

Molecular biology techniques and applications for ocean sensing

J. P. Zehr, I. Hewson, and P. Moisander

Department of Ocean Sciences, University of California Santa Cruz, 1156 High Street E&MS D446, Santa Cruz, CA 95064, USA

Received: 23 September 2008 – Published in Ocean Sci. Discuss.: 27 November 2008

Revised: 14 April 2009 – Accepted: 15 April 2009 – Published: 8 May 2009

Abstract. The study of marine microorganisms using molecular biological techniques is now widespread in the ocean sciences. These techniques target nucleic acids which record the evolutionary history of microbes, and encode for processes which are active in the ocean today. Molecular techniques can form the basis of remote instrumentation sensing technologies for marine microbial diversity and ecological function. Here we review some of the most commonly used molecular biological techniques. These techniques include the polymerase chain reaction (PCR) and reverse-transcriptase PCR, quantitative PCR, whole assemblage “fingerprinting” approaches (based on nucleic acid sequence or length heterogeneity), oligonucleotide microarrays, and high-throughput shotgun sequencing of whole genomes and gene transcripts, which can be used to answer biological, ecological, evolutionary and biogeochemical questions in the ocean sciences. Moreover, molecular biological approaches may be deployed on ocean sensor platforms and hold promise for tracking of organisms or processes of interest in near-real time.

1 Introduction

The biology of the oceans is recorded in the genetic material of organisms and viruses. The nucleotide sequences of chromosomes include genes that encode biological molecules such as proteins and ribosomes, as well as noncoding regions and genetic elements such as mobile elements, repeated DNA and viral genomes and fragments. The function of organisms is determined by the expression of genes into ribonucleic acids, messenger mRNA (mRNA), which is translated into proteins by the ribosome. With methods developed

over the past two decades, it is possible to characterize organisms and microorganisms on the basis of their molecular features, including the sequence of nucleotides composing the chromosome of an individual organism, even from a single cell.

Although there are methods for deoxyribonucleic and ribonucleic acids (DNA and RNA, respectively), proteins and other biological structures, in general “molecular techniques” are considered to be those that are used to characterize the sequence or molecular structure of nucleic acids of organisms. Two common uses of molecular approaches are to identify organisms and to examine the activity of organisms by assaying gene expression (gene transcription) (Zehr and Hiorns, 1998).

DNA sequences evolve by accumulating mutations induced by natural processes, such as UV radiation (Meador et al., 2009). Some of these mutations are selected through evolution because the resulting changes in protein sequences confer an ecological advantage by increasing ecological fitness. Some gene sequences accumulate fewer mutations than others, since the activity of the proteins or structure of the rRNA has specific structural requirements that are sensitive to substitutions of key amino acids or ribonucleotides. These genes are called conserved genes. Accumulated mutations in conserved genes can be used as taxa specific markers and can be used to compare the evolution of organisms using phylogenetic analysis.

In the marine environment, molecular biological methods have been used to study virtually all trophic levels (Cooksey, 1998; Zehr and Voytek, 1999). In the larger expanses of the ocean, and in the context of oceanic sensors and observing networks, molecular biological approaches provide methods for detecting and characterizing the key players in the biogeochemistry of the ocean: the planktonic microbiota. Molecular methods are particularly useful for studying microbial assemblages in the environment, since most environmental microorganisms have not been cultivated (Azam,



Correspondence to: J. P. Zehr
(zehrj@ucsc.edu)

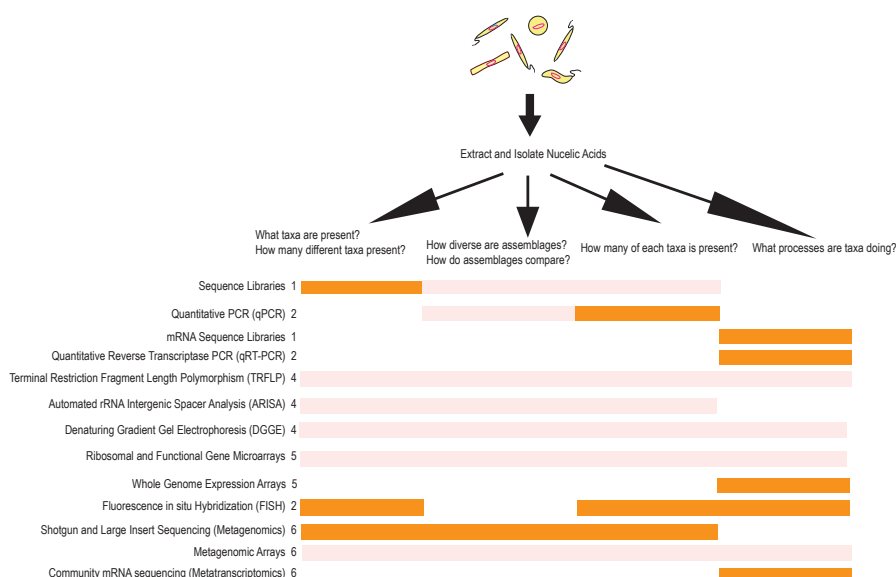


Fig. 1. Ocean scientific questions addressed using molecular biological techniques. Dark grey bars indicate the primary use of the molecular technique, while light grey bars indicate potential applications. The methods are discussed in text (see corresponding numbered sections).

1998). The composition and complexity of microbial assemblages can be identified and compared based on differences in nucleic acid composition and sequence. Analysis of gene transcripts (i.e. mRNAs), can be used to determine whether microorganisms are active, and how they are responding to the environment (Zehr and Hiorns, 1998). Here we review examples of the breadth of molecular techniques currently employed in characterizing marine microorganisms in the ocean.

There are a variety of molecular biological methods that address different types of questions (Fig. 1). Some methods are useful for characterizing overall community assemblage diversity, while others can be used to determine the abundance of microorganisms, or specific enzymatic activities (Fig. 1). The goal of this overview is to acquaint the non-specialist with the breadth of molecular biology techniques, in order to provide the scope and vision for how molecular biological techniques may ultimately be ported to ocean-sensing technology, in situ.

2 Polymerase chain reaction techniques

The development of the polymerase chain reaction (PCR) (Mullis et al., 1986) made it possible to amplify specific genes of interest from very small DNA samples, which facilitated the study of cultivated microorganisms and mixed microbial assemblages. The polymerase chain reaction is an enzymatic method, based on DNA synthesis reactions that enable the geometric amplification of DNA targets using repeated steps of synthesizing DNA. Early applications of PCR targeted specific genes including the universally conserved

small subunit ribosomal RNAs (rRNAs) (Hofle, 1988; Olsen et al., 1986; Pace et al., 1986) in mixed assemblages (Giovannoni et al., 1990) and conserved proteins such as nitrogenase (Zehr and McReynolds, 1989). Subsequently, modifications of the DNA amplification approach have been developed to detect gene expression (mRNA) by reverse-transcription PCR (RT-PCR), and quantitative methods for both DNA and RNA (quantitative PCR or quantitative RT-PCR).

PCR has provided a useful, sensitive and informative approach for studying natural microbial communities. However, there are limitations in application of PCR, including bias introduced by the nonlinear amplification (Polz and Cavanaugh, 1998) which is affected by base pair content (particularly guanosine and cytosine vs adenosine and thymidine content) (Suzuki et al., 1998), the efficiency with which PCR primers bind to their complementary sequence (targets) in the PCR reaction (Suzuki and Giovannoni, 1996), and errors caused by the DNA polymerase enzyme (Barnes, 1992). Additional errors can be introduced due to differential nucleic acid extraction efficiency (Polz et al., 1999) and during DNA sequencing (Acinas et al., 2004).

Typically, PCR is used to amplify rRNA or specific “functional” genes using gene-specific primers, cloning the amplified material into a bacterial host *Escherichia coli* using a plasmid vector, and then growing large numbers of host *E. coli* colonies (called a library) from which individual plasmids can be extracted. The individual DNA molecules can then be sequenced from individual recombinant plasmids (also called the Sanger sequencing). This technique replaced the direct cloning and sequencing approach (Schmidt et al., 1991). The PCR is dependent on designing PCR primers,

which are short single-stranded DNA sequences or oligonucleotides that hybridize to complementary strands of DNA and “prime” the DNA synthesis reaction. PCR primers are designed based upon comparison of available gene sequences from representative organisms, and are usually 16–25 base pair oligonucleotides, which are similar or the same among representative sequences. Since codons for amino acids (three base pair nucleotide sequences) may vary in the third nucleotide base composition yet code for the same amino acid (known as “degeneracy” or “redundancy”), PCR primers can be designed using mixtures of oligonucleotides or bases that hydrogen bond to multiple bases, to create “universal” primers for conserved amino acid sequences that will amplify the same genes from many different microorganisms. This approach has been used for developing PCR primers for “functional” genes, or genes that encode proteins or enzymes that perform an ecologically important function.

PCR has been used to characterize the composition of marine microbial communities, often by targeting rRNA genes. Cloned PCR-amplified 16S rRNA genes have been used to detect the presence of uncultivated Bacteria (Giovannoni et al., 1990) and Archaea (Fuhrman et al., 1992; DeLong, 1992). Clone libraries have revealed that prokaryotic assemblages in seawater are dominated by only a handful of different major groups, however there is a long list of rarer taxa present within assemblages. Amongst dominant taxa, the SAR-11 clade of α -Proteobacteria (Giovannoni, 2004; Giovannoni et al., 1990; Giovannoni and Rappe, 2000), and SAR-202 clade of γ -Proteobacteria (Morris et al., 2004) comprise the dominant fraction of bacterial rRNA clones in surface and deep waters, respectively, which along with the dominance of Archaea in deeper waters (Fuhrman et al., 1993), have been confirmed by hybridization techniques (see next paragraph) (Morris et al., 2002, 2004; Karner et al., 2001). Estimating the richness and diversity of assemblages based upon clone library techniques is problematic since the long tail of the species distribution curve dictates use of non-parametric population sampling techniques, most commonly the Chao1 indicator (Chao, 1987). Because of the historically prohibitive cost of sequencing, creation of sequence libraries large enough to ensure appropriate coverage of microbial assemblages was not generally possible.

The PCR, because of its power in amplifying small amounts of DNA into quantities that can be processed by a variety of techniques, has become a method that underlies many other methods, from RT-PCR to assay mRNA to analyzing thousands of genes on DNA microarrays. It is integrated in many of the techniques described below, particularly those used for diversity analyses.

3 Enumeration of specific microbial targets

Enumeration of individual microbial populations within assemblages has come about through two methods, fluores-

cence in situ hybridization (FISH) and more recently through quantitative PCR. FISH was developed in the early 1990s as a technique for microscopically identifying different groups of microorganisms within assemblages (Amann et al., 1990, 1995). This technique involves permeabilizing prokaryotic cells and hybridizing a fluorescently labeled DNA probe to rRNA or mRNA within the bacterial cells (Daims et al., 2005). The cells are then visualized and can be enumerated with an epifluorescence microscope. There are some biases involved in the FISH approach, notably that not all cells are equally permeable to the hybridization reagents, and that counts of less abundant components of microbial assemblages may not be of high statistical rigor since the absolute number of cells counted is low. The FISH technique has been further expanded to include steps allowing observation of the uptake of specific radioisotopically-labeled compounds (STAR- or MICRO-FISH) (Ouverney and Fuhrman, 1999, 2000; Lee et al., 1999).

Fluorescent substrates and labeled probes have been used to develop methods where the PCR reaction can be monitored in real time, and the PCR product accumulation measured directly by fluorescence (Smith, 2005). In marine applications, quantitative PCR can be used to estimate the abundance of genes per unit volume of seawater. Quantitative PCR (qPCR), unlike sequence libraries and other methods of community analysis (see fingerprinting methods below), is gene-specific and often taxon specific and quantifies the absolute abundance of an organism (specifically its genes). One method used in qPCR reactions uses staining of the double-stranded DNA product with the DNA major groove-binding dye SYBR Green I. The primary drawback of this method is that spurious products and primer dimers can be formed, and lead to overestimates of abundance of the target. The kinetics of “melt curves”, where the amplified products at the end of the qPCR reaction are dissociated slowly and fluorescence changes measured at each degree of dissociation, can be used to distinguish spurious products from the target. More recently, two methods have been published, known as 5'-nuclease (Heid et al., 1996) and hairpin probe hybridization (Tyagi and Kramer, 1996), where the detection of product is based upon hybridization of a probe labeled with fluorophore and quencher to amplification products during each cycle. The nuclease activity of the DNA polymerase degrades the probe and liberates the fluorescence dye from the quencher. The advantages to these chemistries are the very high specificity of the product, since the total matching DNA sequence must be ca. 70–100 bp and match with both probes and primers. In the marine setting, qPCR has been used to estimate the abundance of different phylogenetic groups of organisms based upon rRNA amplification, and different groups of organisms with the ability to conduct particular functions – i.e. amplification of functional genes. One of the first applications of qPCR for detection of phylogenetic groups of marine microorganisms was that of Suzuki et al. (2001) in Monterey Bay, which provided

estimates of the major components of the bacterioplankton community in a coastal environment, without DNA sequencing. In addition, qPCR has been used to estimate the abundance of pathogenic viruses (notably enteric viruses) in seawater samples (Fuhrman et al., 2005; Choi and Jiang, 2005), and has provided a method of detection several thousandfold more sensitive than cultivation-based techniques. Quantitative PCR has also been used for functional genes, including different clades of nitrogen-fixing bacteria targeting dinitrogenase reductase (*nifH*) (Short et al., 2004), ribulose biphosphate carboxylase-oxidase (RuBisCO) in phytoplankton (John et al., 2007) and the photosynthetic reaction center of photoheterotrophs, *pufM* (Schwalbach et al., 2005). These studies have provided the important estimates of functional group abundance in natural samples, since enumeration of functional groups of organisms was previously only possible by using microscopy or cultivation-based techniques which both have biases.

4 Reverse-transcriptase-polymerase chain reaction (RT-PCR)

While DNA-based methods elucidate the genomic potential to carry out particular functions within microbial assemblages, they do not necessarily indicate the active process. Presence of messenger RNA (mRNA) is evidence that genes of interest have been recently transcribed, serving as an essential step towards protein synthesis and an active cellular process. Reverse transcriptase (RT-) PCR involves the conversion of template mRNA into copy DNA (cDNA), which is then used as template material for the PCR, has allowed examination of active processes in environmental samples (Nogales, 2005). RT-PCR, and quantitative RT-PCR (qRT-PCR) have been used to study the presence and expression levels (i.e. abundances of genes and transcripts) of active microorganisms and those conducting specific functions. While qRT-PCR allows enumeration of gene copies in environmental samples, it does not allow enumeration of cells since the number of RNA transcripts per gene and per cell is variable.

RT-PCR and qRT-PCR have been used to study the expression of functional genes, including Rubisco (Wawrik et al., 2002, 2007) and nitrogenase. “Functional” genes is often used to describe genes whose enzymes catalyze a biogeochemical reaction such as carbon fixation, ammonia oxidation or nitrogen fixation, for example. The expression of these genes (i.e. the reading of the gene sequence into a messenger RNA molecule that is used to translate proteins) indicates that the organisms are potentially active in the process. Microorganisms often have genes in their genomes that are not used, and thus, using RT-PCR to determine that the genes are expressed into mRNA can provide information on which organisms are actually involved in observed biogeochemical transformations. The study of nitrogenase (*nifH*) gene expression provides a good example of the kind

of information that RT-PCR can provide. The first studies of *nifH* gene expression in marine bacterioplankton used clone libraries of reverse-transcribed and PCR amplified *nifH* genes to identify active N₂-fixing microorganisms (Zehr et al., 2001; Falcón et al., 2002; Church et al., 2005; Hewson et al., 2007b). In Chesapeake Bay bacterioplankton, it was found that very few of the *nifH* gene types were expressed (Jenkins et al., 2004; Short and Zehr, 2007). Study of the variability in gene expression of different clades of nitrogen-fixing micro-organisms over diel cycles has revealed gene expression patterns that are consistent with measurements of nitrogen fixation in the cyanobacterium *Trichodesmium*, and light- or dark-driven patterns of gene expression for other cultivated and uncultivated groups of N₂-fixing microorganisms (Church et al., 2005; Zehr et al., 2007).

This type of approach can be applied in a number of ways, for example in detecting nitrogen limitation by *ntcA* expression (Lindell et al., 1998), phosphorus stress by phosphorus transporter (*pstS*) gene expression (Dyhrman and Haley, 2006). Such targets may be particularly useful in developing probes for interrogating natural populations on remote instrumentation (see below).

5 “Fingerprinting” methods for diversity studies

Whole community fingerprinting techniques have been used to examine the richness, diversity, and relative assemblage composition of entire assemblages in the marine environment. Similar to sequence library-based approaches, fingerprinting tools provide information on the presence and relative composition of assemblages, where rare components of assemblages (in some cases representing <1% of total template material, Hewson and Fuhrman, 2004b) are detected. While sequence library-based approaches may require thousands of clones to be sequenced to obtain similar resolution, molecular fingerprinting techniques allow rapid and cost-effective estimation of assemblage composition based upon the presence of universal or functional genes. The primary techniques utilized as fingerprinting techniques include terminal restriction fragment length polymorphism (TRFLP) (Avaniss-Aghajani et al., 1994), denaturing gradient gel electrophoresis (DGGE) (Muyzer et al., 1993), automated rRNA intergenic spacer analysis (ARISA) (Fisher and Triplett, 1999), 16S rRNA length heterogeneity PCR (LH-PCR) (Suzuki et al., 1998) and single stranded DNA conformational polymorphism (SSCP) (Lee et al., 1996). Direct sequencing of DGGE fragments can allow identification of different fragments, while fragments in TRFLP, ARISA, and LH-PCR, and perhaps SSCP may be used to identify taxa within fingerprints (Fig. 2).

The TRFLP approach utilizes sequence heterogeneity to separate different taxa within assemblage fingerprints. A conserved gene (which may be rRNA or mRNA) is amplified using PCR with one or two primers with a fluorescent

tag. The amplicons are digested with one or more restriction enzymes (which recognize a particular nucleotide sequence, hence restriction site varies with the sequence of the amplicon) and electrophoretic separation. This technique has been used in study of entire marine bacterial assemblages using 16S rRNA as a target, (Moeseneder et al., 2001; Morris et al., 2005), as well as functional groups such as denitrifying bacteria (Braker et al., 2001), methanogens (Leuders and Friedrich, 2003) and nitrogen-fixers (Hewson and Fuhrman, 2006a). The advantages of this technique include the ability to predict restriction site, and hence restriction fragment length, from public databases and *in silico* restriction digests. The disadvantages include a very restricted number of restriction digest locations in assemblages and hence limited phylogenetic resolution of assemblages.

DGGE relies upon the electrophoretic separation of PCR amplicons based upon their denaturing characteristics, which is determined by their DNA sequence (and more specifically G+C content). In DGGE, PCR is used to amplify targets within mixed assemblages of rRNA or mRNA, where one PCR primer contains a ~50 bp region of G+C, which does not denature under the most denaturing of conditions. The PCR products are then separated on a denaturant gradient gel (typically acrylamide). The amplicons reach the position at which the end of the amplicon is completely denatured (but the G+C region on the primer remains double-stranded) and thus forms a band on the gel. The position at which the amplicons form bands is determined by the sequence of the amplicon (and hence denaturing peculiarity of the sequence). The major advantage of this technique is that the bands on the gel may be excised directly and sequenced using the original PCR primer (i.e. the DNA in the band does not need to be cloned). The disadvantage is that it is very difficult to standardize denaturing conditions, and hence fingerprints in each lane may not be comparable between gels.

ARISA relies upon the length difference in the intergenic region between rRNAs (16S rRNA to 23S RNA) to separate different components within assemblage fingerprints. Like TRFLP, PCR is performed using primers flanking the ITS region, where one primer is fluorescently labeled. The amplified ITS regions are then separated electrophoretically and each ITS length corresponds to a taxonomic unit (analogous but not equal to a taxon). The disadvantage of the technique is that very few genomes or ITS regions have been sequenced, hence it is difficult to assign a particular ITS length to a taxon in fingerprints. However, creation of ITS clone libraries from the same samples, where identification is based on 16S rRNA sequence flanking the ITS, has allowed putative identification of ARISA peaks (Brown et al., 2005). The disadvantage of ARISA is that it can only be applied to rRNAs at present (not functional genes) and that the database of ITS length identities is minimal at present. ARISA has been used to study bacterial assemblage biogeography in sediments (Hewson et al., 2003, 2007a; Luna et al., 2002), microbial mats (Hernandez-Raquet et al., 2006; Leuko et al.,

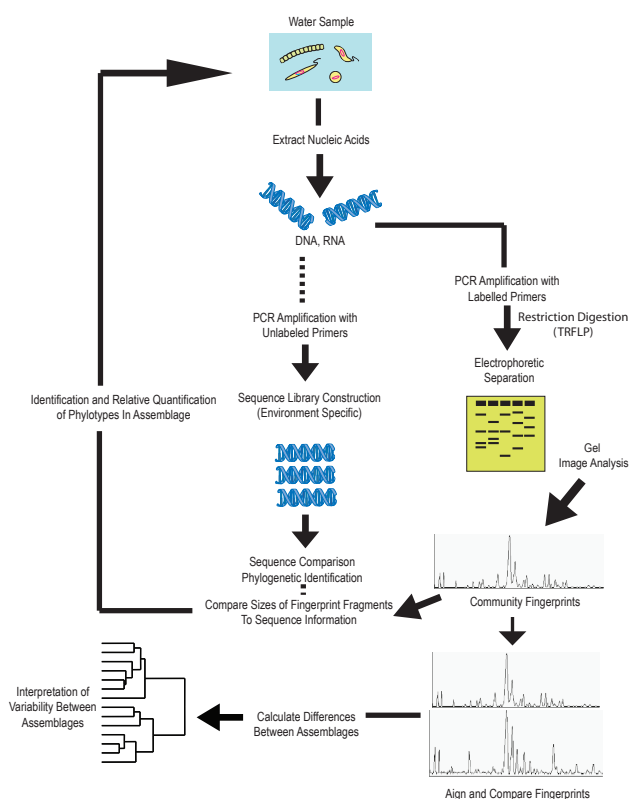


Fig. 2. Comparison of assemblage fingerprints and identification of phylotypes based upon sequence data.

2007) and the water column (Hewson and Fuhrman, 2004a, 2006c; Danovaro and Pusceddu, 2007), where comparisons between assemblage fingerprints has yielded information on the spatial heterogeneity of open ocean bacterioplankton assemblages in surface (Hewson et al., 2006a) and meso- to bathypelagic (Hewson et al., 2006b) assemblages. ARISA has also been used to gauge the impacts of viral pressure upon bacterioplankton assemblages (Hewson et al., 2006c; Hewson and Fuhrman, 2006c; Schwalbach et al., 2004).

LH-PCR examines the length heterogeneity in 16S rRNA PCR amplicons. Because the 16S rRNA is slightly variable in length over hypervariable regions (ca. 50 bp), the presence of different taxa can be examined in fingerprints of rRNA length. This approach was applied in surface waters and observed similar distribution of phylogenetic groups as 16S rRNA sequence libraries (Suzuki et al., 1998).

SSCP relies upon the sequence-driven conformational properties of single stranded 16S rRNA PCR amplicons. These properties result in different migration patterns by gel electrophoresis, and hence organisms with similar sequence comigrate in gels. This approach has been used most recently in the study of sea surface microlayer (Agogue et al., 2005) and Mediterranean plankton (Ghiglione et al., 2005).

6 Microarrays

During the past decade microarray approaches have become increasingly common in microbial ecology. In general, microarrays are arrays of spotted oligonucleotide probes, often on glass slides that can be hybridized to DNA or RNA to identify or characterize genes through hybridization. Microarrays can utilize a variety of platforms and assay chemistries, including arrays spotted on glass slides or membranes, electronic or fiberoptic arrays, and in situ synthesized arrays (such as ones produced by Affymetrix and Nimble-Gen). All microarrays consist of a collection of probes that are designed to match specific microbial target genes. When a sample from a target organism or environment is hybridized with the array, matching nucleotide fragments bind and can be detected by a fluorescent signal. A large variety of probe chemistries and target preparation methods are used affecting the method specificity, sensitivity, and potential for accurate target quantification (Zhou, 2003). Microarrays are divided into different categories based on their design and purpose of use, including (1) whole genome arrays, (2) community genome arrays, (3) phylogenetic arrays and (4) functional gene arrays.

Whole genome information from marine microbes is currently accumulating as a result of ongoing large sequencing projects. These data will allow the design of new microarrays that can be used in full genome expression studies with both laboratory isolates and environment samples. Whole genome microarrays or microarrays constructed with shotgun clones (see below) from microbial isolates or environmental samples can be hybridized with RNA from microbial isolates or environmental transcriptomes (Parro et al., 2007). Linear amplification of RNA is often necessary to reach sufficient detection limit for the mRNA targets. Whole genome microarrays have thus far been applied for marine cyanobacteria *Synechococcus* spp. (Palenik et al., 2003, 2006; Su et al., 2006) and recently a DMSP degrading Proteobacterium *Silicibacter pomeroy* from the *Roseobacter* clade (Burgmann et al., 2007). Genome data and analyses of several marine microbes have been published, including the uncultivated marine Archaeon *Cenarchaeum symbiosum* (Hallam et al., 2006), the anoxygenic anaerobic phototrophic Proteobacterium *Roseobacter denitrificans* (Swingley et al., 2007), the cyanobacterial diazotroph *Crocospaera watsonii* (Zehr et al., 2007), and a marine planctomycete *Pirellula* sp. (Glockner et al., 2003). Many other genome projects near their completion, such as two marine Planctomycetes (Woebken et al., 2007). Additional full genome information has been completed for marine viruses and marine microbial plasmids (Paul et al., 2005, 2007). Genome analyses are highly valuable for microarray development in terms of identifying genome regions of interest and development of optimal microarray probes. Information on genome structure may allow development of partial genome microarrays (Studholme and

Dixon, 2004) with reduced complexity and cost compared to whole genome microarrays.

In addition to the whole genome sequences, partial genome data and analyses are accumulating at an increasing pace from metagenomic and transcriptomic fragments recovered from uncultivated marine micro-organisms (e.g. Frigaard et al., 2006; Suzuki and Béjà, 2007; Vergin et al., 1998). These data provide an enormous resource for development of targeted microarray approaches for detection of expression and distributions of uncultivated micro-organisms in different geographic regions under a range of environmental conditions. Thus far array approaches have been applied to only few oceanic metagenomic datasets. North Pacific Ocean fosmid libraries were spotted on membranes and hybridized with proteorhodopsin targets to screen for presence of proteorhodopsin genes (Frigaard et al., 2006). A similar approach could be useful in other environments if specific gene fragments are sought and sequencing of the entire metagenome or transcriptome is not feasible.

Microarrays have also been applied to microbial diversity studies. The small subunit ribosomal RNA (SSU rRNA) gene sequence has been used to develop phylogenetic microarrays with the goal of detecting microbes representing different phylogenetic groups in the environment. Phylogenetic microarrays have been developed for terrestrial environments (Small et al., 2001; DeSantis et al., 2007), all known sulphate reducers (Loy et al., 2002), and aquatic cyanobacteria (Castiglioni et al., 2004). An electronic microarray (NanoChip) using the SSU rRNA gene sequence for detection of bacteria associated with coastal harmful algal blooms was developed (Barlaan et al., 2007).

Functional gene microarrays may include one or more genes from a variety of micro-organisms, with the goal of targeting diversity or expression of a microbial functional group. These microarrays may include genes encoding key metabolic processes or genes involved in important biogeochemical processes (Zhou, 2003). Functional genes can be used as an approach to address taxonomic composition of microbial communities. Several functional gene microarrays have been designed, but relatively few thus far have been applied to aquatic environments, and even fewer have been used for samples from marine pelagic environments. Examples of functional gene microarrays include one developed to target genes involved in biodegradation of environmental contaminants (Rhee et al., 2004) and another one targeting all known methanotrophs (Bodrossy et al., 2003). Microarrays have been developed to target genes involved in nitrogen cycling (Wu et al., 2001; Taroncher-Oldenburg et al., 2003), primarily in microorganisms that carry out nitrification or denitrification (ammonia oxidation, nitrate reduction). A microarray was also developed for characterization of diazotroph communities in marine environments, based on the *nifH* gene diversity, and the array was applied in studies of communities of nitrogen-fixing micro-organisms in marine microbial mat, estuarine (Moisander et al., 2006, 2007), open

ocean, and coral reef environments (Hewson et al., 2007a, b). Another functional gene microarray targeting marine sediment ammonia-oxidizers has also been developed (Ward et al., 2007a, b). Recently, a functional gene microarray consisting of 24 243 probes covering 10 000 genes involved in cycling of nitrogen, carbon, sulfur, phosphorus among others (He et al., 2007), has been applied in a number of environments, but primarily in soils.

Very recently microarrays have been used to examine the similarity of large fragments of bacterial phylotypes between different locations in the ocean (Rich et al., 2008). The targets on the microarrays were developed around fragments of DNA sequenced as part of a metagenomic survey (see next section) that were closely related to the marine cyanobacterium *Prochlorococcus*. Hybridization with pure culture DNA of several strains of *Prochlorococcus*, and DNA from coastal seawater that was amended with *Prochlorococcus* culture correctly identified the presence of *Prochlorococcus* genes. While field results examining the presence of *Prochlorococcus* in environmental samples has not yet been published, this metagenome array approach holds significant promise in examining gene presence (and expression) of uncultivated microorganisms in marine systems.

The application of microarrays in microbial ecology thus far has been technology and method development driven. The future decade may well witness a surge in marine microbial ecological applications that utilize this technology advancing the knowledge on ecology and physiology of marine micro-organisms.

7 Shotgun cloning, large insert libraries metagenomics and metatranscriptomics

Concern over PCR biases in studies of marine microbial diversity was a strong impetus for development of PCR-independent techniques to study marine microbial diversity. Without amplification of a specific gene of interest, genomic DNA from mixed microbial assemblages needed to be sheared into small fragments, ligated into a vector and then screened for genes of interest. The earliest attempt at this technique by Schmidt and colleagues (Schmidt et al., 1991) used lambda libraries which were screened for 16S rRNA by hybridization prior to sequencing.

Random sequencing of viral DNA was necessary since there are no conserved genes across all viruses, a consequence of the high mutation rate in viral genomes relative to their hosts. As a consequence, a method for randomly sequencing viral DNA was sought. Since large amounts of viral DNA are required for direct sequencing, it was first necessary to amplify viral genome material. This initially took the form of linker-amplified shotgun libraries (LASLs) where PCR priming sites were ligated onto the ends of randomly sheared viral DNA, which was then amplified using PCR and sequenced using Sanger Sequencing. The sequences were

then assembled into contiguous fragments *in silico*. This approach has been used to understand viral composition and diversity in pelagic and benthic compartments of coastal marine waters (Breitbart et al., 2002, 2004). Most recently, direct sequencing of viral DNA has been possible after circularizing and amplifying viral genome fragments using phi29 polymerase, an isothermal and linear amplification employed in genome studies of prokaryotes (Angly et al., 2006).

Because the likelihood of cloning intact genes of interest on small, randomly sheared fragments of microbial DNA, and a desire to obtain further genetic information on the same cloned fragments as universally conserved genes, techniques which aimed to clone much larger fragments of DNA were sought. The application of bacterial artificial chromosomes (BACs) to microbial DNA from the marine environment was first attempted by Stein et al. (1996), where ca. 40 kbp of genetic information was cloned into host *E. coli* using phage transfection. BACs, and later Fosmids (which vary by the method of sheared DNA separation during transfection) have been used to study the genomes of uncultivated microorganisms in marine environments (Béjà et al., 2000; DeLong et al., 2006; Stein et al., 1996). Because the ends of the BACs or Fosmids can be sequenced, this technique may also be used as a method of random sequencing of microbial community DNA, where the sequences obtained may be aligned and overlapping contiguous fragments of DNA sequence assembled into contigs or scaffolds. The application of this type of approach to assemblages of microorganisms is referred to as metagenomics. This has been used most recently by DeLong and colleagues (DeLong et al., 2006) to study the phylogenetic composition and similarity of assemblages at station ALOHA, north of Oahu, Hawaii, over a yearly cycle.

The reduction in cost of Sanger sequencing and the operation of numerous sequencers in series, initially applied to the Human Genome Sequencing project, provided for the first time the ability to sequence extremely high numbers ($>10^6$ individual sequences) of random, sheared, short fragments of microbial DNA from environmental samples. This was first applied simultaneously to microbial DNA from an acid mine drainage site (Tyson et al., 2004) and to bacterioplankton DNA collected at several stations in the Sargasso Sea (Venter et al., 2004). These studies, and later metagenomic studies during the global ocean survey (GOS) project (Rusch et al., 2007) have elucidated high variability in bacterioplankton assemblage structure, in line with observations made using fingerprinting approaches (Hewson and Fuhrman, 2006b), as well as a plethora of new genes and protein families (Yooseph et al., 2007).

Because metagenomics (see Rademaker et al., 2005) elucidates only potential functions, and not necessarily the active utilization of genes in microbial assemblages, there is a need for examination of community transcripts, which has taken the form of random sequencing of mRNAs from assemblages. The first marine metatranscriptomic study utilized

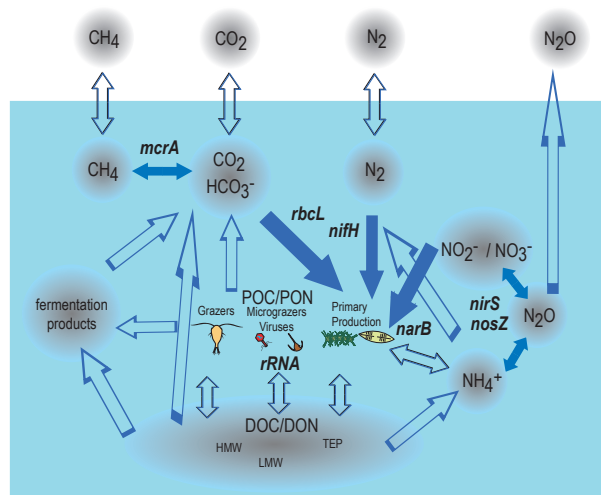


Fig. 3. Examples of dominant nutrient cycling pathways that may be useful targets for molecular detection, including those deployed on environmental sensing platforms. The genes involved in processes are indicated in boldface. *mcrA*=methylcoenzyme reductase A; *rbcL*=ribulose biphosphate carboxylase-oxidase; *nifH*=nitrogenase Fe protein; *narB*=assimilatory nitrate reductase B; *nirS*=nitrite reductase S; *nosZ*=nitrous oxide reductase Z; rRNA=ribosomal RNA.

random amplification via RT-PCR of community RNA extracts which had been treated to remove rRNAs (which comprise >90% of all RNA in an assemblage) to examine diel and tidal influence on estuarine bacterioplankton assemblage expression patterns (Poretsky et al., 2005). The major disadvantage of metatranscriptomic approaches are that sequencing efforts are frequently dominated by rRNAs, since removal via capture hybridization (physical separation) or terminator exonucleic (a way of selectively digesting rRNAs which have different chemical properties than mRNAs) approaches typically remove only a small fraction of rRNAs.

A new type of sequencing which is based upon the fluorescent detection of sequential nucleotide incorporation into amplified DNA was initially developed for use as an alternative to Sanger sequencing (Ronaghi et al., 1996), and later was applied to DNA that is separated onto solid (bead) surfaces in picotiter plates (Margulies et al., 2005). Instead of amplification and the incorporation of terminator dideoxynucleoside triphosphates as in Sanger sequencing, each deoxynucleosidetriphosphate (dNTP) is added in sequence into reactions containing deoxynucleosidetriphosphatase which removes unincorporated dNTPs. The incorporation of the dNTP complementary to the sequenced strand releases pyrophosphate which interacts with adenosine triphosphate (ATP) to produce a flash of light. This is detected and, thus, sequences are determined by the detection of incorporation in relation to the dNTP that is added to the reactions. Unlike Sanger sequencing, which produces sequences

ca. 600–1000 bp in length, the length of sequences is limited to <500 bp. However, because sequencing occurs on a massively parallel scale in picotiter plates, a typical pyrosequencing run generates 100 000–300 000 individual sequences. Several other technologies have also been developed for high throughput sequencing.

To date pyrosequencing has been applied in marine environments to estimate the diversity and identification of microbial assemblages. An early application (Sogin et al., 2006) sequenced a highly variable region within the ribosomal RNA of hydrothermal vent communities and found diversity of assemblages was underestimated in previous studies by 2–3 orders of magnitude. The first shotgun pyrosequencing approach examined virus assemblages in several oceanic regions (Angly et al., 2006) and found a large proportion of viral assemblage genetic composition was not identifiable by comparison to existing sequence data, but that viroplankton assemblages were comprised of very similar species despite large geographic separation. Finally, pyrosequencing technology has recently been applied to coral surface microbial communities (Wegley et al., 2007), where predominately heterotrophs, and sequence data from fungi to viruses were recovered. Pyrosequencing technology holds promise to allow a much larger marine microbial sequence database to be constructed, allowing microbial ecologists to address biogeographical and ecological questions that were not possible with limited numbers of sequences.

8 Summary: applications to sensor technology

It is now possible to develop in situ molecular biological sensors for microorganisms in the environment (Paul et al., 2007). Ultimately, many molecular biological approaches may be applied in some type of remote sensing application. For example, microarray approaches have been proposed for detection and monitoring of pathogens such as *Vibrio cholerae* in coastal environments (Stine et al., 2003). The status of biological sensors is discussed in another chapter (Scholin et al., 2009). Probably the first technologies to be applied will be based on hybridization and the polymerase chain reaction. These techniques facilitate detection and quantification of species, with the potential for detecting organisms with key biochemical functions by their genes or gene expression.

In situ biological sensors require extensive technology development, and development of molecular applications that can be implemented in situ. A relatively small number of probes can be implemented on current technologies, since size and power consumption limit the capacity of remote instrumentation. Probe arrays for hybridization to a suite of probes for microorganisms can be used to detect specific microorganisms or monitor microbial diversity. Gene expression assays using linear amplification (NASBA),

hybridization technologies, or quantitative PCR are likely to be successfully implemented in the near future.

A primary question is how such instruments can best provide information useful for sensing ocean ecosystems. Probes for harmful algal bloom species are an example of an application with clear applied relevance to ocean ecology (Casper et al., 2004). Given the limited number of probe targets that can be implemented with current technology, discussions of useful targets for monitoring ocean ecosystem function are needed.

Microbial components of oceanic systems are the dominant species and ecotypes, and the species involved in biogeochemical cycles. These key components may be useful as targets for monitoring for changes in ocean communities, including regime shifts and dynamics in relation to physical-chemical forcing. Species that may be useful targets are ecotypes of the major oceanic primary producers *Prochlorococcus* and *Synechococcus*, or major lineages of eukaryotic phytoplankton, including the picoeukaryotes. Key ecosystem biogeochemical components are carbon fixation and decomposition, nitrogen fixation and nitrogen assimilation, ammonia oxidation and denitrification (Fig. 3). Targets for these processes could be designed to detect presence (genes) or activity (gene expression as a proxy), by developing probes for specific species, or generic group-level probes. Information on time-series or spatial distribution of these genes could inform decisions on selection of ecosystem probes for implementation on remote instrumentation. Much work needs to be done to select targets that provide useful information on ecosystem structure or function, determine how the spatial and temporal distribution of genes and gene expression relate to ecosystem dynamics, and to design appropriate methods for detection and quantification (Zehr, 2009). Once targets have been identified, and the abundance and dynamics known, appropriate molecular assays can be selected and developed for ocean sensing technology.

Edited by: G. Griffiths

References

- Acinas, S. G., Klepac-Ceraj, V., Hunt, D. E., Pharino, C., Ceraj, I., et al.: Fine-scale phylogenetic architecture of a complex bacterial community, *Nature*, 430, 551–554, 2004.
- Agogue, H., Casamayor, E. O., Bourrain, M., Obernosterer, I., Joux, F., et al.: A survey on bacteria inhabiting the sea surface microlayer of coastal ecosystems, *FEMS Microbiol. Ecol.*, 54, 269–280, 2005.
- Amann, R. I., Krumholz, L., and Stahl, D. A.: Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic and environmental studies in microbiology, *J. Bacteriol.*, 172, 762–770, 1990.
- Amann, R. I., Ludwig, W., and Schleifer, K.-H.: Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation, *Microbiol. Rev.*, 59, 143–169, 1995.
- Angly, F. E., Felts, B., Breitbart, M., Salamon, P., Edwards, R. A., et al.: The marine viromes of four oceanic regions, *PLoS Biology*, 4, 2121–2131, 2006.
- Avaniss-Aghajani, E., Jones, K., Chapman, D., and Brunk, C.: A molecular technique for identification of bacteria using small subunit ribosomal RNA sequences, *Biotechniques*, 17, 144–149, 1994.
- Azam, F.: Microbial control of oceanic carbon flux: The plot thickens, *Science*, 280, 694–696, 1998.
- Barlaan, E. A., Furukawa, S., and Takeuchi, K.: Detection of bacteria associated with harmful algal blooms from coastal and microcosm environments using electronic microarrays, *Environ. Microbiol.*, 9, 690–702, 2007.
- Barnes, W. M.: The fidelity of *Taq* polymerase catalyzing PCR is improved by an N-terminal deletion, *Gene*, 112, 29–35, 1992.
- Béjà, O., Aravind, L., Koonin, E. V., Suzuki, M. T., Hadd, A., et al.: Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea, *Science*, 289, 1902–1906, 2000.
- Bodrossy, L., Stralis-Pavese, N., Murrell, J. C., Radajewski, S., Weilharter, A., et al.: Development and validation of a diagnostic microbial microarray for methanotrophs, *Environ. Microbiol.*, 5, 566–582, 2003.
- Braker, G., Ayala-del-Rio, H. L., Devol, A. H., Fesefeldt, A., and Tiedje, J. M.: Community structure of denitrifiers, bacteria, and archaea along redox gradients in Pacific northwest marine sediments by terminal restriction fragment length polymorphism analysis of amplified nitrite reductase (*nirS*) and 16S rRNA genes, *Appl. Environ. Microb.*, 67, 1893–1901, 2001.
- Breitbart, M., Segall, A., Mead, D., Azam, F., and Rohwer, F.: Genomic analysis of an uncultured marine viral community, *EOS Transactions*, 83, OS111-02, 2002.
- Breitbart, M., Felts, B., Kelley, S., Mahaffy, J. M., Nulton, J., et al.: Diversity and population structure of a near-shore marine-sediment viral community, *P. Roy. Soc. Lond. B Bio.*, 271, 565–574, 2004.
- Brown, M. V., Schwalbach, M. S., Hewson, I., and Fuhrman, J. A.: Coupling 16S-ITS rDNA clone libraries and automated ribosomal intergenic spacer analysis to show marine microbial diversity: development and application to a time series, *Environ. Microbiol.*, 7, 1466–1479, 2005.
- Burgmann, H., Howard, E. C., Ye, W. Y., Sun, F., Sun, S. L., et al.: Transcriptional response of *Silicibacter pomeroyi* DSS-3 to dimethylsulfoniopropionate (DMSP), *Environ. Microbiol.*, 9, 2742–2755, 2007.
- Casper, E. T., Paul, J. H., Smith, M. C., and Gray, M.: Detection and quantification of the red tide dinoflagellate *Karenia brevis* by real-time nucleic acid sequence-based amplification, *Appl. Environ. Microb.*, 70, 4727–4732, 2004.
- Castiglioni, B., Rizzi, E., Frosini, A., Sivonen, K., Rajaniemi, P., et al.: Development of a universal microarray based on the ligation detection reaction and 16S rRNA gene polymorphism to target diversity of cyanobacteria, *Appl. Environ. Microb.*, 70, 7161–7172, 2004.
- Chao, A.: Non-parametric estimation of the number of classes in a population, *Scand. J. Stat.*, 11, 265–270, 1984.
- Choi, S. and Jiang, S. C.: Real-time PCR quantification of human adenoviruses in urban rivers indicates genome prevalence but low infectivity, *Appl. Environ. Microb.*, 71, 7426–7433, doi:10.1128/AEM.71.11.7426-7433.2005, 2005.
- Church, M. J., Jenkins, B. D., Karl, D. M., and Zehr, J. P.: Vertical distributions of nitrogen-fixing phylotypes at Stn ALOHA in the

- oligotrophic north Pacific Ocean, *Aquat. Microb. Ecol.*, 38, 3–14, 2005.
- Cooksey, K. E.: *Molecular Approaches to the Study of the Ocean*, Chapman & Hall, London, UK, 549 pp., 1998.
- Daims, H., Stoecker, K., and Wagner, M.: Fluorescence *in situ* hybridization for the detection of prokaryotes, in: *Advanced Methods in Molecular Microbial Ecology*, edited by: Osborn, A. M. and Smith, C. J., Taylor & Francis, New York, USA, 213–239, 2005.
- Danovaro, R. and Pusceddu, A.: Biodiversity and ecosystem functioning in coastal lagoons: Does microbial diversity play any role?, *Estuarine, Coastal and Shelf Science*, 75, 4–12, 2007.
- DeLong, E. F.: Archaea in coastal marine environments, *P. Natl. Acad. Sci. USA*, 89, 5685–5689, 1992.
- DeLong, E. F., Preston, C. M., Mincer, T., Rich, V., Hallam, S. J., et al.: Community genomics among stratified microbial assemblages in the ocean's interior, *Science*, 311, 496–503, 2006.
- DeSantis, T. Z., Brodie, E. L., Moberg, J. P., Zubietta, I. X., Piceno, Y. M., et al.: High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment, *Microb. Ecol.*, 53, 371–383, 2007.
- Dyhrman, S. T. and Haley, S. T.: Phosphorus scavenging in the unicellular marine diazotroph *Crocospaera watsonii*, *Appl. Environ. Microb.*, 72, 1452–1458, 2006.
- Falcón, L. I., Cipriano, F., Chistoserdov, A. Y., and Carpenter, E. J.: Diversity of diazotrophic unicellular cyanobacteria in the tropical North Atlantic Ocean, *Appl. Environ. Microb.*, 68, 5760–5764, 2002.
- Fisher, M. M. and Triplett, E. W.: Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities, *Appl. Environ. Microb.*, 65, 4630–4636, 1999.
- Frigaard, N. U., Martinez, A., Mincer, T. J., and DeLong, E. F.: Proteorhodopsin lateral gene transfer between marine planktonic bacteria and archaea, *Nature*, 439, 847–850, 2006.
- Fuhrman, J. A., McCallum, K., and Davis, A. A.: Novel major archaeobacterial group from marine plankton, *Nature*, 356, 148–149, 1992.
- Fuhrman, J. A., McCallum, K., and Davis, A. A.: Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific Oceans, *Appl. Environ. Microb.*, 59, 1294–1302, 1993.
- Fuhrman, J. A., Noble, R. T., and Liang, X.: Rapid detection of enteroviruses in small volumes of natural waters by real-time Quantitative Reverse Transcriptase PCR, *Appl. Environ. Microb.*, 71, 4523–4530, 2005.
- Ghiglione, J. F., Larcher, M., and Lebaron, P.: Spatial and temporal scales of variation in bacterioplankton community structure in the NW Mediterranean Sea, *Aquat. Microb. Ecol.*, 40, 229–240, 2005.
- Giovannoni, S. and Rappe, M.: Evolution, diversity, and molecular ecology of marine prokaryotes, in: *Microbial Ecology of the Oceans*, edited by: Kirchman, D. L., Wiley-Liss, New York, USA, 47–84, 2000.
- Giovannoni, S. J., Britschgi, T. B., Moyer, C. L., and Field, K. G.: Genetic diversity in Sargasso Sea bacterioplankton, *Nature (London)*, 345, 60–63, 1990.
- Giovannoni, S. J.: Oceans of bacteria, *Nature*, 430, 515–516, 2004.
- Glockner, F. O., Kube, M., Bauer, M., Teeling, H., Lombardot, T., et al.: Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1, *P. Natl. Acad. Sci. USA*, 100, 8298–8303, 2003.
- Hallam, S. J., Konstantinidis, K. T., Putnam, N., Schleper, C., Watanabe, Y., et al.: Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*, *P. Natl. Acad. Sci. USA*, 103, 18296–18301, 2006.
- He, Z., Gentry, T. J., Schadt, C. W., Wu, L., Liebich, J., et al.: Geochip: A comprehensive microarray for investigating biogeochemical, ecological and environmental processes, *ISME J*, 1, 67–77, doi:10.1038/ismej.2007.2, 2007.
- Heid, C. A., Stevens, J., Livak, K. J., and Williams, P. M.: Real time quantitative PCR, *Genome Res.*, 6, 986–994, 1996.
- Hernandez-Raquet, G., Budzinski, H., Caumette, P., Dabert, P., Le Ménach, K., et al.: Molecular diversity studies of bacterial communities of oil polluted microbial mats from the Etang de Berre (France), *FEMS Microb. Ecol.*, 58, 550–562, doi:10.1111/j.1574-6941.2006.00187.x, 2006.
- Hewson, I., Vargo, G. A., and Fuhrman, J. A.: Bacterial diversity in shallow oligotrophic marine benthos and overlying waters: Effects of virus infection, containment and nutrient enrichment, *Microb. Ecol.*, 46, 322–336, 2003.
- Hewson, I. and Fuhrman, J. A.: Richness and diversity of bacterioplankton species along an estuarine gradient in Moreton Bay, Australia, *Appl. Environ. Microb.*, 70, 3425–3433, 2004a.
- Hewson, I. and Fuhrman, J. A.: Improved strategy for comparing microbial assemblage fingerprints, *Microb. Ecol.*, 51, 147–153, 2006a.
- Hewson, I. and Fuhrman, J. A.: Spatial and vertical biogeography of coral reef sediment bacterial and diazotroph communities, *Mar. Ecol.-Prog. Ser.*, 306, 79–86, 2006b.
- Hewson, I. and Fuhrman, J. A.: Viral impacts upon marine bacterioplankton assemblage structure, *Journal of the Marine Biological Association of the United Kingdom*, 86, 577–589, 2006c.
- Hewson, I., Steele, J. A., Capone, D. G., and Fuhrman, J. A.: Temporal and spatial scales of variation in bacterioplankton assemblages of oligotrophic surface waters, *Mar. Ecol.-Prog. Ser.*, 311, 67–77, 2006a.
- Hewson, I., Steele, J. A., Capone, D. G., and Fuhrman, J. A.: Remarkable heterogeneity in meso- and bathypelagic bacterioplankton assemblage composition, *Limnol. Oceanogr.*, 51, 1274–1283, 2006b.
- Hewson, I., Winget, D. M., Williamson, K. E., Fuhrman, J. A., and Wommack, K. E.: Viral and bacterial assemblage covariance in oligotrophic waters of the West Florida Shelf (Gulf of Mexico), *J. Mar. Biol. Assoc. UK*, 86, 591–603, 2006c.
- Hewson, I., Moisander, P. H., Achilles, K. M., Carlson, C. A., Jenkins, B. D., et al.: Characteristics of diazotrophs in surface to abyssopelagic waters of the Sargasso Sea, *Aquat. Microb. Ecol.*, 46, 15–30, 2007a.
- Hewson, I., Moisander, P. H., Morrison, A. E., and Zehr, J. P.: Diazotrophic bacterioplankton in a coral reef lagoon: phylogeny, diel nitrogenase expression and response to phosphate enrichment, *ISME J*, 1, 78–91, 2007b.
- Hofle, M. G.: Identification of bacteria by low molecular weight RNA profiles: a new chemotaxonomic approach, *J. Microbiol. Meth.*, 8, 235–248, 1988.
- Jenkins, B. D., Steward, G. F., Short, S. M., Ward, B. B., and Zehr, J. P.: Fingerprinting diazotroph communities in the Chesapeake

- Bay by using a DNA microarray, *Appl. Environ. Microb.*, 70, 1767–1776, 2004.
- John, D., Patterson, S., and Paul, J.: Phytoplankton-group specific quantitative polymerase chain reaction assays for RuBisCo mRNA transcripts in seawater, *Mar. Biotechnol.*, 9, 747–759, 2007.
- Karner, M. B., DeLong, E. F., and Karl, D. M.: Archaeal dominance in the mesopelagic zone of the Pacific Ocean, *Nature*, 409, 507–510, 2001.
- Lee, D. H., Zo, Y. G., and Kim, S. J.: Nonradioactive method to study genetic profiles of natural bacterial communities by PCR-single-strand-conformation polymorphism, *Appl. Environ. Microb.*, 62, 3112–3120, 1996.
- Lee, N., Nielsen, P. H., Andreasen, K. H., Juretschko, S., Nielsen, J. L., et al.: Combination of fluorescent in situ hybridization and microautoradiography – a new tool for structure-function analyses in microbial ecology, *Appl. Environ. Microb.*, 65, 1289–1297, 1999.
- Leuders, T. and Friedrich, M. W.: Evaluation of PCR amplification bias by terminal restriction fragment length polymorphism analysis of small-subunit rRNA and *mcrA* genes by using defined template mixtures of methanogenic pure cultures and soil DNA extracts, *Appl. Environ. Microb.*, 69, 320–326, 2003.
- Leuko, S., Goh, F., Allen, M. A., Burns, B. P., Walter, M. R., et al.: Analysis of intergenic spacer region length polymorphisms to investigate the halophilic archaeal diversity of stromatolites and microbial mats, *Extremophiles*, 11, 203–210, 2007.
- Lindell, D., Padan, E., and Post, A. F.: Regulation of *ntcA* expression and nitrite uptake in the marine *Synechococcus* sp. Strain WH 7803, *J. Bacteriol.*, 180, 1878–1886, 1998.
- Loy, A., Lehner, A., Lee, N., Adamczyk, J., Meier, H., et al.: Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment, *Appl. Environ. Microb.*, 68, 5064–5081, 2002.
- Luna, G. M., Manini, E., and Danovaro, R.: Large fraction of dead and inactive bacteria in coastal marine sediments: Comparison of protocols for determination and ecological significance, *Appl. Environ. Microb.*, 68, 3509–3513, 2002.
- Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., et al.: Genome sequencing in microfabricated high-density picolitre reactors, *Nature*, 437, 376–380, 2005.
- Meador, J. A., Baldwin, A. J., Catala, P., Jeffrey, W. H., Joux, F., Moss, J. A., Pakulski, J. D., Stevens, R., and Mitchell, D. L.: Sunlight-induced DNA damage in marine micro-organisms collected along a latitudinal gradient from 70° N to 68° S, *Photochem. Photobiol.*, 85, 412–420, 2009.
- Moeseneder, M. M., Winter, C., and Herndl, G. J.: Horizontal and vertical complexity of attached and free-living bacteria of the eastern Mediterranean Sea, determined by 16S rDNA and 16S rRNA fingerprints, *Limnol. Oceanogr.*, 46, 95–107, 2001.
- Moisander, P. H., Shiue, L., Steward, G. F., Jenkins, B. D., Bebout, B. M., et al.: Application of a *nifH* oligonucleotide microarray for profiling diversity of N₂-fixing microorganisms in marine microbial mats, *Environ. Microbiol.*, 8, 1721–1735, 2006.
- Moisander, P. H., Morrison, A. E., Ward, B. B., Jenkins, B. D., and Zehr, J. P.: Spatial-temporal variability in diazotroph assemblages in Chesapeake Bay using an oligonucleotide *nifH* microarray, *Environ. Microbiol.*, 9, 1823–1835, 2007.
- Morris, R. M., Rappe, M. S., Connon, S. A., Vergin, K. L., et al.: SAR11 clade dominates ocean surface bacterioplankton communities, *Nature*, 420, 806–810, 2002.
- Morris, R. M., Rappe, M. S., Urbach, E., Connon, S. A., and Giovannoni, S. J.: Prevalence of the *Chloroflexi*-related SAR202 bacterioplankton cluster throughout the mesopelagic zone and deep ocean, *Appl. Environ. Microb.*, 70, 2836–2842, 2004.
- Morris, R. M., Vergin, K. L., Cho, J.-C., Rappé, M. S., Carlson, C. A., et al.: Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-series study site, *Limnol. Oceanogr.*, 50, 1687–1696, 2005.
- Mullis, K., Faloona, F., Scharf, S., Saki, R., Horn, G., et al.: Specific enzymatic amplification of DNA *in vitro*: The polymerase chain reaction, *Cold Spring Harbor Symposia on Quantitative Biology*, 51, 263–273, 1986.
- Muyzer, G., De Waal, E. C., and Uitterlinden, A. G.: Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA, *Appl. Environ. Microb.*, 59, 695–700, 1993.
- Nogales, B.: RT-PCR and mRNA expression analysis of functional genes, in: *Molecular Microbial Ecology*, edited by: Osborn, A. M. and Smith, C. J., Taylor & Francis, New York, USA, 135–149, 2005.
- Olsen, G. J., Lane, D. J., Giovannoni, S. J., and Pace, N. R.: Microbial ecology and evolution: A ribosomal RNA approach, *Annu. Rev. Microbiol.*, 40, 337–365, 1986.
- Ouverney, C. C. and Fuhrman, J. A.: Combined microautoradiography-16S rRNA probe technique for determination of radioisotope uptake by specific microbial cell types in situ, *Appl. Environ. Microb.*, 65, 1746–1752, 1999.
- Ouverney, C. C. and Fuhrman, J. A.: Marine planktonic archaea take up amino acids, *Appl. Environ. Microb.*, 66, 4822, 4829–4833, 2000.
- Pace, N. R., Stahl, D. A., Lane, D. J., and Olsen, G. J.: The analysis of natural microbial populations by ribosomal RNA sequences, *Adv. Microb. Ecol.*, 9, 1–55, 1986.
- Palenik, B., Brahamsha, B., Larimer, F. W., Land, M., Hauser, L., et al.: The genome of a motile marine *Synechococcus*, *Nature*, 424, 1037–1042, 2003.
- Palenik, B., Ren, Q., Dupont, C. L., Myers, G. S., Heidelberg, J. F., et al.: Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment, *P. Natl. Acad. Sci. USA*, 103, 13555–13559, 2006.
- Parro, V., Moreno-Paz, M., and Gonzalez-Toril, E.: Analysis of environmental transcriptomes by DNA microarrays, *Environ. Microbiol.*, 9, 453–464, 2007.
- Paul, D., Pandey, G., Pandey, J., and Jain, R. K.: Accessing microbial diversity for bioremediation and environmental restoration, *Trends Biotechnol.*, 23, 135–142, 2005.
- Paul, J., Scholin, C., Van den Engh, G., and Perry, M. J.: *In situ* instrumentation, *Oceanography*, 20, 70–78, 2007.
- Polz, M. and Cavanaugh, C. M.: Bias in template-to-product ratios in multitemplate PCR, *Appl. Environ. Microb.*, 64, 3724–3730, 1998.
- Polz, M. F., Harbison, C., and Cavanaugh, C. M.: Diversity and heterogeneity of epibiotic bacterial communities on the marine nematode *Eubostrius diana*, *Appl. Environ. Microb.*, 65, 4271–4275, 1999.
- Poretzky, R. S., Bano, N., Buchan, A., LeClerc, G., Kleikemper, J.,

- et al.: Analysis of microbial gene transcripts in environmental samples, *Appl. Environ. Microb.*, 71, 4121–4126, 2005.
- Rademaker, J. L. W., Aarts, H. J. M., and Vinuesa, P.: Molecular typing of environmental isolates, in: *Molecular Microbial Ecology*, edited by: Osborn, A. M. and Smith, C. J., Taylor & Francis, New York, USA, 97–134, 2005.
- Rapley, R. and Heptinstall, J.: UV spectrophotometric analysis of ribonucleic acids, in: *RNA Isolation and Characterization Protocols*, *Methods in Molecular Biology*, edited by: Rapley, R. and Manning, D. L., Humana Press Inc., Totowa, New Jersey, USA, *Method. Mol. Cell. Biol.*, 86, 65–68, 1998.
- Rhee, S. K., Liu, X. D., Wu, L. Y., Chong, S. C., Wan, X. F., et al.: Detection of genes involved in biodegradation and biotransformation in microbial communities by using 50-mer oligonucleotide microarrays, *Appl. Environ. Microb.*, 70, 4303–4317, 2004.
- Rich, V. I., Konstantinidis, K., and DeLong, E. F.: Design and testing of “genome-proxy” microarrays to profile marine microbial communities, *Environ. Microbiol.*, 10, 506–521, 2008.
- Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M., and Nyren, P.: Real-time DNA sequencing using detection of pyrophosphate release, *Anal. Biochem.*, 242, 84–89, 1996.
- Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., et al.: *The Sorcerer II* global ocean sampling expedition: Northwest Atlantic through Eastern Tropical Pacific, *PLoS Biology*, 5, 398–431, 2007.
- Schmidt, T. M., DeLong, E. F., and Pace, N. R.: Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing, *J. Bacteriol.*, 173, 4371–4378, 1991.
- Schwalbach, M. S., Hewson, I., and Fuhrman, J. A.: Viral effects on bacterial community composition in marine plankton microcosms, *Aquat. Microb. Ecol.*, 34, 117–127, 2004.
- Schwalbach, M. S., Brown, M., and Fuhrman, J. A.: Impact of light on marine bacterioplankton community structure., *Aquat. Microb. Ecol.*, 39, 235–245, 2005.
- Short, S. M., Jenkins, B. D., and Zehr, J. P.: Spatial and temporal distribution of two diazotrophic bacteria in the Chesapeake Bay, *Appl. Environ. Microb.*, 70, 2186–2192, 2004.
- Short, S. M. and Zehr, J. P.: Nitrogenase gene expression in the Chesapeake Bay estuary, *Environ. Microbiol.*, 9, 1591–1596, 2007.
- Small, J., Call, D. R., Brockman, F. J., Straub, T. M., and Chandler, D. P.: Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays, *Appl. Environ. Microb.*, 67, 4708–4716, 2001.
- Smith, C. J.: Quantitative real-time PCR, in: *Molecular microbial ecology*, edited by: Osborn, A. M. and Smith, C. J., New York, USA, 151–161, 2005.
- Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., et al.: Microbial diversity in the deep sea and the underexplored “Rare biosphere”, *P. Natl. Acad. Sci. USA*, 103, 12115–12120, 2006.
- Stein, J. L., Marsh, T. L., Wu, K. Y., Shizuya, H., and DeLong, E. F.: Characterization of uncultivated prokaryotes: Isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon, *J. Bacteriol.*, 178, 591–599, 1996.
- Stine, O. C., Carnahan, A., Singh, R., Powell, J., Furuno, J. P., et al.: Characterization of microbial communities from coastal waters using microarrays, *Environ. Monit. Assess.*, 81, 327–336, 2003.
- Studholme, D. J. and Dixon, R.: In silico analysis of the sigma(54)-dependent enhancer-binding proteins in *Pirellula* species Strain 1, *FEMS Microbiol. Lett.*, 230, 215–225, 2004.
- Su, Z. C., Mao, F. L., Dam, P., Wu, H. W., Olman, V., et al.: Computational inference and experimental validation of the nitrogen assimilation regulatory network in cyanobacterium *Synechococcus* sp. WH 8102, *Nucleic Acids Res.*, 34, 1050–1065, 2006.
- Suzuki, M. and Giovannoni, S. J.: Bias caused by template annealing in the amplification mixtures of 16S rRNA genes by PCR, *Appl. Environ. Microb.*, 62, 625–630, 1996.
- Suzuki, M., Rappe, M. S., and Giovannoni, S. J.: Kinetic bias in estimates of coastal picoplankton community structure obtained by measurements of small-subunit rRNA gene PCR amplicon length heterogeneity, *Appl. Environ. Microb.*, 64, 4522–4529, 1998.
- Suzuki, M. T., Preston, C. M., Chavez, F. P., and DeLong, E. F.: Quantitative mapping of bacterioplankton populations in seawater: Field tests across an upwelling plume in Monterey Bay, *Aquat. Microb. Ecol.*, 24, 117–127, 2001.
- Suzuki, M. T. and Béjà, O.: An elusive marine photosynthetic bacterium is finally unveiled., *P. Natl. Acad. Sci. USA*, 104, 2561–2562, 2007.
- Swingle, W. D., Sadekar, S., Mastrian, S. D., Matthies, H. J., Hao, J., et al.: The complete genome sequence of *Roseobacter denitrificans* reveals a mixotrophic rather than photosynthetic metabolism, *J. Bacteriol.*, 189, 683–690, 2007.
- Taroncher-Oldenburg, G., Griner, E. M., Francis, C. A., and Ward, B. B.: Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the environment, *Appl. Environ. Microb.*, 69, 1159–1171, 2003.
- Tyagi, S. and Kramer, F. R.: Molecular beacons: Probes that fluoresce upon hybridization, *Nat. Biotechnol.*, 14, 303–308, 1996.
- Tyson, G. W., Chapman, J., Hugenholtz, P., Allen, E. E., Ram, R. J., et al.: Community structure and metabolism through reconstruction of microbial genomes from the environment, *Nature*, 428, 37–43, 2004.
- Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., et al.: Environmental genome shotgun sequencing of the Sargasso Sea, *Science*, 304, 66–74, 2004.
- Vergin, K. L., Urbach, E., Stein, J. L., DeLong, E. F., Lanol, B. D., et al.: Screening of a fosmid library of marine environmental genomic DNA fragments reveals four clones related to members of the order planctomycetales, *Appl. Environ. Microb.*, 64, 3075–3078, 1998.
- Ward, B. B., Capone, D. G., and Zehr, J. P.: What’s new in the nitrogen cycle?, *Oceanography*, 20, 101–109, 2007a.
- Ward, B. B., Eveillard, D., Kirshtein, J. D., Nelson, J. D., Voytek, M. A., et al.: Ammonia-oxidizing bacterial community composition in estuarine and oceanic environments assessed using a functional gene microarray, *Environ. Microbiol.*, 9, 2522–2538, 2007b.
- Wawrik, B., Paul, J. H., and Tabita, F. R.: Real-time pcr quantification of *rbcL* (ribulose-1,5-bisphosphate carboxylase/oxygenase) mRNA in diatoms and pelagophytes, *Appl. Environ. Microb.*, 68, 3771–3779, 2002.
- Wawrik, B., Kutlief, D., Abdivasieva, U. A., Kukor, J. J., Zylstra, G. J., et al.: Biogeography of actinomycete communities and type ii polyketide synthase genes in soils collected in New Jersey and central asia, *Appl. Environ. Microb.*, 73, 2982–2989, doi:10.1128/AEM.02611-06, 2007.

- Wegley, L., Edwards, R., Rodriguez-Brito, B., Liu, H., and Rohwer, F.: Metagenomic analysis of the microbial community associated with the coral *Porites astreoides*, *Environ. Microbiol.*, 9, 2707–2719, 2007.
- Woebken, D., Teeling, H., Wecker, P., Dumitriu, A., Kostadinov, I., et al.: From the Namibian and Oregon coast upwelling systems and their cross-comparison with planctomycete genomes, *ISME J.*, 1, 419–435, 2007.
- Wu, L. Y., Thompson, D. K., Li, G. S., Hurt, R. A., Tiedje, J. M., et al.: Development and evaluation of functional gene arrays for detection of selected genes in the environment, *Appl. Environ. Microb.*, 67, 5780–5790, 2001.
- Yooseph, S., Sutton, G., Rusch, D. B., Halpern, A. L., Williamson, S. J., et al.: The Sorcerer II Global Ocean Sampling expedition: Expanding the universe of protein families, *PLoS Biology*, 5, 432–466, 2007.
- Zehr, J. P. and McReynolds, L. A.: Use of degenerate oligonucleotides for amplification of the *nifH* gene from the marine cyanobacterium *Trichodesmium* spp., *Appl. Environ. Microb.*, 55, 2522–2526, 1989.
- Zehr, J. P. and Hiorns, W. D.: Molecular approaches for studying the activities of marine organisms, in: *Molecular approaches to the study of the ocean*, edited by: Cooksey, K. E., Chapman and Hall, London, UK, 91–112, 1998.
- Zehr, J. P. and Voytek, M. A.: Molecular ecology of aquatic communities: Reflections and future directions, *Hydrobiologia*, 401, 1–8, 1999.
- Zehr, J. P., Waterbury, J. B., Turner, P. J., Montoya, J. P., Omoregie, E., et al.: Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean, *Nature*, 412, 635–638, 2001.
- Zehr, J. P., Montoya, J. P., Jenkins, B. D., Hewson, I., Mondragon, E., et al.: Experiments linking nitrogenase gene expression to nitrogen fixation in the north pacific subtropical gyre, *Limnol. Oceanogr.*, 52, 169–183, 2007.
- Zehr, J. P.: Crystal ball-from microbes to global environmental change: Thinking big about small things, *Environ. Microbiol.*, 1, 24–26, 2009.
- Zhou, J. H.: Microarrays for bacterial detection and microbial community analysis, *Curr. Opin. Microbiol.*, 6, 288–294, 2003.