Role of Genome based Tools in Environmental Microbiology

Srilatha B*
Presidency College, Bangalore University, India

Abstract
Genomic investigations into the diversity of environmental microbes are leading to insights into ecological dynamics, the evolution of new forms of biological systems, and the discovery of new functions that might be exploited for biotechnological and biomedical purposes. It is now clear that an understanding of the community structure, function and evolution of bacteria in their natural environments is required to meet the promise of microbial biotechnology. To meet these new challenges, microbiologists are applying the tools of genomics and related high-throughput technologies to both cultured microbes and environmental samples. This work will lead to new views on ecosystems and biological function together with the biotechnology enabled by this science.

Keywords: Environment; Microbes; Genomics; Biotechnology; Biomedical.

Abbreviations: (EEM): Environmental effects monitoring; (EXPASY): Expert Protein Analysis System; (2DE): Two-dimensional gel electrophoresis; (PCR): polymerase chain reaction;

Introduction
Microorganisms are found almost in every habitat present in nature due to their ubiquitous property. Some microorganisms are free living where as others are parasites, they exhibit various beneficial properties in food production, in cleaning up of environment, human health [1]. The biotechnology potential is increasing exponentially with the identification of organisms, isolation of novel compounds and their pathways, and the molecular and biochemical characterization of cellular components [2].

Microorganisms of various types are present in the environment [3]. Environmental microorganisms, especially those living under extreme conditions, cannot be cultured easily under laboratory conditions. Genomes of uncultured organisms have remained mostly uncharacterized and are thought to contain a wide range of novel genes of scientific and industrial interest. Metagenomics approaches, which are analyses of mixed populations of uncultured microbes, have been developed to identify novel and industrially useful genes and to study microbial diversity in a wide variety of environments [4]. As the role of the environment is accorded a more prominent role in modifying the relationship between genetic variants and clinical measures of disease, consideration of gene-environment interactions is a must [5].

Microbial ecology examines the diversity and activity of microorganisms in Earth’s biosphere. In the last 20 years, the application of genomics tools have revolutionized microbial ecological studies and drastically expanded our view on the previously underappreciated microbial world. This review first introduces the basic concepts in microbial ecology and the main genomics methods that have been used to examine natural microbial populations and communities. In the ensuing three specific sections, the applications of the genomics in microbial ecological research are highlighted. The four specific genomics methods (phylogenetic analysis of ribosomal RNA, DNA–DNA re-association kinetics, Metagenomics, and micro-arrays) in analyzing the diversity and potential activity of microbial populations and communities from a variety of terrestrial and aquatic environments [6].

Genomics technologies for environmental science
Molecular techniques are powerful tools for monitoring environmental effects and characterizing microbial diversity. The Environmental Microbiology group is undertaking a series of projects using such molecular techniques. The research team has undertaken a project to apply gene arrays for environmental effects monitoring (EEM), which is vital to ecosystem protection. Molecular tools such as rRNA probes, DNA extraction and analyses, i.e., denaturing gradient gel electrophoresis and microarrays, can effectively monitor changes in, and improve the understanding of microbial communities involved in vital ecosystem processes.

Gene inventory and Metagenomics techniques have allowed rapid exploration of bacterial diversity and the potential physiologies present within microbial communities. However, it remains nontrivial to discover the identities of environmental bacteria carrying two or more genes of interest [7]. However, the complexities and structural characteristics of these genes are still unknown among species; the information analysis in entropy view may thus help us elucidate mechanisms [8]. Comparison with existing approaches shows that this method can achieve better performance in terms of environment [9]. As a result of these properties, in recent years projection methods are being successfully applied to biological data such as DNA microarrays and proteomic data [10].

Molecular biology, in the genome era, does not refer to studies involving just single macromolecules, it actually involves the study of complete cellular pathways, and why not, even entire organisms. Indeed, the world-wide genome-sequencing projects revolutionized the field and are producing unimaginable amount of biological data, providing a near complete list of the components that are present in an organism [11]. There are many ways of recovering biological products...
and the decision the scientist or engineer has is to decide what method is best to achieve the most efficient separation process to meet the growing demands of the biotechnology industry [12].

**DNA sequencing**

High-throughput DNA sequencing has enabled systems biology to begin to address areas in health, agricultural and basic biological research. Concomitant with the opportunities is an absolute necessity to manage significant volumes of high-dimensional and inter-related data and analysis. The output from these technologies currently ranges from 1-20 gigabases of raw sequence information per experiment, with a relatively high error rate compared to Sanger sequencing. The sheer quantity of output, the relative shortness of reads and the frequency of errors have created problematic areas for data management in terms of organization, analysis and information extraction [13].

Sequencing using both reverse and forward primers was carried out and the sequence so obtained was translated using EXPASY (Expert Protein Analysis System) tool. Nucleotide and protein sequence data was analyzed using BLAST program at NCBI website whereas Clustal W was performed for multiple sequence alignment of the test sequence [14]. These new technologies are rapidly evolving, and near-term challenges include the development of robust protocols for generating sequencing libraries, building effective new approaches to data-analysis, and often a rethinking of experimental design.DNA sequencing has the potential to dramatically accelerate biological and biomedical research in environmental microbes by enabling the comprehensive analysis of genomes, transcriptomes and interactomes to become inexpensive, routine and widespread, rather than requiring significant production-scale efforts [15].

Methods that are based on DNA sequencing circumvent these obstacles, as DNA can be isolated directly from living or dead cells in various contexts. Such methods have led to the emergence of a new field, which is referred to as Metagenomics [16].

**Metagenomics**

Metagenomics is a rapidly growing field of research that has had a dramatic effect on the way we view and study the microbial world. Environmental microbiology that using conventional methods, cultured microorganisms represents no more than 1% of the microorganisms present in the vast majority of environmental habitats. Consequently, a huge metabolic diversity still remains to be explored and discovered. With the aim to study and use the information contained in the genomes of uncultured microbes, environmental microbiologists have been investigating microbial communities since the 1990’s applying Metagenomics based approaches. Sequence annotation by gene function revealed specific adaptive capabilities enriched in the air environment, including genes potentially involved in resistance to desiccation and oxidative damage [17].

Metagenomics is the genomic analysis of microorganisms by direct extraction and cloning of DNA from an assemblage of microorganisms. The development of Metagenomics stemmed from the ineluctable evidence that as-yet-uncultured microorganisms represent the vast majority of organisms in most environments on earth. Novel genes and gene products discovered through Metagenomics. The application of Metagenomics sequence information will facilitate the design of better culturing strategies to link genomic analysis with pure culture studies [18]. With improved genotyping technologies and the growing number of available markers, case-control Genome Wide Association Studies (GWAS) have become a key tool for investigating complex diseases [19].

DNA typing techniques can cause problems when evidence samples are inadvertently contaminated with DNA from another source. Therefore, precautions need to be taken to minimize the risk of contamination [20]. Such population-specific disease-gene and genetic damage association studies can provide disease-damage susceptibility/resistance information which can be useful for exploring target specific DNA-safe therapeutics [21]. Proteins are far more complex than the genome and a proper analysis can be extremely expensive and time consuming [22].

Metagenomics-based approaches have led to the accumulation of an increasing number of DNA sequences, but until this time the sequences retrieved have been those of uncultured microbes. These genomic sequences are currently exploited for novel biotechnological and pharmaceutical applications and to increase our knowledge on microbial ecology and physiology of these microbes. Using the Metagenomics sequences to fully understand how complex microbial communities function and how microbes interact within these niches represents a major challenge for microbiologists today [23].

**DNA microarrays**

Although DNA microarray technology has been used successfully to analyze global gene expression in pure cultures, it has not been rigorously tested and evaluated within the context of complex environmental samples. Adapting microarray hybridization for use in environmental studies faces several challenges associated with specificity, sensitivity and quantization [24]. DNA microarray technology permits high-throughput identification of differentially expressed genes [25].

DNA microarrays have emerged as one of the most promising methods for the analysis of gene expression. This technique allows the study of an immense amount of genes (over 10,000) with only one experiment and therefore can draw a picture of a whole genome. Anyway, the huge number of data coming out from microarray experiments may often raise experimental complications and difficulties in the analysis [26]. Microarrays are a novel platform for analysis of genes and genomes in microbes [27].

DNA microarrays can proportionate an instant picture about the preferential gene expression between two different environmental samples. However, this type of analysis is very difficult and complex in natural ecosystems, mainly because of the broad biodiversity and multiple environmental parameters that may affect gene expression. As a result of these properties, in recent years projection methods are being successfully applied to biological data such as DNA microarrays and proteomic data [28].

Nowadays there is an increasing interest in the development of more efficient and less time-consuming methods to assess the presence of microorganisms, as well as their viability for bioprocess control and improvement. Rapid detection of microorganisms in samples is one of the key questions to obtain real-time data for the development of more accurate quality control programs [29]. Microarrays can be coupled with PCR where they serve as a set of parallel dot-blot tests to enhance
product detection and identification. Finally, microarrays can also be used to “fingerprint” bacterial isolates and they can be used to identify diagnostic markers suitable for developing new PCR-based detection assays [30].

DNA microarrays exploit primary sequence data to measure transcript levels and detect sequence polymorphisms, for every gene, simultaneously. The design and construction of a DNA microarray for any given microbial genome are straightforward. By monitoring microbial gene expression, one can predict the functions of uncharacterized genes, probe the physiologic adaptations made under various environmental conditions, identify virulence-associated genes, and test the effects of drugs [31].

Microarrays and Metagenomics to investigate the genetic diversity of environmentally relevant micro-organisms and identify new functional genes involved in the catabolism of xenobiotics [32]. Finally, we outline scenarios for an innovative combination of microarrays with other molecular tools for structure-function analysis of complex microbial communities [33].

Bioinformatics

Bioinformatics based analysis and prediction is playing a pivotal role in understanding and capturing the in-depth knowledge of biological molecules particularly with reference to proteomics and genomics. Although with this advancement, there have been only limited efforts on the collection of all relevant information for a specific field of interest. With this realization, present study focuses on the wide spread data and information related to the occurrence and potential of degrading bacteria. The information and detailed account on these bacteria are quite limited and scattered in scientific journals [34]. Knowledge of the three-dimensional structure of a protein would be an invaluable aid to understand the details of a particular protein [35]. Further studies are needed to elucidate the precise contributions of each of these proteins and to determine their possible relevance in the targeting of new therapeutic interventions [36].

Metabolic networks are complex and highly interconnected, thus systems-level computational approaches are required to elucidate and to understand metabolic genotype-phenotype relationships [37]. The implications of this hypothesis in genetic diversity, protein antigenic properties and diseases are discussed [38]. Computational biology technology has facilitated an increase in the successful rate of genetic association study and reduced the cost of genotyping. In the present study, we applied various bioinformatics tools for the selection of high potential microbes [39]. The differentially expressed genes identified in this study are also considered as biomarkers [40]. These tools are being used toward the development of novel therapies, for the utilization of systems models, and to help guide experimental investigations [41].

Massively parallel pyrosequencing of the small subunit (16S) ribosomal RNA gene has revealed that the extent of rare microbial populations in several environments [42]. The evaluation of the various molecular properties of these populations helps in the discovery of new microbes [43].

We describe a novel approach that eliminates costly and time-consuming probe selection and testing by applying data mining and common bioinformatics tools. Similar to a rational drug design process in which drug-protein interactions are modeled in the computer, the rational probe design described here uses a set of criteria and publicly available bioinformatics software to select the desired probe molecules from libraries comprised of hundreds of thousands of probe molecules [44]. Summary data from recent studies provide overwhelming evidence that bioinformatics tools are useful in protein interactions [45]. Microorganisms provide a large pool of bioactive compounds, and the intensive search for new drugs leads to the identification and structure determination of many novel compounds from these organisms [46].

Proteomics

Protein function can be understood in terms of its structure. Indeed, the three-dimensional structure of a protein is closely related to its biological function. Proteins that perform similar functions tend to show a significant degree of structural homology [47]. Protein structure has always been a significant concern among molecular biologists because it provides intimate information regarding the function and mechanism of the given protein. This knowledge regarding proteins, which are key molecules in the biology of living organisms, can be used in a variety of ways, ranging from protein structure modeling to structural genomics [48]. This technology was choosing because of high performance in application development. This application was tested in order to analyze their performance for accessing heterogeneous biological data [49].

Existing protein-protein interactions databases cover only a portion of the interactomes and interaction information on protein isoforms is underrepresented [50]. Existing algorithms that are based on sequence homology (ortholog conservation) or protein structural data are not necessarily superior [51]. Proteomic approach to identify proteins [52]. Two-dimensional gel electrophoresis (2DE) still plays a key role in proteomics for exploring the protein content of complex biological mixtures. However, the development of fully automatic strategies in extracting interpretable information from gel images is still a challenging task [53].

Proteomics is to advance knowledge in the field of environmental biotechnology, including studies of bacterial physiology, metabolism and ecology. Bacteria are widely applied in environmental biotechnology for their ability to catalyze dehalogenation, methanogenesis, denitrification and sulfate reduction, among others. Environmental samples are often highly complex, which makes proteome studies in this field especially challenging. Some of these challenges are the lack of genome sequences for the vast majority of environmental bacteria, difficulties in isolating bacteria and proteins from certain environments, and the presence of complex microbial communities. Despite these challenges, proteomics offers a unique dynamic view into cellular function [54]. Improving these aspects of cell-based proteomics is essential for improving the stringency and efficacy of [55]. Environmental proteomics enables simple protein cataloging, comparative and semi-quantitative proteomics, analyses of protein localization, discovery of post-translational modifications, and even determination of amino-acid sequences and genotypes by strain-resolved Proteogenomics [56].

Functional genomics approaches, such as proteomics, greatly enhance the value of genome sequences by providing a global level assessment of which genes are expressed, when genes are expressed and at what cellular levels gene products are synthesized. With over 1000 complete genome sequences of different microorganisms available,
and DNA sequencing for environmental samples (Meta-omics) producing vast amounts of gene sequence data, there is a real opportunity and a clear need to generate associated functional genomic data to learn about the source microorganisms [57].

PCR

The polymerase chain reaction (PCR) is an enzymatic reaction which follows simple, predictable and well understood principles. Selective amplification of nucleic acid molecules, that are initially present in minute quantities, provides a powerful tool for analyzing nucleic acids [58]. PCR offers certain advantages over conventional methods for the diagnosis and characterization of microbes. When approximately applied, PCR can be more specific, sensitive, versatile, and rapid than conventional methods; in addition, genetic information can be obtained in the process [59].

PCR technology provides potential for a powerful diagnostic tool in detection of pathogenic microorganisms [60]. Thus, real-time PCR assay can be used as a rapid and effective procedure that can detect minute amounts of microbes from complex environments [61]. Remarkably, PCR-based markers linked to malt trait could have been used for evaluating the genetic diversity and determining the genetic relationships among these accessions [62].

Microbial populations in complex environmental samples are difficult to characterize; current techniques are incomplete and time-consuming. We investigated a polymerase chain reaction (PCR)-based method for rapidly comparing bacterial communities independent of culture or cloning. This recent increase in awareness of our inability to cope with microbial diversity is due to a quantum leap in methodologies (e.g. molecular cloning, polymerase chain reaction (PCR), DNA probing etc.) and in the development of concepts that allowed biologists to come to a unified view of the genealogy of all living material, i.e. the use of semantic molecules for phylogenetic studies [63]. Microbial degradation and decolorization is an environment friendly and cost-competitive alternative to chemical decomposition processes [64].

Identifications of microbial organisms are now usually done by comparing their SSU rRNA gene sequences to those of known organisms. The usual application is to study the composition of the microbial community within a given environmental or clinical sample. SSU rRNA gene sequences are thus obtained, either after cloning the PCR products and random sequencing a set of clones) or by pyrosequencing). The questions are to find out if these sequences are related to other sequences already found in environmental samples, and/or related to well known cultured microorganisms and eventually a type strain [65].

In conclusion, the results of the present investigation clearly indicate that [66] the direct sequencing results, therefore, shall help in understanding the molecular events associated with environmental microbes [67]. In addition, gene expression profiling may provide mechanistic insights that may subsequently be employed to develop biomarkers to detect chemical toxicity as well as strategies to intervene chemical toxicity [68].

There is an increasing demand worldwide for the application of intelligent, fast and inexpensive measurement systems in clinical diagnosis. In the field of Clinical Microbiology, current techniques generally require 24–48 hours to identify and characterize a pathogenic microorganism following a series of biochemical tests. Although new molecular biological and serological test have been introduced recently, they still have not replaced cultural methods and microscopy [69].

Revolutionary advancements in molecular tools to understand the structure and function of microbial communities are bolstering the power of microbial ecology. A push from advances in modern materials along with a pull from a societal need to become more sustainable is enabling environmental biotechnology to create novel processes [70].

Conclusion

Molecular biology has revolutionized the study of microorganisms in the environment and improved our understanding of the composition, phylogeny, and physiology of microbial communities. The current molecular toolbox encompasses a range of DNA-based technologies and new methods for the study of RNA and proteins extracted from environmental samples. Currently there is a major emphasis on the application of "Omics" approaches to determine the identities and functions of microbes inhabiting different environments. The key to this approach will be the integration of gene expression, proteomics, physiological, mutant phenotype, and metabolic data into working cellular models that can accurately predict the response of the organism to a given environment. High-throughput sequencing and advances in DNA sequencing and amplification technology, coupled with genomic tools, are enabling holistic views into the composition and dynamics of predominantly microbial communities.

References

1. Naga Deepthi C, Phani Santosh, Prithivi raj kodi (2011) Evaluation of the Various uses of Microorganisms with Emphasis on Probiotics. J Microbial Biochem Technol R1: 004.
2. Chellapandi P, Sivaramakrishnan S, Viswanathan MB (2010) Systems Biotechnology: an Emerging Trend in Metabolic Engineering of Industrial Microorganisms. J Comput Sci Syst Biol 3: 043-049.
3. Augusto da Costa AC, da Silva Lino LA, Hannesch O (2011) Total Microbial Populations in Air-Conditioned Spaces of a Scientific Museum: Precautions Related to Biodeterioration of Scientific Collections. J Bioprocess Biotechniq 1:106.
4. Abe T, Sugawara H, Kinouchi M, Kanaya S, Ikemura T (2005) Novel phylogenetic studies of genomic sequence fragments derived from uncultured microbial mixtures in environmental and clinical samples. DNA Res 12: 281-290.
5. Lee YC, Lai CQ, Ordovas JM, Parnell LD (2011) A Database of Gene-Environment Interactions Pertaining to Blood Lipid Traits, Cardiovascular Disease and Type 2 Diabetes. J Data Mining in Genom Proteomics 2: 106.
6. Xu J (2006) Microbial ecology in the age of genomics and Metagenomics: concepts, tools, and recent advances. Mol Ecol 15: 1713-1731.
7. Oltesen EA, Hong JW, Quake SR, Leadbetter JR (2006) Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. Science 4: 1464-1467.
8. Xie X, Yu Y, Liu G, Yuan Z, Song J (2010) Complexity and Entropy Analysis of DNA Methylationtransferase. J Data Mining in Genom Proteomics 1:105.
9. Liu R, Hu J (2011) Prediction of Discontinuous B-Cell Epitopes Using Logistic Regression and Structural Information. J Proteomics Bioinform 4: 010-015.
10. Mazzara S, Cerutti S, Iannaccone S, Conti A, Olivieri S, et al. (2011) Application of Multivariate Data Analysis for the Classification of Two Dimensional Gel Images in Neuroproteomics. J Proteomics Bioinform 4: 016-021.
11. Piccoli S, Giorgetti A (2011) Perspectives on Computational Structural Bio-Systems. J Bioprocess Biotechniq 1:104e.
12. Lee T, Amore TD (2011) Membrane Separation Theoretical and Applicable
55. Balashova EE, Lokhov PG (2010) Proteolytically-cleaved Fragments of Cell-surface Proteins from Live Tumor Cells Stimulate Anti-tumor Immune Response In vitro. J Carcinogene Mutagene 1:103.

56. Schneider T, Riedel K (2010) Environmental proteomics: analysis of structure and function of microbial communities. Proteomics 10: 785-798.

57. Burg D, Ng C, Ting L, Cavicchioli R (2011) Proteomics of extremophiles. Environ Microbiol 13: 1934-1955.

58. Garg N, Pundhir S, Prakash A, Kumar A (2008) PCR Primer Design: DREB Genes. J Comput Sci Syst Biol 1: 21-40.

59. Eroglu F, Koltas IS, Genc A (2011) Identification of Causative Species in Cutaneous Leishmaniasis Patients Using PCR-RFLP. J Bacteriol Parasitol 2: 113.

60. Banerjee HN, Gramby M, Hawkins Z (2011) Molecular Diagnosis of Helicobacter Pylori Strain by 16S rDNA PCR Amplification and Direct Sequencing. J Bioprocess Biotechniq 1: 105e.

61. Mesapogu S, Babu BK, Bakshi A, Reddy SS, Saxena S (2011) Rapid detection and quantification of Fusarium udum in soil and plant samples using realtime PCR. J Plant Pathol Microbiol 2: 107.

62. Qian G, Ping J, Wang D, Zhang Z, Luo S, et al. (2011) Malt Genotypic Screening of Polymorphism Information Content (PIC) of PCR-based Marker in Barley, Based on Physiological Traits. Molecular Biology 1: 101.

55. Von Wintzingerode F Göbel UB, Stackebrandt E (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS Microbiol Rev 21: 213-229.

64. Ramya M, Jayapani S, Manju A, Jiffe JS (2010) Biodegradation and Decolorization of Acid Red by Acinetobacter radioresistens. J Bioremed Biodegrad 1: 105.

65. Croce O, Chevenet F, Christen R (2010) A New Web Server for the Rapid Identification of Microorganisms. J Microbial Biochem Technol 2: 84-88.

66. Bishayee A, Petit DM, Santani K (2010) Angioprevention is implicated in Resveratrol Chemoprevention of Experimental Hepatocarcinogenesis. J Carcinogene Mutagene 1: 104.

67. Abdullah S, Sameer S A, Dil-Afroz, Syeed N, Das BC, et al. (2010) PS3- The Molecular Guardian Crashes in Gastric Adenocarcinomas - A Study in an Ethnic Kashmiri Population. J Carcinogene Mutagene 1: 106.

68. Sellamuthu R, Umbright C, Chapman R, Leonard S, Li S, et al. (2011) Transcriptomics Evaluation of Hexavalent Chromium Toxicity in Human Dermal Fibroblasts. J Carcinogene Mutagene 2: 116.

69. Banerjee HN, Gramby M, Hawkins Z (2011) Molecular Diagnosis of Helicobacter Pylori Strain by 16S rDNA PCR Amplification and Direct Sequencing. J Bioprocess Biotechniq 1: 105e.

70. Rittmann BE (2006) Microbial ecology to manage processes in environmental biotechnology. Trends Biotechnol 24: 261-266.