ACTA ABP BIOCHIMICA POLONICA

Vol. 59, No 4/2012
501–505
on-line at: www.actabp.pl

Review

Metagenomic approach in the investigation of new bioactive compounds in the marine environment

Agnieszka Felczykowska✉, Sylwia Kamila Bloch, Bożena Nejman-Faleńczyk and Sylwia Barańska

Department of Molecular Biology, University of Gdańsk, Gdańsk, Poland

The marine environment is estimated to be one of the most significant sources of biological activity in the world. In the last few decades an increase in the research intensity conducted on marine microorganisms has been observed, which confirms the great potential of these organisms in the field of bioactive compounds' production. In order to efficiently use the natural resources of the marine environment, metagenomics can be applied. This powerful technique allows for efficient screening of microbial biodiversity for bioactive compounds. The primary aim of this review is to present some aspects of the construction of metagenomic libraries, and strategies of screening for novel bioactives in the marine surrounding. This paper also illustrates several examples of the application of metagenomic methods in the discovery of novel enzymes and drugs in various marine environments.

Key words: metagenomics, metagenomic library, function-based screening, marine habitat

Received: 18 November, 2012; revised: 11 December, 2012; accepted: 15 December, 2012; available on-line: 19 December, 2012

BACKGROUND

Oceans and seas that cover more than 70% of the Earth's surface are now one of the most mysterious environments in the world (Zhao, 2011). They were formerly considered to be uninhabited by any life forms (Zobell, 1946). However, recent studies have proven that oceans constitute a rich source of living microorganisms, with cell counts of 10^6–10^9 cells per milliliter (Rheinheimer, 1992; Fenical & Jensen, 2006). Extreme conditions like low temperatures and high pressures, which are present in the marine habitat, support the production of bioactive compounds (bioactives). The ability of marine microorganisms to adapt to the difficult life conditions has resulted in the development of unique biochemical features which may find their application in virtually all industries (Andersen *et al*., 1974). Simultaneously, adjustment to the challenges of the deep sea has led to greater adaptability in both signaling pathways involved in the detection of nutrients and greater metabolic capabilities in utilizing them (Kennedy *et al*., 2010). Furthermore, the marine surfaces are proven to be a significant source of bioactives due to the existing competition between microorganisms associated on the surfaces of eukaryotes (Penesyan *et al.,* 2010).

The last ten years have exposed a new tendency in the exploration of the microorganisms' biodiversity, which is based on culture-independent techniques and the examination of the complex microbial population for basic and applied research (Konig *et al*., 2006). In the study of microbial communities, this approach was developed in order to supplement or replace culture-based technologies, especially since nowadays it is commonly known that only 1–5% of the consortia of living microorganisms can be cultivated under laboratory conditions (Handelsman, 2004). A crucial method of exploring the DNA in uncultivated microorganisms is metagenomics. The microbial DNA isolated from the sample obtained from natural habitats is called the metagenome (Mitra *et al.,* 2010). This sample of DNA theoretically represents all microbes contained in the environmental sample, whose number can vary from several hundred to several thousand of species in a single metagenome (Raes *et al.,* 2007; Gurgui & Piel, 2010; Thomas *et al.,* 2012).

CONSTRUCTION OF A METAGENOMIC LIBRARY

DNA isolation and selection of the vector system

The first step in planning the construction of a metagenomic library is the selection of a vector system. This choice is related to the quality of the extracted environmental DNA, which can be described by means of four parameters: yield, purity, fragment size and representativeness (Morgan *et al.,* 2009; Ekkers *et al.,* 2012). In practice, it is very difficult to achieve satisfactory values of all these factors due to the negative effect that one may inflict on the other. For instance, high yield of DNA is often related to an increased contamination and low average fragment size of the DNA sample. Consequently, small average DNA fragment size limits the analysis of larger operons, in the case of which capital DNA fragments are required (Williamson *et al.,* 2005).

The isolation of DNA from a specific environment, which is destined to be used to construct a metagenomic library, can be performed by using one of the following strategies: direct or indirect extraction of the DNA from the environmental sample. The former approach is based on the isolation of the DNA directly from the sample, while the latter requires a prior selection of the cells of interest (Foote *et al.,* 2012). The indirect method of environmental DNA isolation was proposed by Faegri *et al.* (1977) and applies four steps: dispersion of the en-

✉e-mail: agnieszka.felczykowska@biol.ug.edu.pl

Abbreviations: MDA, multiple displacement amplification; BACs, bacterial artificial chromosomes; CUB, codon usage differences; eDNA, environmental DNA, SIGEX, substrat-induced gene expression; GFP, green fluorescent protein; FACs, fluorescence-activated cell sorter; kb, kilobase.

vironmental sample, cell separation, cell lysis and DNA isolation. However, Roh *et al.* (2006) have proven that direct isolation methods allow for receiving a hundred fold higher amount of DNA on average than the indirect ones.

Occasionally standard techniques of DNA isolation are not sufficiently effective in obtaining the satisfactory yield of DNA for further analysis. In these cases in order to enhance the DNA weight, pre-amplification is required, by means of the multiple displacement amplification (MDA) technique (Woyke *et al.,* 2010). This technique permits whole genome amplification of small amounts of DNA. Spits *et al.* (2006) developed the three-hour procedure of MDA-based-amplification of DNA from a single cell with the use of Phi 29 polymerase to generate 1–2 g of DNA. A possible disadvantage of MDA is the risk of artifact inherence, such as chimeric rearrangements (Neufeld *et al.,* 2008, Gilbert *et al.,* 2010).

The size of the obtained DNA fragment is crucial in the selection of the vector system used in the construction of the metagenomic library. Small-size metagenomic libraries require the DNA sample of about 10 kb and exploit the high-copy-number plasmids with leaky-promoters such as pUC19 (Handelsman, 2004; Godiska, 2009). DNA fragments between 20 to 40 kb require to be inserted in fosmids or cosmids (Hallam *et al.,* 2004; Kim *et al.,* 2006). Bacterial artificial chromosomes (BACs) are commonly used to construct metagenomic libraries with the 100-200 kb DNA fragments (Rondon *et al.,* 2000; Uchiyama & Miyazaki, 2009).

Selection of the host(s)

A standard host in most studies based on metagenomic techniques is *Escherichia coli*. However, it is predicted that the use of *E. coli* as the host in metagenomic library construction allows to express only 40% of genes contained in the eDNA of an average sample (Craig *et al.,* 2010; Parachin & Gorwa-Grauslund, 2011; McMahon *et al.*, 2012). Due to this limitation, an alternative host strain from *Bacillus*, *Pseudomonas* or *Streptomyces* genera can be applied (Lorenz & Eck, 2005; Aavick *et al.,* 2009). There are also few archaeal genera (*Methanococcus*, *Pyrococcus*, *Sulfolobus*, *Thermococcus*), which have been successfully employed in designing of the stable host-vector expression system (Angelov & Liebl, 2010).

Another strategy is the application of broad-host-range vectors, which is related to the concept that a significant number of genes presented in a metagenomic library cannot be expressed in a single host. Thus, this approach may have a positive effect on the detection frequency of genes that confer novel functions (Simon *et al.,* 1983; Wexler *et al.,* 2005). For instance, the application of bacterial genera such as *Pseudomonas*, *Rhizobium* or *Streptomyces*, which have over 15 RNA polymerase σ factors (*Escherichia coli*, which is the most commonly used host in the metagenomic library construction has only seven), may be crucial in the expression of genes that require specialized σ factors (Gabor *et al.*, 2004b).

SCREENING STRATEGY

When the metagenomic library is constructed, a screening strategy should be selected. There are two fundamentally different approaches — activity-based screening (alternatively named functional screening) and sequence-based screening — that may be applied to detect novel bioactives in metagenomic libraries (Fig. 1). Selection of a screening method is connected with various factors such as knowledge about the sequence similarity of the target gene to known genes, type of desired gene product activity, or availability of rapid methods of identifying the transformants of interest (Knietsch *et al.,* 2003; Yun & Ryu, 2005).

Activity-based screening

Functional screening of metagenomic libraries exploits the direct detection of activity of interest. The screen of a metagenomic library is not dependent on the sequence information or sequence similarity of known genes (Lorenzo, 2005). However, this method, being distinct from the sequence-based screening, allows for the discovery of new classes of genes that encode both known and new activities (Suenaga *et al.,* 2007). This approach is strictly bounded by the limitation of eDNA expression in a selected host. All the investigated genes in a given DNA fragment are expected to be expressed independently of its size and structure (Mori *et al.,* 2008). However, there are many factors, which influence an efficient expression: recognition of promoters given in the gene library, regulatory agents of the transcription system of the host, toxicity of gene products, codon usage differences (CUB), correct protein folding, presence of proper initiation factors, or the capacity of the host to secret the gene expression product (Lorenz *et al.*, 2002; Craig *et al.,* 2010). Moreover, activity-based screening methods are directed at the discovery of already known compounds and limited by available assays and detection sensitivity (Kowalchuk *et al,* 2007).

The main advantage of applying this strategy of investigation of novel bioactives is that this approach allows to identify the inherence of a gene product, which cannot be detected by means of the sequence-based method due to the absence of, or insufficient homology to

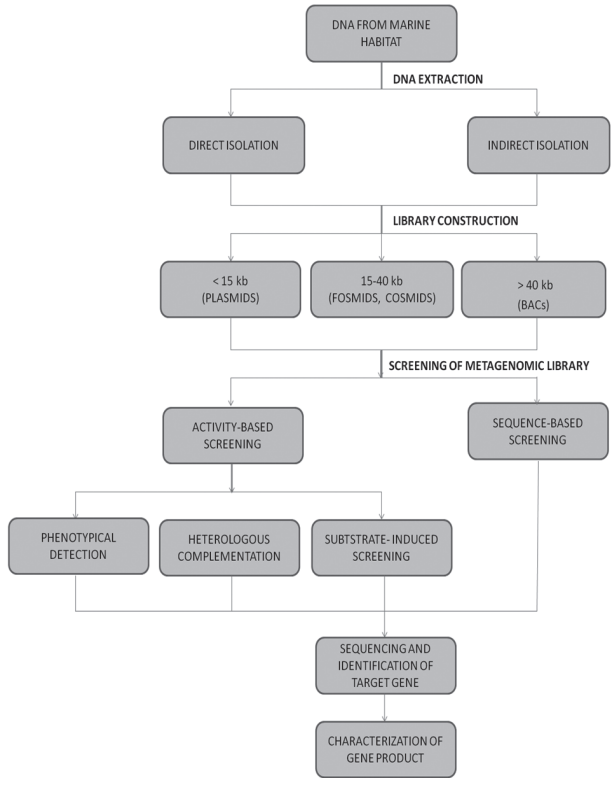

Figure 1. Scheme of a gene library construction and approaches of the discovery of bioactive compounds from the marine environment.

Table 1. Examples of the application of metagenomic studies in the investigation of new bioactive compounds

Target	Origin	Metagenomic library type	Number of clones	Screening approach	Reference
Lipase	Baltic sea sediment (Sweden)	Fosmids	7000	Phenotypical selection	Hårdeman & Sjöling, 2007
Esterase	Neritic sediments of the South China Sea	Plasmids	60000	Phenotypical detection	Peng *et al.*, 2011
Alkaline hydroxylanase	Pacific ocean	Cosmids	–	Heterologous complementation	Xu *et al.*, 2008
Cellulase	Aquatic community (Germany)	Cosmids	3744	Phenotypical selection	Pottkämper *et al.*, 2009
protease	Deep-sea sediment (Korea)	Fosmids	30000	Phenotypical setection	Lee *et al.*, 2007
Serine Protease inhibitor	South China sea water	Plasmids	50000	Sequence-based	Jiang *et al.*, 2011
Chitinase	Arctic Sea	eDNA	–	Sequence-based	LeCleir *et al.*, 2004
Amidase	Soil and enrichment cultures from Marine sediment (Netherlands)	Plasmids	193000	Heterologous complementation	Gabor *et al.*, 2004a
Magnetosome Island Gene cluster	Aquatic sediments (Germany)	Fosmids	5823	Sequence-based	Jogler *et al.*, 2009
Termostable Viral DNA polymerase	Yellowstone hot Springs (USA)	Viral DNA	–	Sequence-based	Schoenfeld *et al.*, 2010

known genes. Furthermore, there is an assurance that the gene product is correctly expressed, which hastens the process of protein characterization (Jones *et al.,* 2007; Park *et al.,* 2007).

The activity-based analysis of eDNA in a metagenomic library might be divided into three different strategies: (1) direct detection of gene products in individual clones; (2) heterologous complementation of host strains or mutants; (3) induced gene expression (Rolf, 2005; Ferrer *et al.,* 2005).

The first approach to the investigation of enzymatic activities of positive clones commonly applies chromophore-containing derivatives of enzyme substrates, which allow for observing the enzymatic reaction. For example, Lämmle *et al.* (2007) have successfully employed the fluorescent methods in the detection of novel lipolytic enzymes and phosphatases.

The second type of metagenomic library screening is based on the use of the host, which requires heterologous complementation by target genes to grow under selective conditions (Carola & Rolf, 2011). This technique is designed for screens of specific bioactive compounds, provided the gene product is in an active form. As an example for this strategy, Riesenfeld *et al.* (2004) successfully applied this method in the discovery of nine clones expressing resistance to aminoglycoside antibiotics, and one expressing resistance to tetracycline.

A different approach is presented in substrate-induced gene expression (SIGEX) screening of gene libraries designed by Uchiyama and Miyazaki (2009). This method is based on the identification of positive clones with a reporter system, which "traps" genes encoding biodegradative pathways located in the "operon-trap" vector. The screen of eDNA clones is connected with the assumption that operons encoding these pathways may be induced by the substrate of the pathway. Reported gene is immediately included downstream of the cloning site for the DNA insert, hence the expression of the insert gene is coupled with that of the reporter gene. The system is addressed especially for screening the catabolic genes which induction is connected with metabolically relevant compounds. Positive clones are detected by a reported signal coming from the fluorescence-activated cell sorter (FACS) in combination with a fluorescent reporter protein, such as the green fluorescent protein (GFP), which production is induced by the substrate of interest (Handelsman, 2005).

Sequence-based screening

Two different strategies of screening of gene libraries are employed in the sequence-based approach: (1) PCR-based and (2) hybridization-based techniques (Weiland *et al.*, 2010). Both methods use the acquirement about conserved DNA sequences present in Databases. Thus, knowledge about the conserved domains of genes is the main limitation of the sequence-based screens. Additionally, the detection of sequence of interest does not guarantee an efficient expression of the target gene. Consequently, this approach is successfully applied in studies based on biodiversity research of environmental samples and identification of exact strains using the bacterial gene encoding 16S rRNA.

However, there are some examples proving that the sequence-based screening of gene libraries may be used successfully in investigating new bioactive compounds. Park *et al.* (2008) created a format of microarray (metagenome microarray technique — MGA), which is arrayed with fosmid library clone DNA samples on a glass slide. They evaluated the MGA applying random prime labeled fluorescent probes prepared using the PCR products of the target gene. The observation was that it is possible to obtain specific hybridization signals only for the DNA fragments that contained the target gene.

METAGENOMICS AND MARINE HABITATS

Microbial communities inhabiting the marine environment have evolved during extended evolutionary processes of physiological adaptation under the influence of a variety of selection pressures. It is estimated that marine habitat is nowadays one of the richest sources of

bioactive compounds due to the diversity of metabolically complex microorganisms (Schloss & Handelsman, 2003). During the last five decades natural products from the marine environment have become a fascinating target for scientists in the discovery of lead compounds applied in medicine and industry (Cooper, 2004; Jones *et al.* 2009; Mayer *et al.,* 2011). Recent examples of marine enzymes that can be potentially applied in different industries, such as the food industry (amylase), chemical industry (lipase, protease, esterase) or agriculture (cellulase), are presented in Table 1.

CONCLUSIONS

Oceans and seas appear to be especially rich in bioactive compounds produced by various marine organisms. Since most of marine microbes are not cultivable, metagenomic approaches can be considered as very promising in discovering previously unknown applicable compounds of important biological activities. It is likely that further research will provide new fascinating discoveries in this field.

Acknowledgments

This work was supported by the European Commission (grant no. 245137, Project acronym: MAREX, large scale integrating project within KBBE-2009-3-2-01) and Ministry of Science and Higher Education, Poland (grant no. 1544/7.PR UE/2010/7).

REFERENCES

Aakvik T, Degnes KF, Dahlsrud R, Schmidt F, Dam R, Yu L, Völker U, Ellingsen TE, Valla S (2009) A plasmid RK2-based broad-host-range cloning vector useful for transfer of metagenomic libraries to a variety of bacterial species. *FEMS Microbiol Lett* **296**: 149–158.

Andersen RJ, Wolfe MS, Faulkner DJ (1974) Autotoxic antibiotic production by a marine *Chromobacterium*. *Mar Biol* **27**: 281–285.

Angelov A, Liebl W (2010) Heterologous gene expression in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Methods Mol Biol* **668**: 109-116.

Carola S, Rolf D (2011) Achievements and new knowledge unraveled by metagenomic approaches. *Appl Microbiol Biotechnol* **85**: 265–276.

Cooper EL (2004) Drug discovery, CAM and natural products. *BMC Complement Altern Med* **1**: 215–217.

Craig JW, Chang FY, Kim JH, Obiajulu SC, Brady SF (2010) Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse *Proteobacteria*. *Appl Environ Microbiol* **5**: 1633–1641.

de Lorenzo V (2005) Problems with metagenomic screening. *Nat Biotechnol* **23**: 1045–1046.

Ekkers DM, Cretoiu MS, Kielak AM, van Elsas JD (2012) The great screen anomaly — a new frontier in product discovery through functional metagenomics. *Appl Microbiol Biotechnol* **93**: 1005–1020.

Faegri A, Torsvik VL, Goksoyr J (1977) Bacterial and fungal activities in soil: separation of bacteria and fungi by a rapid fractionated centrifugation technique. *Soil Biol Biochem* **9**: 105-112.

Fenical, W, Jensen PR (2006) Developing a new resource for drug discovery: Marine actinomycete bacteria. *Nat Chem Biol* **2:** 666–673

Ferrer M, Martínez-Abarca F, Golyshin PN (2005) Mining genomes and 'metagenomes' for novel catalysts. *Curr Opin Biotechnol* **6**: 588–593.

Foote AD, Thomsen PF, Sveegaard S, Wahlberg M, Kielgast J, Kyhn LA, Salling AB, Galatius A, Orlando L, Gilbert TP (2012) Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. *PLoS ONE* **7**: e41781.

Gabor EM, de Vries EJ, Janssen DB (2004a) 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* **6**: 948–958.

Gabor EM, Alkema WBL, Janssen DB (2004b) Quantifyling the accessibility of the metagenome by random expression cloning techniques. *Environ Microbiol* **6**: 879–886.

Gilbert JA, Zhang K, Neufeld JD (2010) Multiple Displacement Amplification. *Handbook of Hydrocarbon and Lipid Microbiology*, Kenneth N, eds. pp 4255–4263. Springer Press.

Godiska R (2008) Bias-free cloning of "unclonable" DNA for simplified genomic finishing. *DNA sequencing III: Dealing with difficult templates*. Kieleczawa J, eds. Jones and Bartlett Publishers, Sudbury, MA.

Gurgui C, Piel J (2010) Metagenomic approaches to identify and isolate bioactive natural products from microbiota of marine sponges. *Methods Mol Biol* **668**: 247–64.

Hallam SJ, Putnam N, Preston CM, Detter JC, Rokhsar D, Richardson PM, DeLong EF (2004) Reverse methanogenesis: testing the hypothesis with environmental genomics. *Science* **305**: 1457–1462.

Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* **68**: 669–685.

Handelsman J (2005) Sorting out metagenome. *Nat Biotechnol* **23**: 506–512.

Hårdeman F, Sjöling S (2007) Metagenomic approach for the isolation of a novel low-temperature-active lipase from uncultured bacteria of marine sediment. *FEMS Microbiol Ecol* **59**: 524–534.

Jiang CJ, Hao ZY, Zeng R, Shen PH, Li JF, Wu B (2011) Characterization of a novel serine protease inhibitor gene from a marine metagenome. *Mar Drugs* **9**: 1487–1501.

Jogler C, Lin W, Meyerdierks A, Kube M, Katzmann E, Flies C, Pan Y, Amann R, Reinhardt R, Schüler D (2009) Towards cloning the magnetotactic metagenome: Identification of magnetosome island gene clusters in uncultivated magnetotactic bacteria from different aquatic sediments. *Appl Environ Microbiol* **75**: 3972–3979.

Jones BV, Sun F, Marchesi JR (2007) Using skimmed milk agar to functionally screen a gut metagenomic library for proteases may lead to false positives *Lett Appl Microbiol* **45:** 418–420.

Jones C, Gu L, Sorrels CM, Sherman DH, Gerwick WH (2009) New tricks from ancient algae: natural products biosynthesis in marine cyanobacteria. *Curr Opin Chem Biol* **13:** 216–223.

Kennedy J, Flemer B, Jackson SA, Lejon DPH, Morrissey JP, O'Gara F, Dobson ADW (2010) Marine metagenomics: new tools for the study and exploitation of marine microbial metabolism. *Mar Drugs* **8**: 608–628.

Kim JY, Choi GS, Kim SB, Yoon GS, Kim YS, Ryu YW (2006) Screening and characterization of a novel esterase from a metagenomic library. *Protein Expr Purif* **45:** 315–323.

Knietsch A, Waschkowitz T, Bowien S, Henne A (2003) Construction and screening of metagenomic libraries derived from enrichment cultures: generation of a gene bank for genes conferring alcohol oxidoreductase activity on *Escherichia coli*. *Appl Environ Microbiol* **69:** 1408-141.

Konig GM, Kehraus S, Seibert SF, Abdel-Lateff A, Muller D (2006) Natural products from marine organisms and their associated microbes. *Chembiochem* **7***:* 229–238.

Kowalchuk GA, Speksnijder AG, Zhang K, Goodman RM, van Veen JA (2007) Finding the needles in the metagenome haystack. *Microb Ecol* **53**: 475–485.

Lämmle K, Zipper H, Breuer M, Hauer B, Buta C, Brunner H, Rupp S (2007) Identification of novel enzymes with different hydrolytic activities by metagenome expression cloning. *J Biotechnol* **127**: 575–592.

LeCleir GR, Buchan A, Hollibaugh JT (2004) Chitinase Gene sequences retrieved from diverse aquatic habitats reveal environment-specific distributions. *Appl Environ Microbiol* **70**: 6977–6983.

Lee DG, Jeon JH, Jang MK, Kim NY, Lee JH, Lee JH, Kim SJ, Kim GD, Lee SH (2007) Screening and characterization of a novel fibrinolytic metalloprotease from a metagenomic library. *Biotechnol Lett* **29**: 465–472.

Lorenz P, Liebeton K, Niehaus F, Eck J (2002) Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Curr Opin Biotechnol* **13**: 572–577.

Lorenz P, Eck J (2005) Metagenomics and industrial applications. *Nature Reviews Microbiology* **3**: 510–516.

McMahon MD, Guan C, Handelsman J, Thomasa MG (2012) Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression. *Appl Environ Microbiol* **6**: 3622–3629

Mayer AMS, Rodríguez AD, Berlinck RGS, Fusetani N (2011) Marine pharmacology in 2007-8: marine compounds with antibacterial, anticoagulant, antifungal, anti-inflammatory, antimalarial, antiprotozoal, antituberculosis, and antiviral activities; Affecting the immune and nervous system, and other miscellaneous mechanisms of action. *Comp Biochem and Physiol* **153**: 191–222.

Mitra S, Gilbert JA, Field D, Huson DH (2010) Comparison of multiple metagenomes using phylogenetic networks based on ecological indices. *ISME J* **4**: 1236–1242.

Morgan JL, Darling AE, Eisen JA (2009) Metagenomic sequencing of an *in vitro*-simulated microbial community. *PLoS ONE* **5**: e10209.

Mori T, Mizuta S, Suenaga H, Miyazaki K (2008) Metagenomic screening for bleomycin resistance genes. *Appl Environ Microbiol* **74**: 6803–6805.

Neufeld JD, Chen Y, Dumont MG, Murrell JC (2008) Marine methylotrophs revealed by stable-isotope probing, multiple displacement amplification and metagenomics. *Environ Microbiol* **10**: 1526–1535.

Parachin NS, Gorwa-Grauslund MF (2011) Isolation of xylose isomerases by sequence- and function-based screening from a soil metagenomic library. *Biotechnol Biofuels* **4**: 9.

Park HJ, Jeon JH, Kang SG, Lee JH, Lee SA, Kim HK (2007) Functional expression and refolding of new alkaline esterase, EM2L8 from deep-sea sediment metagenome. *Protein Expr Purif* **52**: 340–347.

Penesyan A, Kjelleberg S, Egan S (2010) Development of novel drugs from marine surface associated microorganisms. *Mar Drugs* **8**: 438–459.

Peng Q, Zhang Z, Shang M, Wang X, Wang G, Li B, Guan G, Li J, Wang Y (2011) A novel esterase gene cloned from a metagenomic library from neritic sediments of the South China Sea. *Microb Cell Fact* **10**: 95.

Pottkämper J, Barthen P, Ilmberger N, Schwaneberg U, Schenk A, Schulte M, Ignatiev N, Streit WR (2009) Applying metagenomics for the identification of bacterial cellulases that are stable in ionic liquids. *Green Chem* **11**: 957–965.

Raes J, Korbel JO, Lercher MJ, von Mering C, Bork P (2007) Prediction of effective genome size in metagenomic samples. *Genome Biol* **8**: 10.

Rheinheimer G (1992) *Aquatic Microbiology*, pp 56–82. Willey J & Sons, New York, USA.

Riesenfeld CS, Goodman RM, Handelsman J (2004) Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ Microbiol* **6**: 981–989.

Roh C, Villatte F, Kim BG, Schmid RD (2006) Comparative study of methods for extraction and purification of environmental DNA from soil and sludge Samales. *Appl Biochem Biotechnol* **134**: 97–112.

Rolf D (2005) The metagenomics of soil. *Nat Rev Microbiol* **3**: 470–478.

Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C, Tiong CL, Gilman M, Osburne MS, Clardy J, Handelsman J, Goodman RM (2000) Cloning the Soil Metagenome: a Strategy for Accessing the Genetic and Functional Diversity of Uncultured Microorganisms. *Appl Environ Microbiol* **66**: 2541–2547.

Schloss PD, Handelsman J (2003) Biotechnological prospects from metagenomic. *Curr Opin Biotechnol* **14**: 303–310.

Schoenfeld T, Liles M, Wommack KE, Polson SW, Godiska R, Mead D (2010) Functional viral metagenomics and the next generation of molecular tools. *Trends Microbiol* **18**: 20–29.

Simon C, Daniel R (2011) Metagenomic analyses: past and future trends. *Appl Environ Microbiol* **77**: 1153–1161.

Spits C, Le Caignec C, De Rycke M, Van Haute L, Van Steirteghem A, Liebaers I, Sermon K (2006) Whole-genome multiple displacement amplification from single cells. *Nat Protoc* **1**: 1965–1970.

Suenaga H, Ohnuki T, Miyazaki K (2007) Functional screening of a metagenomic library for genes involved in microbial degradation of aromatic compounds. *Environ Microbiol* **9**: 2289–2297.

Thomas T, Gilbert J, Meyer F (2012) Metagenomics — a guide from sampling to data analysis. *Microb Inform Exp* **2**: 3.

Uchiyama T, Miyazaki K (2009) Functional metagenomics for enzyme discovery: challenges to efficient screening. *Curr Opin Biotechnol* **20**: 616–622.

Weiland N, Löscher C, Metzger R, Schmitz R (2010) Construction and screening of marine metagenomic libraries. *Methods Mol Biol* **668**: 51–65.

Wexler M, Bond PL, Richardson DJ, Johnson AW (2005) A wide host-range metagenomic library from a waste water treatment plant yields a novel alcohol/aldehyde dehydrogenase. *Environ Microbiol* **7**: 1917–1926.

Williamson LL, Borlee BR, Schloss PD, Guan CH, Allen HK, Handelsman J (2005) Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor. *Appl Environ Microbiol* **71**: 6335–6344.

Woyke T, Tighe D, Mavromatis K, Clum A, Copeland A, Schackwitz W, Lapidus A, Wu D, McCutcheon JP, McDonald BR, Moran NA, Bristow J, Cheng JF (2010) One bacterial cell, one complete genome. *PLoS ONE* **5**: e10314.

.Xu M, Xiao X, Wang F (2008) Isolation and characterization of alkane hydroxylases from a metagenomic library of Pacific deep-sea sediment. *Extremophiles* **12**: 255–262.

Yun J, Ryu S (2005) Screening for novel enzymes from metagenome and SIGEX, as a way to improve it. *Microb Cell Fact* **4**: 8.

Zhao XQ (2011) Genome-based studies of marine microorganisms to maximize the diversity of natural products discovery for medical treatments. *Evid-Based Complement Alternat Med* **2011**: 384572

Zobell CE (1946) Characteristics of water bacteria. *Sanitarian* **8**: 211–214.