

## Metagenome-based analysis: A promising direction for plankton ecological studies

YAN QingYun & YU YuHe\*

*Key Laboratory of Biodiversity and Conservation of Aquatic Organisms, Institute of Hydrobiology,  
Chinese Academy of Sciences, Wuhan 430072, China*

Received September 9, 2010; accepted October 14, 2010; published online November 19, 2010

The plankton community plays an especially important role in the functioning of aquatic ecosystems and also in biogeochemical cycles. Since the beginning of marine research expeditions in the 1870s, an enormous number of planktonic organisms have been described and studied. Plankton investigation has become one of the most important areas of aquatic ecological study, as well as a crucial component of aquatic environmental evaluation. Nonetheless, traditional investigations have mainly focused on morphospecies composition, abundances and dynamics, which primarily depend on morphological identification and counting under microscopes. However, for many species/groups, with few readily observable characteristics, morphological identification and counting have historically been a difficult task. Over the past decades, microbiologists have endeavored to apply and extend molecular techniques to address questions in microbial ecology. These culture-independent studies have generated new insights into microbial ecology. One such strategy, metagenome-based analysis, has also proved to be a powerful tool for plankton research. This mini-review presents a brief history of plankton research using morphological and metagenome-based approaches and the potential applications and further directions of metagenomic analyses in plankton ecological studies are discussed. The use of metagenome-based approaches for plankton ecological study in aquatic ecosystems is encouraged.

**metagenome, plankton community, morphospecies, molecular microbial ecology**

**Citation:** Yan Q Y, Yu Y H. Metagenome-based analysis: A promising direction for plankton ecological studies. *Sci China Life Sci*, 2011, 54: 75–81, doi: 10.1007/s11427-010-4103-4

Plankton, the key structural and functional component of aquatic ecosystems, is generally considered to be composed of both producers (phytoplankton) and consumers (zooplankton), but also includes bacterial decomposers (bacterioplankton) [1]. Therefore, it plays important roles in aquatic biogeochemical cycles [2]. Plankton investigation has been a major area of ecological study since the very start of marine research expeditions (e.g., Challenger Expedition, 1872–1876; German Plankton Expedition, 1889). The increasing awareness of their roles in the ecosystem has led to a rapid growth in descriptive, experimental and theo-

retical studies concerning plankton composition, abundance, biomass and activities. However, traditional plankton ecological studies have largely depended upon morphological methods and morphological identification has historically been a difficult task because of the small size of organisms and the paucity of readily observable characteristics of taxonomic value [3]. Therefore, independent investigations may get different results from the same samples; sometimes with significant discrepancy. In addition, different groups of phytoplankton and zooplankton have always been investigated separately and bacterioplankton was often excluded from morphological investigations because of the difficulties in identification. These factors and others have made it

\*Corresponding author (email: yhyu@ihb.ac.cn)

difficult to reach higher-order ecological conclusions about general plankton trends and patterns.

Morphological approaches can only provide a few clues for identification and related functional studies of most microbes. The need for nontraditional techniques to reveal the microbial world was highlighted in 1986 [4]. One year earlier Pace *et al.* [5] had proposed the idea of cloning DNA directly from environmental samples and the first such clone library was reported in 1991 [6]. Since then, natural microorganisms collected from different environments have been investigated at the DNA level and progress has been made in microbial ecology and environmental microbiology [7]. In 1998, a professional term ‘metagenomics’ was coined [8] to describe natural habitat based microbial investigations, which provided powerful tools to extract valuable biological information quickly and effectively from complex environments [7,9]. For plankton research, Giovannoni *et al.* [10] were the first to phylogenetically analyze clone libraries of 16S rRNA genes from a natural bacterioplankton community. To date, large numbers of metagenome-based studies on bacterioplankton and more recently on eukaryotic picoplankton have been undertaken (e.g., [11–14]). Targeting the plankton community as a system, Yan *et al.* [15–18] and Yu *et al.* [19] currently applied fingerprinting techniques to investigate the relationships between DNA polymorphisms and species composition of plankton communities in different habitats and environments (i.e., lakes, rivers, reservoirs, ponds, enclosures, and artificial systems). As community composition is largely dictated by environmental factors, it has been a natural extension of this relationship to use plankton community fingerprints as an indication of environmental conditions [17–21]. In contrast to traditional morphological analyses in ecological study, multiple samples can be analyzed simultaneously with standardized sets of metagenome-based procedures, which make it practical to perform comparative analyses aimed at elucidating general microbial ecological rules. Moreover, by comparing target genes’ expression under different environments, community functions can be explored using function-based metagenomics.

## 1 Morphospecies-based plankton investigation

### 1.1 A brief history of early plankton research

The term ‘plankton’ was coined by the famous plankton ecologist Victor Hensen in 1887, but plankton research can be traced back to when Antonie van Leeuwenhoek observed ‘small infusoria’ with his own microscope. Johannes Müller pioneered the systematic study of small planktonic organisms with hand-held nets in the 1840s [22]. The interests of plankton research initially focused on naming and classifying the species according to their morphological characteristics (i.e., morphospecies). After a period of taxonomic

investigation, numerous planktonic species had been described and more and more new species were reported. During this period, interests in plankton research also extended to the relationships between community structure and the environment. This led directly to the ecological aspects of the field. Furthermore, Victor Hensen applied the quantitative methods to gauge the distribution, abundance and productivity of the microscopic plankton in the open sea [22].

### 1.2 Limitations of traditional plankton investigations

As the basis of biodiversity, species diversity was the first to be investigated and attracted the attention of many taxonomists and ecologists. A large number of species diversity indices have been proposed [23,24]. However, the mechanisms regulating diversity in most systems are not completely understood [25], and questions of how to determine the real diversity of plankton communities and to explore the potential ecological functions present tremendous challenges. These challenges not only include the need for extensive taxonomic identification, but also problems of how to compare and summarize related studies to generate higher-order ecological conclusions. Since the very start of plankton ecological investigation in the 1870s, plankton research has largely depended upon traditional morphological methods. However, morphological classification of planktonic organisms has historically been a difficult task (even for a seasoned taxonomist) because of the lack of distinguishing features, especially for the multitudes of minute, nondescript organisms. Furthermore, planktonic bacteria and some small phytoplankton cannot be classified to the species level or even to the genus level using morphological methods alone. Therefore, it is not so much a lack of ideas but inadequate methodologies and instruments that have limited progress in understanding of the diversity of plankton. Despite more than 100 years for plankton research, our knowledge of plankton diversity and its roles in the natural environment has only increased modestly. Therefore, more powerful techniques are urgently needed to reveal their diversity and explore their ecological functions in ecosystems. Fortunately, recent technical developments in molecular biology have found extensive applications in the field of studying natural community ecology [26–28].

## 2 Metagenome-based plankton study

### 2.1 Metagenome

The term ‘metagenome’ was first introduced to describe the combined genomes of natural microbes in the soil environment [8]. Since then, microbial communities from diverse niches (including oceans, lakes, soils, thermal vents, hot springs, the mouth and gastrointestinal tract) have been involved in metagenome-based studies over a very short pe-

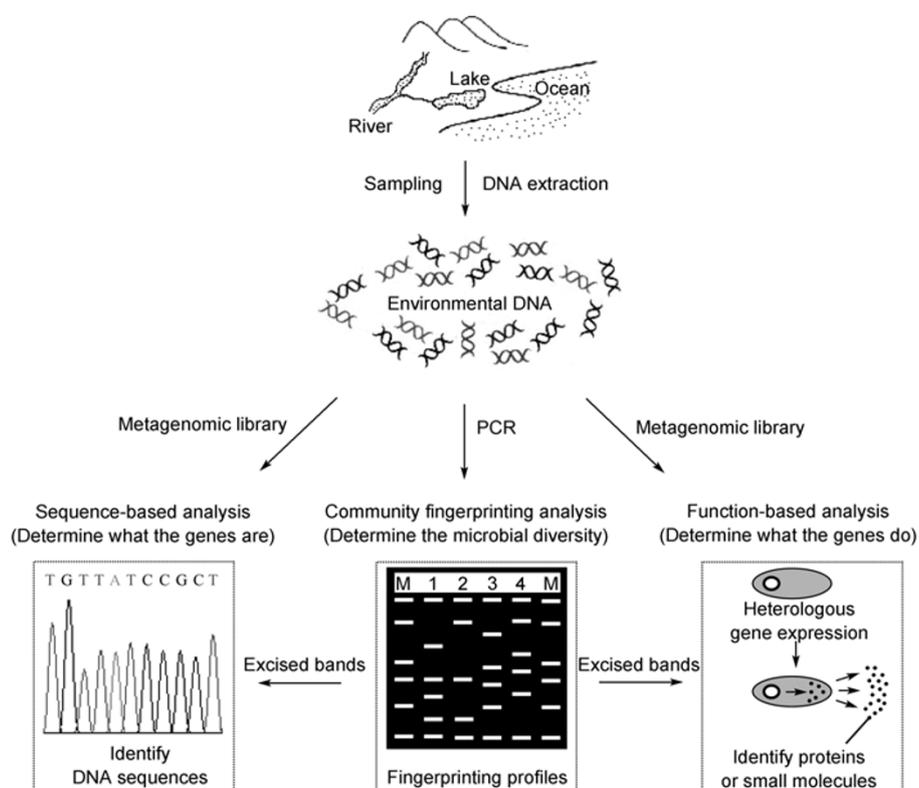
riod of time [9,29]. Now, the term metagenome has been extended to refer to all of the genetic material recovered directly from environmental samples, which can be analyzed in a way analogous to the study of a single genome [7,9]. The genomic study of single organisms had extended our understanding from single genes to the collective genes of an organism, whereas metagenomics goes beyond the genome of single organisms and enables us to study genomes of the community [30]. More importantly, metagenomics transcends the limitations of classical genomics and microbiology and provides valuable strategies to study the relationships among genes, genomes, organisms and communities present in the natural world [7]. Since the term ‘metagenome’ was first published in 1998 [8], the number of annually published SCI publications in this field has increased exponentially and reached a level of approximately 400 in 2010 (according to searches using the online version of the ISI Web of Science).

## 2.2 Metagenomics and its applications in microbial ecology

Metagenomics is the study of genomic material collected directly from natural environmental samples. It is also termed as environmental genomics, ecogenomics and

community genomics. However, metagenomics is most commonly used in describing this emerging field of study. Although the term ‘metagenomics’ was coined in 1998 [8], the concept that organisms could be identified without cultivation by retrieving and sequencing their DNA directly from natural samples dates much earlier (e.g., [4,5]). Another breakthrough was the amplification of 16S rRNA genes from natural picoplankton and the use of amplified environmental genes to perform phylogenetic analysis [10].

All metagenomic studies begin with community DNA extraction from diverse members present in a particular environment. Metagenome information is then analyzed directly or cloned prior to investigations of structure, function and dynamics of natural microbial communities (Figure 1). Over the past two decades, metagenomic studies have provided valuable information in a number of fields including microbiology, medicine, energy, biotechnology, agriculture, environmental remediation and gut microbial ecology. In fact, metagenomics has bridged the gap between genetics and ecology, demonstrating that the genes of single organisms are related to the genes of other species or even to the entire community [7]. It also offers a powerful lens for viewing the microbial world which in turn presents potential to revolutionize understanding of the entire living world.



**Figure 1** Schematic showing the major process of metagenome-based study. All metagenome-based studies involve the extraction of genomic DNA directly from environmental samples. The temporal or spatial heterogeneity of a community can be effectively compared by analyzing the fingerprinting profiles. After the genes are sequenced and compared with identified sequences, the functions of these genes can be determined.

### 3 Metagenome-based studies of plankton communities

#### 3.1 Community fingerprinting analysis

Community fingerprinting, a quick way of viewing the spatiotemporal patterns and succession of target microbial communities, has been widely applied to study communities from a variety of environments (e.g., oceans, lakes, reservoirs, rivers, artificial niches, soils, thermal vents, hot springs, the mouth and gastrointestinal tract) [29]. If comparable samples collected from different locations harbor similar communities in the same area, the fingerprinting patterns are theoretically expected to be similar. To date, the fingerprinting techniques commonly used in aquatic habitats include denaturing/temperature gradient gel electrophoresis (DGGE/TGGE), single strand conformation polymorphism (SSCP), randomly amplified polymorphic DNA (RAPD), terminal-restriction fragment length polymorphism (T-RFLP), amplified ribosomal DNA restriction analysis (ARDRA), ribosomal intergenic spacer analysis (RISA) and automated ribosomal intergenic spacer analysis (ARISA) (see [31] for details). These methods provide a rapid means for screening microbial communities from different environments or for comparing community dynamics at different spatial/temporal scales. Among all of the available fingerprinting methods, PCR-DGGE is one of the most commonly used approaches for screening environmental microbial communities [32], although it cannot resolve all parts of microbial community in complex communities.

Metagenome-based analysis in plankton research has become very common, although studies and publications are primarily focused on the pico- and nanoplankton [12–14]. PCR-DGGE fingerprinting patterns of bacterioplankton communities in lakes with different nutrient content and water color have been found to be strongly correlated with the biomass of microzooplankton, cryptophytes, and chrysophytes [33]. With PCR-DGGE fingerprinting, Lindström [34] also found that the introduction of allochthonous bacteria and their subsequent interaction with other planktonic organisms affected the composition of bacterioplankton. Although eukarya-specific primers have also been developed and applied in aquatic environmental studies (e.g., [11]), comparatively few studies have targeted eukaryotic plankton. Moreover, prokaryotic and eukaryotic groups have always been investigated separately.

Yu *et al.* [35] targeted the entire plankton community as a system (including both planktonic prokaryotes and eukaryotes) and explored the feasibility of applying DNA fingerprinting to community-level plankton studies. Furthermore, Yan *et al.* [15–18,20,21] applied different fingerprinting approaches to investigate relationships among the genetic diversity of plankton communities, species composition and environmental factors in aquatic environments of lakes, rivers, and reservoirs. Results indicated that the DNA

fingerprints of target communities were generally correlated with species composition of plankton and their environments [15–18,20,21]. The concentration of TP was found to be the major factor in determining plankton communities derived from eutrophic environments [19]. PCR-DGGE fingerprinting was further applied to address plankton succession in simulated niches. Plankton succession was not significantly affected by cyanobacterial bloom removal using chitosan-modified local soils [18]. Additionally, specific plankton cyanobacterial groups were explored using PCR-DGGE fingerprinting to monitor succession *in situ* (Taihu Lake, unpublished data). Community fingerprinting analyses have significantly enhanced our ability to understand the non-cultured plankton world [11,12], and have made it possible to elucidate mechanisms of ecosystem functions with molecular insights.

#### 3.2 Sequence-based and function-based analysis

In addition to using fingerprint patterns as indicators of the diversity and succession of target communities, fingerprinting methods have often been used in combination with selectively excised, sequenced bands of interest (e.g., [3,12]). However, the information provided by the excised bands may be insufficient when bands are short (e.g., less than 500 bp for DGGE/TGGE and SSCP). Therefore, genomic DNA recovered from environmental samples also needs to be directly cloned into an appropriate vector and transformed into a bacterial host, and a metagenomic library constructed. The genetic information held within the metagenomic library is then studied using sequence-based or function-based screening (Figure 1) to examine the characteristics and functions of the microbial community. In principle, the clones of a metagenomic library represent the genetic complement derived from the target community in the investigated habitat.

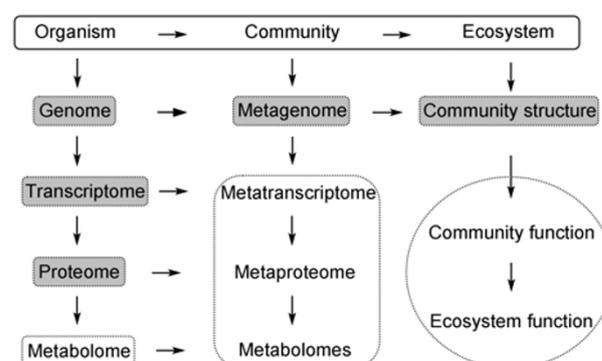
In sequence-based analysis, the clone library is sequenced and analyzed to obtain information regarding the structure and organization of the metagenome; sequence information is then compared with known DNA sequences deposited in public databases (e.g., GenBank, <http://www.ncbi.nlm.nih.gov>; EMBL, <http://www.ebi.ac.uk/embl>; DDBJ, <http://www.ddbj.nig.ac.jp>). This can involve the complete sequencing of clones containing phylogenetic anchors indicating taxonomic groups. Alternatively, random sequencing can also be conducted to identify the gene of interest [7]. By using this strategy, more and more new groups of bacteria/archaea have been identified [10,36]. Moreover, Stein *et al.* [37] isolated a large genome fragment from uncultured planktonic archaeon. After 1997, the culture-independent 16S rRNA gene sequences derived from environmental clones in GenBank began to exceed those from cultivated bacteria and archaea [38]. Since Giovannoni *et al.* [10] first phylogenetically analyzed 16S rRNA genes amplified from natural bacterioplankton, phylogenetic information directly

derived from environmental samples has greatly enhanced our understanding of biodiversity in the biosphere. More recently, phylogenetic analysis of the amplified 18S rDNA, internal transcribed spacer (ITS) and other regions mimicking the 16S rDNA analysis in prokaryotes, have indicated high diversity of small planktonic eukaryotes [13,14]. These monumental studies have provided new insights for characterizing planktonic organisms from natural assemblages. Moreover, accompanying the revolution driven by the platforms of Roche 454, Illumina Genome Analyzer and ABI-SOLiD, the high-throughput next-generation sequencing will realize direct and cost-effective sequencing of complex samples at an unprecedented scale and speed [39,40]. In addition, open-source metagenomics RAST (MG-RAST) is also available and creating expanded opportunities for the annotation and analysis of metagenomic sequences [41]. As sequencing technology is further refined, sequence-based analyses will continue to enrich our understanding of the diversity and functions of non-cultivated plankton in the natural world.

In function-based analysis, the metagenomic library is screened to identify functions of interest, such as vitamin production and antibiotic resistance [42]. Additionally, function-based methods also enable us to directly extract and identify novel proteins and metabolites from microbial communities [42]. Thus, the capabilities of a community can be assessed without having prior knowledge of the gene sequences and researchers can identify entirely new classes of genes for target functions. To date, functional analysis has identified antibiotics, antibiotic resistance genes, lipases, chitinases, membrane proteins, degradative enzymes and genes encoding the biotin synthesis pathways (see [43] for details). However, current techniques cannot express multi-genes in any particular host bacterium and the frequency of metagenomic clones that express a given activity is generally very low because the host may not have the correct microenvironment for target expression. With development of microarray-based genomic technology, He *et al.* [44,45] developed the GeoChip for studying biogeochemical processes and functional activities of microbial communities. The current GeoChip contains probes for different genes involved in biogeochemical cycling, making it possible to use functional gene arrays to analyze complex environmental samples. The functional anchors, which are the functional analogs of phylogenetic anchors, are also emerging in metagenomic analysis [7]. However, most function-based studies have focused on soil environments and planktonic organisms and aquatic ecosystems have rarely been involved. Bacterio-, phyto- and zooplankton are essential components of the aquatic food web and generally contribute the dominant portion of productivity in aquatic ecosystems [46]. Therefore, further efforts are urgently needed to address functional expression for planktonic communities, to advance function-based plankton metagenomics.

## 4 Perspectives

Metagenomics is a burgeoning and exciting field that has generated enormous amounts of valuable biological information over a relatively short period. As a result of recent developments, future directions for plankton research at the metagenome level can be illustrated as in Figure 2. Although current efforts in microbial genomics have mainly focused on single organisms, from genome to transcriptome and also to proteome (as indicated by the shadows at organism level, Figure 2), some breakthrough has also been realized recently at the community level [14,47]. Metagenome-based techniques are now widely available for the study fields of biodiversity, agriculture, biotechnology, human health and generally for deeper understanding of the biosphere. Current trends suggest that opportunities for further development are broad (Figure 2). In brief, expansion to the community level will lead to comprehensive understanding of community functions; expansion from single organisms to ecosystems will advance understanding of biotic and abiotic interactions and ecosystem evolution. The continuing development of current techniques and the introduction of more refined techniques will increase the potency of our understanding of plankton in the natural world (e.g., structure, abundance, biomass and functions in ecosystems). This will help us to address a variety of ecological questions. For example, high-throughput sequencing of metagenomic DNA derived from plankton samples will greatly advance our understanding of the plankton community genetic diversity, metabolic functions, their ecology and evolution [39,40]. Another representative example is the newly developed GeoChip, which has proved to be a powerful, high-throughput metagenomic tool for analyzing microbial communities (including diversity, metabolic capability, and their function activity) [45]. This type of functional gene array can also be used to study plankton community functional structure and to link plankton communities to ecosystem processes and functioning. Nonetheless, it should be



**Figure 2** Schematic showing current status and future directions for the development of genomic-related plankton research. The shadows indicate the efforts accomplished in current studies; the others await future developments.

acknowledged that there are some challenges ahead. First, to define and elucidate genes' functions: Although sequence homology comparisons are useful, they only provide clues because homology does not necessarily mean the same function. A further challenge for metagenomic studies is to understand the consequences of genetic capacity and interactions at the community/system-level, beyond the individual/population-level.

*This work was supported by the National Natural Science Foundation of China (Grant No. 30770298). We thank Prof. James A. Nienow (Valdosta State University) and Prof. Brendan J. Hicks (University of Waikato) for their help with the English.*

- 1 Callieri C, Stockner J G. Freshwater autotrophic picoplankton: A review. *J Limnol*, 2002, 61: 1–14
- 2 Roemmich D, McGowan J. Climatic warming and the decline of zooplankton in the California current. *Science*, 1995, 267: 1324–1326
- 3 Savin M C, Martin J L, LeGresley M, et al. Plankton diversity in the Bay of Fundy as measured by morphological and molecular methods. *Microb Ecol*, 2004, 48: 51–65
- 4 Pace N R, Stahl D A, Lane D J, et al. The analysis of natural microbial populations by rRNA sequences. *Adv Microb Ecol*, 1986, 9: 1–55
- 5 Pace N R, Stahl D A, Lane D J, et al. Analyzing natural microbial populations by rRNA sequences. *ASM News*, 1985, 51: 4–12
- 6 Schmidt T M, DeLong E F, Pace N R. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J Bacteriol*, 1991, 173: 4371–4378
- 7 Handelsman J. Metagenomics: Application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev*, 2004, 68: 669–685
- 8 Handelsman J, Rondon M R, Brady S F, et al. Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. *Chem Biol*, 1998, 5: R245–249
- 9 Xu J. Microbial ecology in the age of genomics and metagenomics: Concepts, tools, and recent advances. *Mol Ecol*, 2006, 15: 1713–1731
- 10 Giovannoni S J, Britschgi T B, Moyer C L, et al. Genetic diversity in Sargasso Sea bacterioplankton. *Nature*, 1990, 345: 60–63
- 11 van Hannen E J, van Agterveld M P, Gons H J, et al. Revealing genetic diversity of eukaryotic microorganisms in aquatic environments by denaturing gradient gel electrophoresis. *J Phycol*, 1998, 34: 206–213
- 12 Díez B, Pedrós-Alió C, Massana R. Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Appl Environ Microbiol*, 2001, 67: 2932–2941
- 13 López-García P, Rodríguez-Valera F, Pedrós-Alió C, et al. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature*, 2001, 409: 603–607
- 14 Moon-van der Staay S Y, De Wachter R, Vaulot D. Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature*, 2001, 409: 607–610
- 15 Yan Q Y, Yu Y H, Feng W S. Corresponding relations between DNA fingerprinting topological structure and species composition of plankton community in Dongting Lake (in Chinese). *Acta Ecol Sin*, 2005, 25: 2221–2226
- 16 Yan Q Y, Yu Y H, Feng W S, et al. DNA polymorphism of the plankton community and its relationships to species composition in the Three Gorges Reservoir Region of the Yangtze River. *J Freshw Ecol*, 2006, 21: 385–390
- 17 Yan Q Y, Yu Y H, Feng W S, et al. Genetic diversity of plankton community as depicted by PCR-DGGE fingerprinting and its relation to morphological composition and environmental factors in Lake Donghu. *Microb Ecol*, 2007, 54: 290–297
- 18 Yan Q Y, Yu Y H, Feng W S, et al. Plankton community succession in artificial systems subjected to cyanobacterial blooms removal using chitosan-modified soils. *Microb Ecol*, 2009, 58: 47–55
- 19 Yu Y H, Yan Q Y, Feng W S. Spatiotemporal heterogeneity of plankton communities in Lake Donghu, China, as revealed by PCR-denaturing gradient gel electrophoresis and its relation to biotic and abiotic factors. *FEMS Microbiol Ecol*, 2008, 63: 328–337
- 20 Yan Q Y, Yu Y H, Zhang W J. The DNA polymorphism of plankton community and eutrophication in Lake Donghu, Wuhan (in Chinese). *Acta Ecol Sin*, 2005, 25: 461–465
- 21 Yan Q Y, Yu Y H, Feng W S, et al. Relationship between DNA fingerprinting of plankton community and physico-chemical factors (in Chinese). *Acta Hydrobiol Sin*, 2005, 29: 601–606
- 22 Smetacek V, Montresor M, Verity P. Marine productivity: Footprints of the past and steps into the future. In: Williams P J le B, Thomas D N, Reynolds C S, eds. *Phytoplankton Productivity*. Oxford: Blackwell, 2002. 350–369
- 23 Lewis C E, Swindel B E, Tanner G W. Species diversity and diversity profiles: Concept, measurement and application to timber and range management. *J Range Manageme*, 1988, 41: 466–469
- 24 Magurran A E. *Ecological Diversity and its Measurement*. Princeton: Princeton University Press, 1988
- 25 Interlandi S J, Kilham S S. Limiting resources and the regulation of diversity in phytoplankton communities. *Ecology*, 2001, 82: 1270–1282
- 26 Lee S H, Fuhrman J A. Spatial and temporal variation of natural bacterioplankton assemblages studied by total genomic DNA cross-hybridization. *Limnol Oceanogr*, 1991, 36: 1277–1287
- 27 Santo-Domingo J W, Kaufman M G, Klug M J, et al. Influence of diet on the structure and function of the bacterial hindgut community of crickets. *Mol Ecol*, 1998, 7: 761–767
- 28 Zehr J P, Voytek M A. Molecular ecology of aquatic communities: Reflections and future directions. *Hydrobiologia*, 1999, 401: 1–8
- 29 Riesenfeld C S, Schloss P D, Handelsman J. Metagenomics: Genomic analysis of microbial communities. *Annu Rev Genet*, 2004, 38: 525–552
- 30 Sharma P, Kumari H, Kumar M, et al. From bacterial genomics to metagenomics: Concept, tools and recent advances. *Indian J Microbiol*, 2008, 48: 173–194
- 31 Dorigo U, Volatier L, Humbert J F. Molecular approaches to the assessment of biodiversity in aquatic microbial communities. *Water Res*, 2005, 39: 2207–2218
- 32 Muyzer G. DGGE/TGGE a method for identifying genes from natural ecosystems. *Curr Opin Microbiol*, 1999, 2: 317–322
- 33 Lindström E S. Bacterioplankton community composition in five lakes differing in trophic status and humic content. *Microb Ecol*, 2000, 40: 104–113
- 34 Lindström E S. Investigating influential factors on bacterioplankton community composition: Results from a field study of five mesotrophic lakes. *Microb Ecol*, 2001, 42: 598–605
- 35 Yu Y H, Zhang W J, Yan Q Y. The feasibility for application of DNA fingerprinting to community-level life system (in Chinese). *Acta Hydrobiol Sin*, 2004, 28: 457–463
- 36 Fuhrman J A, McCallum K, Davis A A. Novel major archaeobacterial group from marine plankton. *Nature*, 1992, 356: 148–149
- 37 Stein J L, Marsh T L, Wu K Y, et al. Characterization of uncultivated prokaryotes: Isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. *J Bacteriol*, 1996, 178: 591–599
- 38 Rappé M S, Giovannoni S J. The uncultured microbial majority. *Annu Rev Microbiol*, 2003, 57: 369–394
- 39 Tringe S G, Rubin E M. Metagenomics: DNA sequencing of environmental samples. *Nat Rev Genet*, 2005, 6: 805–814
- 40 Marguerat S, Wilhelm B T, Bähler J. Next-generation sequencing: Applications beyond genomes. *Biochem Soc Trans*, 2008, 36: 1091–1096

- 41 Meyer F, Paarmann D, D'Souza M, *et al.* The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, 2008, 9: 386
- 42 Committee on Metagenomics: Challenges and functional applications, National Research Council (USA). *The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet*. Washington DC: National Academies Press, 2007
- 43 Schloss P D, Handelsman J. Biotechnological prospects from metagenomics. *Curr Opin Biotechnol*, 2003, 14: 303–310
- 44 He Z L, Gentry T J, Schadt C W, *et al.* GeoChip: A comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J*, 2007, 1: 67–77
- 45 He Z L, Deng Y, van Nostrand J D, *et al.* GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. *ISME J*, 2010, 4: 1167–1179
- 46 Barber R T, Hiking A K. History of the study of plankton productivity. In: Williams P J le B, Thomas D N, Reynolds C S, eds. *Phytoplankton Productivity*. Oxford: Blackwell, 2002. 16–43
- 47 Giovannoni S J, Stingl U. Molecular diversity and ecology of microbial plankton. *Nature*, 2005, 437: 343–348

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.