families. Science 278, 631–637.
Tavormina, P.L., Ussler III, W., Joye, S.B., Harrison, B.K., Orphan, V.J., 2010. Distributions of putative aerobic methanotrophs in diverse pelagic marine environments. ISME J. 4, 700–710.
Taylor, B.L., Zhulin, I.B., 1999. PAS domains: internal sensors of oxygen, redox potential, and light. Microbiol. Mol. Biol. Rev. 63, 479–506.
Thrash J.C., Temperton B., Swan B.K., Landry Z.C., Woyke T., DeLong E.F., Stepanauskas R. Giovannoni S.J., Single-cell enabled comparative genomics of a deep ocean SAR11 bathytype, ISME J.
Tringe, S.G., von Mering, C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W., 2005. Comparative metagenomics of microbial communities. Science 308, 554–557.
Varela, M.M., van Aken, H.M., Sintes, E., Herndl, G.J., 2008. Latitudinal trends of *Crenarchaeota* and *Bacteria* in the meso- and bathypelagic water masses of the Eastern North Atlantic. Environ. Microbiol. 10, 110–124.
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., et al., 2004. Environmental genome shotgun sequencing of the Sargasso Sea. Science 304, 66–74.
Vezzi, A., Campanaro, S., D'Angelo, M., Simonato, F., Vitulo, N., Lauro, F.M., Cestaro, A., Malacrida, G., Simionati, B., Cannata, N., et al., 2005. Life at depth: *Photobacterium profundum* genome sequence and expression analysis. Science 307, 1459–1461.
Villanueva, L., Speth, D.R., van Alen, T., Hoischen, A., Jetten, M.S., 2014. Shotgun metagenomic data reveals significant abundance but low diversity of "*Candidatus scalindua*" marine anammox bacteria in the Arabian Sea oxygen minimum zone. Front. Microbiol. 5, 31.
von Mering, C., Hugenholtz, P., Raes, J., Tringe, S.G., Doerks, T., Jensen, L.J., Ward, N., Bork, P., 2007. Quantitative phylogenetic assessment of microbial communities in diverse environments. Science 315, 1126–1130.
Wang, F., Wang, J., Jian, H., Zhang, B., Li, S., Wang, F., Zeng, X., Gao, L., Bartlett, D.H., Yu, J., et al., 2008. Environmental adaptation: genomic analysis of the piezotolerant and psychrotolerant deep-sea iron reducing bacterium *Shewanella piezotolerans* WP3. PLoS One 3, e1937.
Wang, D.Z., Xie, Z.X., Zhang, S.F., 2014. Marine metaproteomics: current status and future directions. J. Proteomics 97, 27–35.
Whitman, W.B., Coleman, D.C., Wiebe, W.J., 1998. Prokaryotes: the unseen majority. Proc. Natl. Acad. Sci. U. S. A. 95, 6578–6583.
Wilmes, P., Bond, P.L., 2004. The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. Environ. Microbiol. 6, 911–920.
Witte, U., Wenzhöfer, F., Sommer, S., Boetius, A., Heinz, P., Aberle, N., Sand, M., Cremer, A., Abraham, W.R., Jørgensen, B.B., et al., 2003. *In situ* experimental evidence of the fate of a phytodetritus pulse at the abyssal sea floor. Nature 424, 763–766.
Wolanin, P.M., Thomason, P.A., Stock, J.B., 2002. Histidine protein kinases: key signal transducers outside the animal kingdom. Genome Biol. 3. http://dx.doi.org/10.1186/gb-2002-3-10-reviews3013 (REVIEWS3013).
Woyke, T., Teeling, H., Ivanova, N.N., Huntemann, M., Richter, M., Gloeckner, F.O., Boffelli, D., Anderson, I.J., Barry, K.W., Shapiro, H.J., et al., 2006. Symbiosis insights through metagenomic analysis of a microbial consortium. Nature 443, 950–955.
Woyke, T., Xie, G., Copeland, A., González, J.M., Han, C., Kiss, H., Saw, J.H., Senin, P., Yang, C., Chatterji, S., et al., 2009. Assembling the marine metagenome, one cell at a time. PLoS One 4, e5299.
Wu, J., Gao, W., Johnson, R.H., Zhang, W., Meldrum, D.R., 2013. Integrated metagenomic and metatranscriptomic analyses of microbial communities in the meso- and bathypelagic realm of North Pacific Ocean. Mar. Drugs 11, 3777–3801.
Wuchter, C., Abbas, B., Coolen, M.J., Herfort, L., van Bleijswijk, J., Timmers, P., Strous, M., Teira, E., Herndl, G.J., Middelburg, J.J., et al., 2006. Archaeal nitrification in the ocean. Proc. Natl. Acad. Sci. U. S. A. 103, 12317–12322.
Xu, K., Ma, B.G., 2007. Comparative analysis of predicted gene expression among deep-sea genomes. Gene 397, 136–142.
Yakimov, M.M., La Cono, V., Denaro, R., D'Auria, G., Decembrini, F., Timmis, K.N., Golyshin, P.N., Giuliano, L., 2007. Primary producing prokaryotic communities of brine, interface and seawater above the halocline of deep anoxic lake L'Atalante, Eastern Mediterranean Sea. ISME J. 1, 743–755.
Yakimov, M.M., La Cono, V., Smedile, F., DeLuca, T.H., Juárez, S., Ciordia, S., Fernández, M., Albar, J.P., Ferrer, M., Golyshin, P.N., et al., 2011. Contribution of crenarchaeal autotrophic ammonia oxidizers to the dark primary production in Tyrrhenian deep waters (Central Mediterranean Sea). ISME J. 5, 945–961.
Yakimov M.M., La Cono V., Smedile F., Crisafi F., Arkadi E., Leonardi M., Decembrini F., Catalfamo M., Bargiela R., Ferrer M., Golyshin P.N., Giuliano L., Heterotrophic bicarbonate assimilation is the main process of *de novo* organic carbon synthesis in hadal zone of the Hellenic Trench, the deepest part of Mediterranean Sea, Environ. Microbiol. (in press).
Zaballos, M., López-López, A., Ovreas, L., Bartual, S.G., D'Auria, G., Alba, J.C., Legault, B., Pushker, R., Daae, F.L., Rodríguez-Valera, F., 2006. Comparison of prokaryotic diversity at offshore oceanic locations reveals a different microbiota in the Mediterranean Sea. FEMS Microbiol. Ecol. 56, 389–405.
Zengler, K., Toledo, G., Rappe, M., Elkins, J., Mathur, E.J., Short, J.M., Keller, M., 2002. Cultivating the uncultured. Proc. Natl. Acad. Sci. U. S. A. 99, 15681–15686.