© 2017 IJSRSET | Volume 3 | Issue 6 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section: Science and Technology

Bioinformatics in Environmental Bioremediation - A Review

Neerja Shukla
Department of Chemistry, University of Allahabad, Allahabad, Uttar Pradesh, India

ABSTRACT

Bioremediation is a technology that uses metabolic processes to degrade or transform contaminants, so that they remain no longer in harmful form. Microbial mediated bioremediation has a great potential to effectively restore contaminated environment, but the lack of information about factor regulating the growth and metabolism of various microbial communities in polluted environment often limits its implementation. Newly seeded technology such as Bioinformatics offer remarkable promise as tools to address longstanding questions regarding the molecular mechanisms involved in the control of mineralization pathways. Bioinformatics requires the study of microbial genomics, proteomics, interactomics, computational biology, phylogeny and application of bioinformatics tools for determining the structure and development of databases which would be further utilized for better upliftment. The bioinformatics web-servers plays key role in the application of bioremediation. This paper highlights the significance of bioinformatics concepts utilized over the bioremediation fields.

Keywords: Bioremediation, Bioinformatics, RNA, Proteomics, Genomics, Interactomics, Computational Biology, Phylogeny

I. INTRODUCTION

Environmental pollutants have become a major global concern, given their undesirable recalcitrant and xenobiotic compounds. A variety of polynuclear aromatic hydrocarbons (PAHs), xenobiotics, chlorinated and nitro-aromatic compounds were depicted to be highly toxic, mutagenic and carcinogenic for living organisms [1, 2]. As a result of their diversity, versatility and adaptability, a number of microorganisms are considered to be the best candidates among all living organisms to remediate most of the environmental contaminants into the natural biogeochemical cycle. These microorganisms display a remarkable range of contaminant degradable ability [3] that can efficiently restore natural environmental conditions [4, 5]. However, a variety of contaminants have been shown to be unusually recalcitrant, i.e. microorganisms either do not metabolize or transform them into certain other metabolites that again accumulate in the environment [6]. Therefore, it may be more productive to explore new catabolic pathways that might lead towards complete mineralization of these pollutants. One of the reasons, our knowledge of microbial degradation pathways is so incomplete is the immense complexity of microbial physiology that allows response and adaptability to various internal and external stimuli.

An understanding of these molecular and physiological mechanisms in any site-specific bacterium is fundamental for rational development, and the study of RNA and protein expression patterns has made and continues to make, critical contributions [7–11]. Dramatic advances in the nature and molecular technologies are making a global gene expression profile possible; genome-wide analysis of DNA (genomics), RNA expression (transcriptomics) and protein expression (proteomics) as well as exploring complexes of protein aggregation such as protein–protein interaction (interactomics) create the opportunity to systemically study the physiological expressions of an organisms.

Gene microarrays and proteomics technologies have recently been refined and made commercially available. The functional genomics approach in biological science has revolutionized research at the forefront of

environmental cleanup [12]. This technology was set up to elucidate the role of the many genes of unknown function that have been identified in the numerous genome sequencing projects [13–15]. In response to different environmental conditions, the genome and proteome directly reflect the physiological status of an organism.

Bioremediation offers many interesting possibilities from a bioinformatics point of view still slightly explored. This discipline requires the integration of huge amounts of data from various sources: chemical structure and reactivity of organic compounds; sequence, structure and function of proteins (enzymes); comparative genomics; environmental microbiology and so on. The accumulation of huge amounts of data on individual genes and proteins allowed the first studies of biology from a 'Systems' perspective (61–65). From this point of view, biological systems are modeled as being composed of components in complex relationships whose ultimate properties cannot be understood by studying these components separately and later 'summing' their properties, but only by studying the system as a whole. In a similar manner, data related to bioremediation (genome sequences, structures of chemical compounds, enzyme sequences and structures, etc.) are being accumulated in public databases (66). This allows the first studies of bioremediation from Systems Biology perspective (67,68), which complement the traditional approach focused on individual components (microorganisms, enzymes, etc.). The bioinformatics resources devoted to bioremediation are still scarce.

Bioremediation

Bioremediation is a technology that uses metabolic processes to degrade or transform contaminants, so that they remain no longer in harmful form. In some cases, the contaminant is the primary part of the metabolic process, acting as a main source of carbon and energy for the microbial cell. In others, it is transformed into a second substance, serves as a primary energy or carbon source. Recently, biological remediation process has also been devised to either precipitate effectively or immobilize inorganic pollutants such as heavy metals [1].

Intrinsic bioremediation technique has gained popularity, as the contaminant in the place itself and cuts down the excavation cost. Intrinsic bioremediation is done in-situ and relies on naturally occurring biological processes carried out by indigenous microorganisms. Intrinsic bioremediation is a component of natural attenuation, which includes physical and chemical processes (Hinchee and Wilson 1995). This technique is very successful in organically polluted soil especially with PAHs [5]. Bio-stimulation is the addition of nutrients (usually sources of carbon, nitrogen, phosphorus), oxygen or other electron donors or acceptors helps to increase the number or activity of naturally occurring microorganisms available for bioremediation. Another type is Bio-augmentation is the addition of microorganisms that can bio-transform or biodegrade a particular contaminant. Such process can be enhanced by the continuous addition of microorganisms to a bioreactor for the treatment of ground waters.

Bioinformatics

Bioinformatics has evolved into a full-fledged multidisciplinary subject that integrates developments in information and computer technology as applied to Biotechnology and Biological Sciences. Bioinformatics uses computer software tools for database creation, data management, data warehousing, data mining and global communication networking.

Bioinformatics is the recording, annotation, storage, analysis, and searching/retrieval of nucleic acid sequence (genes and RNAs), protein sequence and structural information. This includes databases of the sequences and structural information as well methods to access, search, visualize and retrieve the information. Bioinformatics concern the creation and maintenance of databases of biological information whereby researchers can both access existing information and submit new entries. Function genomics, biomolecular structure, proteome analysis, cell metabolism, biodiversity, downstream processing in chemical engineering, drug and vaccine design are some of the areas in which Bioinformatics is an integral component. There are three important sub-disciplines within bioinformatics involving computational biology:

- The development of new algorithms and statistics with which to assess relationships among members of large data sets
- The analysis and interpretation of various types of data including nucleotide and amino acid sequences, protein domains, and protein structures and

- The development and implementation of tools that enable efficient access and management of different types of information

Potential of Bioinformatics

The potential of Bioinformatics in the identification of useful genes leading to the development of new gene products, drug discovery and drug development has led to a paradigm shift in biology and biotechnology-these fields are becoming more & more computationally intensive. The new paradigm, now emerging, is that all the genes will be known "in the sense of being resident in database available electronically", and the starting point of biological investigation will be theoretical and a scientist will begin with a theoretical conjecture and only then turning to experiment to follow or test the hypothesis. With a much deep understanding of the biological processes at the molecular level, the Bioinformatics scientist have developed new techniques to analyse genes on an industrial scale resulting in a new area of science known as 'Genomics'.

The shift from gene biology has resulted in the development of strategies-from lab techniques to computer programmes to analyse whole batch of genes at once. Genomics is revolutionizing drug development, gene therapy, and our entire approach to health care and human medicine.

The genomic discoveries are getting translated in to practical biomedical results through Bioinformatics applications. Work on proteomics and genomics will continue using highly sophisticated software tools and data networks that can carry multimedia databases. Thus, the research will be in the development of multimedia databases in various areas of life sciences and biotechnology. There will be an urgent need for development of software tools for data-mining, analysis and modelling, and downstream processing. Security of data, data transfer and data compression, auto checks on data accuracy and correctness will also be major research area of bioinformatics. The use of virtual Reality in drug design, metabolic pathway design, and unicellular organism design, paving the way to design and modification of multi-cellular organisms, will be the challenges which Bioinformatics scientist and specialist have to tackle. It has now been universally recognized that Bioinformatics is the key to the new grand data-intensive molecular biology that will take us into 21 century.

Proteomics

The name proteomics is traditionally associated with the display of large sets of protein from a given organism or cell line on two-dimensional (2D) polyacrylamide gels. The ability to associate a spot on a 2D gel with known proteins is used to create databases of proteins that are expressed in an organism or cell line under defined experimental condition. This approach is complementary to the generation of databases of mRNA expression levels by microarray hybridization. So proteomics are the large scale study of protein properties such as expression levels, post-translational modifications and interactions with other molecules to obtain a global view of cellular processes at the protein level. Because the tools for high throughput DNA and RNA analysis are not available for protein analysis, the emphasis of functional genomics has been on the mRNA message. So it is the product of mRNA, i.e. the protein, which actually carries out the majority of the reaction of the cell. Proteomics is therefore a complementary approach to genomics and mRNA expression using microarrays. The activity of most associated proteomics is fractionating and visualizing large number of proteins from cell on two-dimensional (2D) polyacrylamide gel. These type of experiment have been performed from more than Twenty years to build databases of proteins expressed from certain cell or tissues type (Anderson and Anderson; 1996; O'Farrell; 1975).

Role of Proteomics in Bioremediation

The cellular expression of proteins in an organism varies with environmental conditions. The changes in physiological response may occur due to the organism's adaptive responses to different external stimuli, such as the presence of toxic chemicals in the environment. The advent of proteomics has allowed an extensive examination of global changes in the composition or abundance of proteins, as well as identification of key proteins involved in the response of microorganisms in a given physiological state [46, 47]. A number of reports have described sets of proteins that are up- or down-regulated cells. In agreement with the sequential regulation of in response to the presence of specific pollutants[46–48].

Applications of proteomics to study the physiology of B. s ubtilis have been reviewed [69]. Extensive proteomic work has been performed to understand the tolerance of bacilli to extreme environments [70–73], the allocation of stimulons and regulons [74, 75], and biofilm formation [76, 77], as well as full proteome [78–80] and secretome [81–83] mapping.

Genomics

The study of genes and their role in an organism's structure; growth health disease or resistance to disease etc. are known as Genomics. Genomics is now becoming a powerful computer technology which is used to understand the structure and function of all genes in the organism which is based on knowing the organism's entire DNA sequence. The field of genomics determines the entire sequence of DNA of an organism and fine-scale genetic mapping efforts. Genomics also depends over the study of intragenomic phenomena i.e. heterosis; epistasis; pleiotropy other interaction between Loci and Alleles within the genome.

Role of Genomics for Bioremediation

The regulation of gene expression is one of the key processes for adapting to changes in environmental conditions and thus for survival. DNA microarrays are extremely powerful platform that enable determination of the mRNA expression level of every gene of an organism [16–18]. The most challenging issue in microarray experiments is elucidation of data [19]. Often, hundreds of genes may be up- and/or down-regulated in a particular stress condition.

Even with the complete genome sequences of microorganisms with the potential for bioremediation [17, 20–23], studies are not accelerating in a rapid manner. With the completed genome sequences, it is possible to analyse the expression of all genes in each genome under various environmental conditions using whole-genome DNA microarrays [14, 24, 25]. Such genome-wide expression analysis provides important data for identifying regulatory circuits in these organisms [12, 22, 23]. In the past, DNA microarrays have been used to evaluate the physiology of pure environmental cultures [7] and to monitor the catabolic gene expression profile in mixed microbial communities [26]. More than 100 genes were found to be affected by oxygen-limiting conditions when a DNA microarray was used to study changes in mRNA expression levels in Bacillus subtilis grown under anaerobic conditions [27]. Sensitivity may often be a part of the problem in PCR-based cDNA microarrays, since only genes from populations contributing to more than 5% of the community DNA can be detected. Several parameters were evaluated to validate the sensitivity of spotted oligonucleotide DNA microarrays and their applicability for bacterial functional genomics [28].

Interactomics

Genome-wide mRNA profiling is unable to provide any information about the activity, arrangement, or final destination of the gene products, the proteins. Various proteomic approaches, on the other hand, can successfully provide the straight answers. It is very rare that any protein molecule acts as a unique pillar during the physiological response in bioremediation process of any contaminant when cellular proteins and various other related cellular expressions are on crest [11, 14, 15, 52]. In general, cellular life is organized through a complex protein interaction network, with many proteins taking part in multicomponent protein aggregation. The detection of these aggregated proteins, i.e. 'interactomics', is usually based upon affinity tag/pull down/MS/MS approaches at a proteome level [53–55]. Studies on protein–protein interaction and super molecular complex formation represent one of the main directions of functional proteomics.

Genomics versus Proteomics and Interactomics for Bioremediation

Based on an overall analysis of genomics and proteomics, the comprehensive analysis of whole-genome sequencing is especially helpful to understand bioremediation-relevant microorganisms whose physiology has not yet been studied in detail. Global gene expression using DNA microarray technology, very much depends on the degree of coverage of the cellular mRNA and cellular proteins, whereas the coverage of the whole genome represents all the genes of an organism by definition. Cellular mRNA levels do not display as wide a dynamic range as the encoded proteins [58]. Thus, whole genome arrays are believed to provide a much more comprehensive overview of the actual gene expression pattern than proteomic studies. According to global gene expression studies, both

genomics and proteomics support the view that the DNA array technologies record changes in gene expression more completely than the proteomics [11, 14, 59]. Therefore, genomics data is deemed necessary to complement the proteomics approach [60].

However, proteomics would retain its central position in functional genomics. The protein molecules, but not the mRNAs, are the key players in an on-site microbial mineralization reaction; the later are one of the highly unstable transmitters on the path from the genes to the ribosome, but each protein molecule represents the end product of gene expression [11]. Complete protein profiling provides not only information on the individual organism, but also information on the fate and destination of protein molecules inside and outside the cell that can only be discovered via a joint genomics, proteomics and interactomics approach.

Computational Biology

Bioinformatics and computational biology each maintain close interactions with life sciences to realize their full potential. Bioinformatics applies principles of information sciences and technologies to make the vast, diverse, and complex life sciences data more understandable and useful. Computational biology uses mathematical and computational approaches to address theoretical and experimental questions in biology. Although bioinformatics and computational biology are distinct, there is also significant overlap and activity at their interface.

Phylogeny

Phylogenetic inference is the process of developing hypotheses about the evolutionary relatedness of organisms based on their observable characteristics. Traditionally, phylogenetic analyses have been based on the gross anatomy of species. When Linneaus developed the system of classification into kingdoms, phyla, genera, and species, the early biologists sorted living things into a symbolic Tree of Life. This tree-based representation of the relationships among species is a phylogenetic tree; it has since been adopted as a convenient schematic for depicting evolutionary relatedness based on sequence similarity. The quantitative nature of sequence relationships has allowed the development of more rigorous methods and rules for tree drawing.

Conception of evolutionary relationships, modern phylogenetic trees are strictly binary; that is, at any branch point, a parent branch splits into only two daughter branches. Binary trees can approximate any other branching pattern, and the assumption that trees are binary greatly simplifies the tree-building algorithms. The length of branches in a quantitative phylogenetic tree can be determined in more than one way. Evolutionary distance between pairs of sequences, relative to other sequences in an input data set, is one way to assign branch length.

While a phylogeny of species generally has a root, assuming that all species have a specific common ancestor, a phylogenetic tree derived from sequence data may be rooted or unrooted. It isn't too difficult to calculate the similarity between any two sequences in a group and to determine where branching points belong. It is much harder to pinpoint which sequence in such a tree is the common ancestor, or which pair of sequences can be selected as the first daughters of a common ancestor.

II. METHODS AND MATERIAL

Major categories of Bioinformatics Tools:

There are both standard and customized products to meet the requirements of particular projects. There are data-mining software that retrieves data from genomic sequence databases and also visualization tools to analyze and retrieve information from proteomic databases. These can be classified as homology and similarity tools, protein functional analysis tools, sequence analysis tools and miscellaneous tools.

Here is a brief description of a few of these, everyday bioinformatics is done with sequence search programs like BLAST, sequence analysis programs, like the EMBOSS and Staden packages, structure prediction programs like THREADER or PHD or molecular imaging/modelling programs like RasMol and WHATIF.

- **Homology and Similarity Tools:**

Homologous sequences are sequences that are related by divergence from a common ancestor. Thus the degree of similarity between two sequences can be measured while their homology is a case of being either true of false. This set of tools can be used to identify similarities between novel query sequences of unknown structure and function and database sequences whose structure and function have been elucidated.

❖ **Protein Function Analysis:**

These groups of programs allow you to compare your protein sequence to the secondary (or derived) protein databases that contain information on motifs, signatures and protein domains. Highly significant hits against these different pattern databases allow you to approximate the biochemical function of your query protein.

❖ **Structural Analysis:**

These sets of tools allow you to compare structures with the known structure databases. The function of a protein is more directly a consequence of its structure rather than its sequence with structural homologs tending to share functions. The determination of a protein's 2D/3D structure is crucial in the study of its function.

❖ **Sequence Analysis:**

This set of tools allows you to carry out further, more detailed analysis on your query sequence including evolutionary analysis, identification of mutations, hydropathy regions, CpG islands and compositional biases. The identification of these and other biological properties are all clues that aid the search to elucidate the specific function of your sequence.

Bioinformatics Web Server used for Bioremediation

❖ **MetaRouter**

MetaRouter, a system for maintaining heterogeneous information related to bioremediation in a framework that allows its query, administration and mining (application of methods for extracting new knowledge). MetaRouter is an application intended for laboratories working in biodegradation and bio- remediation, which need to maintain and consult public and private data, linked internally and with external databases, and to extract new information from it. Among the data-mining features is a program included for locating biodegradative pathways for chemical compounds according to a given set of con- straints and requirements. The integration of biodegradation information with the corresponding protein and genome data provides a suitable framework for studying the global properties of the bioremediation network. The system can be accessed and administrated through a web interface.

The University of Minnesota Biocatalysts/ Biodegradation Database (http://www.labmed. umn.edu/umbbd) begins its fifth year having met its initial goals. It contains approximately 100 pathways for microbial catabolic metabolism of primarily xenobiotic organic compounds, including information on approximately 650 reactions, 600 compounds and 400 enzymes, and containing approximately 250 microorganism entries. It includes information on most known microbial catabolic reaction types and the organic functional groups they transform. Having reached its first goals, it is ready to move beyond them. It is poised to grow in many different ways, including mirror sites; fold prediction for its sequenced enzymes; closer ties to genome and microbial strain databases; and the prediction of biodegradation pathways for compounds it does not contain (Ellis et al. 2000).

❖ **BASys**

BASys (Bacterial Annotation System) is a web server that supports automated, in-depth annotation of bacterial genomic (chromosomal and plasmid) sequences. It accepts raw DNA sequence data and an optional list of gene identification information and provides extensive textual annotation and hyper- linked image output. BASys uses >30 programs to determine $60 annotation subfields for each gene, including gene/protein name, GO function, COG function, possible paralogues and orthologues, molecular weight, isoelectric point, operon structure, sub- cellular localization, signal peptides, transmembrane regions, secondary structure, 3D structure, reactions and pathways.

A fully navigable graphical map, which is hyper- linked to textual gene descriptions, is also generated to allow results to be easily browsed and evaluated. BASys' strengths are in its web accessibility, its depth and comprehensiveness of annotation and its user-friendly

graphical interface. BASys does not (yet) handle partially assembled genomes nor does it offer the sophisticated ORF handling of stand-alone programs like GenDB or Magpie. The BASys web server is freely accessible at http://wishart.biology.ualberta.ca/basys.

❖ NAST

The NAST (Nearest Alignment Space Termination) web server is available for creating MSAs of small and large 16S rRNA gene sequencing projects. NAST allows retention of fixed MSA column counts regardless of the quantity of records added to a profile alignment. This permit ongoing MSA curation, and supports collaborative efforts in comparative genomics. NAST and supporting tools at the Greengenes website enable microbiologists to rapidly compare sampled sequences to publicly available reference sequences as well as to each other.

III. CONCLUSION

Till the long study it was found that Bioremediation is a new technique in Bioinformatics particularly genomics, proteomics and interactomics which offer remarkable tools to study the mechanism involved in the regulation of the mineralization pathways. The application of this technique are still been in their first stage, but the amount of data that is continuously being generated by today's genomics and proteomics technocrats needs to be organised in a step wise manner within information databases. The strategies need to be refined in which genomics and proteomics data are combined together in order to understand the mineralization process in a meaningful way. These techniques show great promise in their ability to predict organisms metabolism in contaminated environments and to predict the microbial assisted attenuation of contaminants to accelerate bioremediation.

IV. REFERENCES

[1]. Zhang C, Bennett GN. Biodegradation of xenobiotics by anaerobic bacteria. Appl Microbiol Biotechnol 2005;67: 600-18.

[2]. Samanta SK, Singh OV, Jain RK. Polycyclic aromatic hydrocarbons: environmental pollution and bioremediation. Trends Biotechnol 2002;20:243-8.

[3]. Pandey G, Paul D, Jain RK. Branching of o-nitrobenzoate degradation pathway in Arthrobacter protophormiae RKJ100: identification of new intermediates. FEMS Microbiol Lett2003; 229:231-6.

[4]. Labana S, Singh OV, Basu A, et al. A microcosm study on bioremediation of p-nitrophenol contaminated soil using Arthrobacter protophormiae RKJ100. Appl Microbiol Biotechnol 2005;68:417-24.

[5]. Labana S, Pandey G, Paul D, etal. Plot and field studies on bioremediation of p-nitrophenol contaminated soil using Arthrobacter protophormiae RKJ100. Environ Sci Technol 2005; 39:3330-7.

[6]. Esteve-Nunez A, Caballero A, Ramos JL. Biological degradation of 2,4,6-trinitrotoluene. Microbiol Mol Biol Rev 2001;65:335-52.

[7]. Schut GJ, Zhou J, Adams MW. DNA microarray analysis of the hyperthermophilic archaeon Pyrococcus furiosus: evidence for a new type of sulfur-reducing enzyme complex. JBacteriol 2001;183 :7027-36.

[8]. Krivobok S, Kuony S, Meyer C, et al. Identification of pyrene-induced proteins in Mycobacterium spp. strain 6PY1: evidence for two ring-hydroxylating dioxygenases. JBacteriol 2003;185 :3828-41.

[9]. Rhee SK, Liu X, Wu L, etal. Detection of genes involved in biodegradation and biotransformation in microbial communities by using 50-mer oligonucleotide microarrays. Appl EnvironMicrobiol 2004;70:4303-17.

[10]. Kim SJ, Jones RC, Cha CJ, et al. Identification of proteins induced by polycyclic aromatic hydrocarbon in Mycobacterium vanbaalenii PYR-1 using two-dimensional polyacrylamide gel electrophoresis and de novo sequencing methods. Proteomics 2004;4:3899-908.

[11]. Kuhner S, Wohlbrand L, Fritz I, et al. Substrate-dependent regulation of anaerobic degradation pathways for toluene and ethylbenzene in a denitrifying bacterium, strain EbN1. JBacteriol 2005;18 7:1493-503.

[12]. Lovley DR. Cleaning up with genomic: applying molecular biology to bioremediation. Nat Rev Mi crobiol 2003;1:35-44.

[13]. Santos PM, Benndorf D, Sa-Correia I. Insights into Pseudomonas putida KT2440 response to

phenol-induced stress by quantitative proteomics. Proteomics 2004;4: 2640-52.
[14]. Gao H, Wang Y, Liu X, etal. Global transcriptome analysis of the heat shock response of Shewanella oneidensis. JBacteriol 2004;186 :7796-803.
[15]. Eyers L, George I, Schuler L, et al. Environmental genomics: exploring the unmined richness of microbes to degrade xenobiotics. Appl Microbiol Biotechnol 2004;66: 123-30.
[16]. Schena M, Heller RA, Theriault TP, et al. Microarrays: biotechnology's discovery platform for functional genomics. Trends Biotechnol 1998;16:301-6.
[17]. Golyshin PN, Martins Dos Santos VA, Kaiser O, et al. Genome sequence completed of Alcanivorax borkumensis,a hydrocarbon-degrading bacterium that plays a global role in oil removal from marine systems. JBiotechnol2003;10 6: 215-20.
[18]. Diaz E. Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility. Int Microbiol 2004;7:173-80.
[19]. Dharmadi Y, Gonzalez R. DNA microarrays: experimental issues, data analysis, and application to bacterial systems. Biotechnol Prog 2004;20:1309-24.
[20]. Tiedje JM. Shewanella—the environmentally versatile genome. Nat B iotechnol 2002;20:1093-4.
[21]. Heidelberg JF, Paulsen IT, Nelson KE, et al. Genome sequence of the dissimilatory metal ion-reducing bacterium Shewanella oneidensis. Nat Biotechnol 2002;20:1118-23.
[22]. Seshadri R, Adrian L, Fouts DE, et al. Genome sequence of the PCE-dechlorinating bacterium Dehalococcoides ethenogenes. Science 2005;307:105-8.
[23]. Rabus R, Kube M, Heider J, etal. The genome sequence of an anaerobic aromatic-degrading denitrifying bacterium, strain EbN1. Arch Microbiol 2005;183:27-36.
[24]. Muffler A, Bettermann S, Haushalter M, et al. Genome- wide transcription profiling of Corynebacterium glutamicum after heat shock and during growth on acetate and glucose. J Biotechnol 2002;98:255-68.
[25]. Schut GJ, Brehm SD, Datta S, et al. Whole-genome DNA microarray analysis of a hyperthermophile and an archaeon: Pyrococcus furiosus grown on carbohydrates or peptides. J Bacteriol 2003;18 5:3935-47.
[26]. Dennis P, Edwards EA, Liss SN, et al. Monitoring gene expression in mixed microbial communities by using DNA microarrays. Ap pl E nviron Microbiol 2003;69:769-78.
[27]. Ye RW, Tao W, Bedzyk L, et al. Global gene expression profiles of Bacillus subtilis grown under anaerobic conditions. J Bacteriol 2000;18 2:4458-65.
[28]. Denef VJ, Park J, Rodrigues JL, et al. Validation of a more sensitive method for using spotted oligonucleotide DNA microarrays for functional genomics studies on bacterial communities. E nviron Microbiol 2003;5:933-43.
[29]. Cho JC, Tiedje JM. Quantitative detection of microbial genes by using DNA microarrays. Appl Environ Microbiol 2002;68:1425-30.
[30]. Greene EA, Voordouw G. Analysis of environmental microbial communities by reverse sample genome probing. JMicrobiol Methods 2003;53:211-9.
[31]. Zhou J, Thompson DK. Challenges in applying microarrays to environmental studies. Curr Opin Biotechnol 2002;13: 204-7.
[32]. Burgmann H, Widmer F, Sigler WV, et al. mRNA extraction and reverse transcription-PCR protocol for detection of nifH gene expression by Azotobacter vinelandii in soil. Appl Environ Microbiol 2003;69:1928-35.
[33]. Urakawa H, El Fantroussi S, Smidt H, et al. Optimization of single-base-pair mismatch discrimination in oligo- nucleotide microarrays. Appl Environ Microbiol 2003;69: 2848-56.
[34]. Wu L, Thompson DK, Li G, et al. Development and evaluation of functional gene arrays for detection of selected genes in the environment. Appl Environ Microbiol 2001;67: 5780-90.
[35]. Wasinger VC, Cordwell SJ, Cerpa-Poljak A, et al. Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. Electrophoresis 1995;16:1090-4.
[36]. Hochstrasser DF. Proteome in perspective. Clin Chem Lab Med 1998;36:825-36.
[37]. Sikkema J, deBont JAM, Poolman B. Mechanisms of membrane toxicity of hydrocarbons. Microbiological Rev1995;59 :201-22.
[38]. Paoletti AC, Zybailov B, Washburn MP. Principles and applications of multidimensional

protein identification technology. Expert Rev Proteomics 2004;1:275-82.
[39]. Aebersold R, Mann M. Mass spectrometry-based proteomics. Nature 2003;422:198-207.
[40]. Aitken A, Learmonth M. Protein identification by in-gel digestion and mass spectrometric analysis. Mol Biotechnol 2002;20:95-7.
[41]. Landry F, Lombardo CR, Smith JW. A method for application of samples to matrix-assisted laser desorption ionization time-of-flight targets that enhances peptide detection. Anal Biochem 2000;279:1-8.
[42]. Merchant M, Weinberger SR. Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. Electrophoresis 2000;21:1164-77.
[43]. Seibert V, Ebert MP, Buschmann T. Advances in clinical cancer proteomics: SELDI-ToF-mass spectrometry and bio- marker discovery. BriefFunct Genomic Proteomic 2005;4:16-26.
[44]. Knigge T, Monsinjon T, Andersen OK. Surface-enhanced laser desorption/ionization-time of flight-mass spectrometry approach to biomarker discovery in blue mussels (Mytilus edulis) exposed to polyaromatic hydrocarbons and heavy metals under field conditions. Proteomics 2004;4:2722-7.
[45]. Joo WA, Kim CW. Proteomics of Halophilic archaea. JChromatogr BAnalytTechnolBiomed Life Sci 2005;815:237-50.
[46]. Vasseur C, Labadie J, Hebraud M. Differential protein expression by Pseudomonas fragi submitted to various stresses. Electrophoresis 1999;20:2204-13.
[47]. Wilkins JC, Homer KA, Beighton D. Altered protein expression of Streptococcus oralis cultured at low pH revealed by two-dimensional gel electrophoresis. Appl Environ Microbiol 2001;67:3396-405.
[48]. Kim S II, Kim SJ, Nam MH, et al. Proteome analysis of aniline-induced proteins in Acinetobacter lwoffi K24. Curr Microbiol 2002;44:61-6.
[49]. Johnsen AR, Wick LY, Harms H. Principles of microbial PAH-degradation in soil. Environ Pollut 2005;133:71-84.
[50]. Wang RF, Wennerstrom D, Cao WW, et al. Cloning, expression, and characterization of the katG gene, encoding catalase-peroxidase, from the polycyclic aromatic hydro- carbon-degrading bacterium Mycobacterium sp. strain PYR-1. Appl EnvironMicrobiol 2000;66:4300-4.
[51]. Khan AA, Wang RF, Cao WW, et al. Molecular cloning, nucleotide sequence, and expression of genes encoding a polycyclic aromatic ring dioxygenase from Mycobacterium sp. strain PYR-1. Appl EnvironMicrobiol 2001; 67:3577-85.
[52]. Segura A, Godoy P, van Dillewijn P, et al. Proteomic analysis reveals the participation of energy- and stress-related proteins in the response of Pseudomonas putida DOT-T1E to toluene. JBacteriol 2005;187 :5937-45.
[53]. Lee WC, Lee KH. Applications of affinity chromatography in proteomics. Anal Biochem 2004;324:1-10.
[54]. Coulombe B, Jeronimo C, Langelier MF, et al. Interaction networks of the molecular machines that decode, replicate, and maintain the integrity of the human genome. Mol Cell Prot eomics 2004;3:851-6.
[55]. Gingras AC, Aebersold R, Raught B. Advances in protein complex analysis using mass spectrometry. JPhysiol 2005; 563:11-21.
[56]. Labaer J, Ramachandran N. Protein microarrays as tools for functional proteomics. Curr Opin Chem Biol 2005; 9:14-9.
[57]. Liu WT, Zhu L. Environmental microbiology-on-a-chip and its future impacts. Trends Biotechnol 2005;23:174-9.
[58]. Gygi SP, Rochon Y, Franza BR, etal. Correlation between protein and mRNA abundance in yeast. Mol Cell Biol 1999; 19:1720-30.
[59]. Eymann C, Homuth G, Scharf C, et al. Bacillus subtilis functional genomics: global characterization of the stringent response by proteome and transcriptome analysis. JBacteriol 2002;18 4:2500-20.
[60]. Hegde PS, White IR, Debouck C. Interplay of transcriptomics and proteomics. Curr Opin Biotechnol 2003;14: 647-51.
[61]. Ideker,T., Thorsson,V., Ranish,J.A., Christmas,R., Buhler,J., Eng,J.K., Bumgarner,R., Goodlett,D.R., Aebersold,R. and Hood,L. (2001) Integrated genomic and proteomic analysesof a systematically perturbed metabolic network. Science, 292, 929-934.
[62]. Jeong,H., Mason,S.P., Baraba si,A.L. and Oltvai,Z.N. (2001) Lethality and centrality in protein networks. Nature, 411, 41-42.

[63]. Alves,R., Chaleil,R.A.G. and Sternberg,M.J.E. (2002) Evolution of enzymes in metabolism: a network perspective. J. Mol. Biol., 320, 751-770.
[64]. Fraser,H.B., Hirsh,A.E., Steinmetz,L.M., Scharfe,C. and Feldman,M.W.(2002) Evolutionary rate in the protein interaction network. Science, 296, 750-752.
[65]. Rison,S.G.C. and Thornton,J.M. (2002) Pathway evolution, structurally speaking. Curr. Opin. Struct. Biol., 12, 374-382.
[66]. Hecker M, Volker U. Towards a comprehensive understanding of Bacillus subtilis cell physiology by physiological proteomics. Proteomics 2004;4:3727-50.
[67]. Antelmann H, Scharf C, Hecker M. Phosphate starvation inducible proteins of Bacillus subtilis: Proteomics and transcriptional analysis. Journal of Bacteriology 2000;182: 4478-90.
[68]. Bernhardt J, Weibezahn J, Scharf C, et al. Bacillus subtilis during feast and famine: visualization of the overall regulation of protein synthesis during glucose starvation by proteome analysis. Genome Res 2003;13:224-37.
[69]. Eymann C, Mach H, Harwood CR, et al. Phosphate- starvation-inducible proteins in Bacillus subtilis: A two- dimensional gel electrophoresis study. Microbiology-UK1996;142:3163-70.
[70]. Hoper D, Bernhardt J, Hecker M. Salt stress adaptation of Bacillus subtilis: a physiological proteomics approach. Prot eomics 2006;6:1550-62.
[71]. Hecker M, Engelmann S. Proteomics, DNA arrays and the analysis of still unknown regulons and unknown proteins of Bacillus subtilis and pathogenic Gram-positive bacteria. IntJ Med Microbiol 2000;290:123-34.
[72]. Tam le T, Antelmann H, Eymann C, et al. Proteome signatures for stress and starvation in Bacillus subtilis as revealed by a 2-D gel image color coding approach. Prot eomics 2006;6:4565-85.
[73]. Morikawa M, Kagihiro S, Haruki M, et al. Biofilm formation by a Bacillus subtilis strain that produces gamma- polyglutamate. Microbiology-SGM 2006;152 :2801-7.
[74]. Oosthuizen MC, Steyn B, Theron J, et al. Proteomic analysis reveals differential protein expression by Bacillus cereus during biofilm formation. Appl Environ Microbiol2002;68:2770-80.
[75]. Buttner K, Bernhardt J, Scharf C, et al. A comprehensive two-dimensional map of cytosolic proteins of Bacillus of the versatile microbe Rhodopseudomonas palustris subtilis. Electrophoresis 2001;22:2908-35.
[76]. Eymann C, Dreisbach A, Albrecht D, etal. A comprehensive proteome map of growing Bacillus subtilis cells. Proteomi cs 2004;4:2849-76.
[77]. Wolff S, Antelmann H, Albrecht D, et al. Towards the entire proteome of the model bacterium Bacillus subtilis by gel-based and gel-free approaches. J Chromatogr B Analyt Te c h n o l B i o m e d L i f e S ci 2007;849:129-40.
[78]. Lamonica JM, Wagner MA, Eschenbrenner M, et al. Comparative secretome analyses of three Bacillus anthracis strains with variant plasmid contents. Infect Immun 2005;73: 3646-58.
[79]. Tjalsma H, Antelmann H, Jongbloed JDH, etal. Proteomics of protein secretion by Bacillus subtilis: Separating the ‘‘secrets’’ of the secretome. Mi crobiol Mol Biol Rev 2004;68: 207-33.
[80]. Antelmann H, Williams RC, Miethke M, et al. The extracellular and cytoplasmic proteomes of the non-virulent Bacillus anthracis strain UM23C1-2. Proteomics 2005;5: 3684-95.
[81]. Dassarma S, Berquist BR, Coker JA, et al. Post-genomics of the model haloarchaeon Halobacterium sp. NRC-1. Saline Syst 2006;2:3.
[82]. Kennedy SP, Ng WV, Salzberg SL, etal. Understanding the adaptation of Halobacterium species NRC-1 to its extremeenvironment through computational analysis of its genomesequence. Genome Res 2001;11:1641-50.
[83]. Joo WA, Kim CW. Proteomics of Halophilic archaea. JChromatogr BAnalytTechnolBiomed LifeSci 2005;815:237-50.
[84]. Kim SI, Choi JS, Kahng HY. A proteomics strategy for the analysis of bacterial biodegradation pathways. Omics 2007; 11 :280-94.
[85]. Krayl M, Benndorf D, Loffhagen N, etal. Use of proteomics and physiological characteristics to elucidate ecotoxic effects of methyl tert-butyl ether in Pseudomonas putida KT2440. Prot eomics 2003;3:1544-52.

[86]. Lupi CG, Colangelo T, Mason CA. Two-dimensional gel electrophoresis analysis of the response of Pseudomonas putida KT2442 to 2-Chlorophenol. Appl EnvironMicrobiol 1995;61: 2863-72.
[87]. Reardon KF, Kim KH. Two-dimensional electrophoresis analysis of protein production during growth of Pseudomonas putida F1 on toluene, phenol, and their mixture. Electrophoresis 2002;23:2233-41.
[88]. Arevalo-Ferro C, Reil G, Go¨rg A, et al. Biofilm formation of Pseudomonas putida IsoF: the role of quorum sensing as assessed by proteomics. Syst Appl Mi crobi ol 2005;28:87-114.
[89]. Sauer K, Camper AK. Characterization of phenotypic changes in Pseudomonas putida in response to surface- associated growth. JBacteriol 2001;183 :6579-89.
[90]. Sauer K, Camper AK, Ehrlich GD, et al. Pseudomonas aeruginosa displays multiple phenotypes during development as a biofilm. J Bacteriol 2002;184:1140-54.
[91]. Sauer K, Cullen MC, Rickard AH, et al. Characterization of nutrient-induced dispersion in Pseudomonas aeruginosa PAO1 biofilm. J Bacteriol 2004;186:7312-26.
[92]. Kertesz MA, Leisinger T, Cook AM. Proteins induced by sulfate limitation in Escherichia coli, Pseudomonas putida, or Staphylococcus aureus. J Bacteriol 1993;175:1187-90.
[93]. VerBerkmoes NC, Shah MB, Lankford PK, et al. Determination and comparison of the baseline proteomes under its major metabolic states. JProteomeRes 2006;5: 287-98.
[94]. Elias DA, Monroe ME, Marshall MJ, et al. Global detection and characterization of hypothetical proteins in Shewanella oneidensis MR-1 using LC-MS based proteomics. Proteomics 2005;5:3120-30.
[95]. Elias DA, Monroe ME, Smith RD, et al. Confirmation of the expression of a large set of conserved hypothetical proteins in Shewanella oneidensis MR-1. J Microbiol Methods2006;66:223-33.
[96]. Fang R, Elias DA, Monroe ME, etal. Differential label-free quantitative proteomic analysis of Shewanella oneidensis cultured under aerobic and suboxic conditions by accurate mass and time tag approach. Mol C el l Proteomics 2006;5: 714-25.
[97]. Masselon C, Pasa-Tolic L, Tolic N, et al. Targeted comparative proteomics by liquid chromatography- tandem Fourier ion cyclotron resonance mass spectrometry. Anal Chem 2005;77:400-6.
[98]. Mohan D, Pasa-Tolic L, Masselon CD, et al. Integration of electrokinetic-based multidimensional separation/concentration platform with electrospray ionization-Fourier transform ion cyclotron resonance-mass spectrometry for proteome analysis of Shewanella oneidensis. Anal Chem 2003; 75:4432-40.
[99]. Romine MF, Elias DA, Monroe ME, et al. Validation of Shewanella oneidensis MR-1 small proteins by AMT tag-based proteome analysis. OMICS 2004;8:239-54.