

## TRENDS IN BACTERIAL DIVERSITY STUDY AND ITS PROSPECTS

RATUL SAIKIA\*, RUPAK K. SARMA\*, RAJAL DEBNATH\* AND TARUN C. BORA\*

---

*Bacteria are the most abundant and species-rich groups of organisms, and they mediate many critical ecosystem processes. In spite of the ecological importance of bacteria, past practical and theoretical limitations have limited our ability to study the patterns of bacterial diversity and to understand the processes that determine these patterns. Recent advances in molecular techniques have focused on multi-locus approaches and on combined analysis of molecular and ecological data. In this article we outline some molecular techniques used to document bacterial diversity in the environment.*

---

**M**icrobial diversity is an unseen international resource that deserves greater attention. Soils sustain an immense diversity of microbes, of which, a large extent still remains unexplored. Bacteria are an important constituent of the soil microflora having large species diversity (a minimum of 4 000–7 000 different bacterial genomes per gram of soil). It is estimated that 30 g of forest soil contains over half a million species<sup>1</sup>. The bacterial species existing in natural ecosystems might be more than a thousand million or even as high as a trillion. This microbial biota have key role in maintaining the biogeochemical cycles and recycling the trace elements (iron, nickel, mercury, etc.) and are therefore heavily implicated in energy and nutrient exchanges within the soil.

Every bacterium in soil exhibits different characteristics which creates problems for the taxonomists for proper classification and identification. Microbial taxonomists, from its beginning, battled with many difficulties and often incorrect identification of newly discovered bacteria. For long time, only isolation of bacterial cells from natural samples and further characterization through biochemical or immunochemical techniques signified the only way for categorization of

microbes. From present point of view all these morphological methods are insufficient and incomplete. Again diversity study of soil bacterial constituents is hampered not only by the technical inability, high heterogeneity of the environment, with its changing temporal and spatial microhabitat also creates numerous problems<sup>2</sup>. Similarly, the very tiny morphology, adhesion property with soil, uneven distribution of nutrients or toxicants, genome changes markedly affects the diversity pattern<sup>3</sup>.

### ***Why to Explore Bacterial Diversity?***

Microorganisms represent a huge reservoir of resources providing innovative applications useful to human society. They are capable of exploiting a vast range of energy sources having adaptability in almost every habitat. In this new era of science people have to explore this unseen natural bioresource and to study the complexity and vitality of bacterial communities as well as their myriad medicinal, agricultural and industrial application. The opportunity for discovery of new industrial applications from microorganisms is as large as the variety of environments they confront. Much of the environmental variation experienced by microorganisms is produced by the chemical complexity of the world around them. It was estimated that single sediment may contain as many as 100000 distinct organic chemical structures<sup>4</sup>.

---

\* Biotechnology Division, CSIR-North-East Institute of Science & Technology, Jorhat 785 006, Assam, India  
(e-mail: saikia\_ratul@rii.jorhat.res.in; rsaikia19@gmail.com)

The majority of existing biotechnological applications are of microbial origin, and it is widely appreciated that the microbial world contains by far the greatest fraction of biodiversity in the biosphere. For 100 years onwards Pure cultures of microorganisms have successfully been used in different industrial field and its annual added value to the chemical industry alone is about \$10 billion and, in the next 10 years, 60% of synthetic fine chemicals will rely on microbial conversions<sup>5</sup>. Now efforts towards the industrial application of redox conversions have mainly focused on whole resting microbial cells; well-designed microbial routes to produce building blocks with a number of functional groups and chiral centers via a nonchiral or asymmetric synthesis, racemic and dynamic resolutions, and conversion of low-cost aromatic and non aromatic substrates into valuable alcohols and reactive intermediates, have successfully been achieved at the industrial scale <sup>6</sup>.

The rational combination of enzymatic catalysis and modern techniques of molecular biology creates new trend in industrial microbiology. New powerful natural enzymes can be rapidly found through metagenomics, customized for specific applications via directed evolution. In this area, a new chemo-enzymatic process for the production of glycolic acid has recently been established, to further improve the economical efficiency of the process the enzyme was cloned and efficiently over-expressed in *E. coli* cells. Metagenomics is a culture independent approach to sample and characterize microbial genomes and the enormous genetic pool of 1033 putative genes in the environment. In the past few years, 140 metagenome-derived enzymes have been described, some of which may have biotech potential.

In sustainable agriculture, the natural roles of microorganisms in maintaining soil fertility and biocontrol of plant pathogens may be more important than in conventional agriculture by high inputs of agrochemicals. Many studies of the influence of plant growth promoting rhizobacteria (PGPR) on proper growth and development of agricultural crops have been conducted; however, the mechanisms of interaction are still poorly understood. Novel approaches including PCR-based methods, stable isotope profiling, and molecular markers have begun to shed light on the activity, identity and role of bacteria in the sustainable agriculture.

### **Bacterial Diversity Analysis**

In recent decades, the modern approaches of bacterial taxonomy includes characterization of bacterial DNA or RNA for proper identification of bacteria, as well as illustration of bacterial DNA diversity in the case of

individual bacteria, or DNA of entire microbial communities, besides this, there are also other techniques based on analysis of characteristics and mostly highly conservative properties. Utilization of powerful technologies i.e. , MALDI-TOF mass spectrometry or analysis of the fatty acid methyl esters composition can accurately identify organisms to species level, that could provide the best choice for routine microbiology. Biochemical and immunochemical tests are the oldest methods of microbial taxonomy. The commercial kits exists now a days, allows determination of biochemical characteristics of individual pure cultures. The preliminary step in culture-based detection methods is to culture the microbial population in nutrient enrichment medium. Selective and differential media is now needed for isolation of the targeted bacterium and identified on the basis of multiple biochemical properties. Selective media are those allowing only certain organisms to grow because they have specific inhibitors added to the media (e.g., the bile salts in MacConkey agar). But the biochemical methods suffer from some major drawbacks. First, they can be used only for organisms that can be cultivated *in vitro*, secondly specific strains exhibit unique biochemical characteristics and cannot be considered as a characteristic of any known genus and species. Again since these are very time consuming approach taxonomists are not so much satisfied with these. Further identification performed by immunochemical techniques creates an additional drawback other than the time factor, is the necessity for special chemical consumables with the logistic burden and costs. These can increase operational costs by hundreds of dollars per hour.

Presently, broad-scale molecular approaches are applied for assessing the bacterial diversity. In the percentage of G + C profiling, DNA are extracted from the total community in the soil. Guanine plus cytosine abundance (G + C) of the DNA is measured through thermal denaturation process<sup>7</sup>, density gradient centrifugation, cross hybridization of different DNA samples and reassociation kinetics of denatured DNA. Since, different bacterial species bear different extent of the nucleotides (from ca. 24 to 76% G1C), %G+C distribution is a sensitive measure of bacterial diversity<sup>8</sup>.

Genetic fingerprint techniques are based on PCR amplification. Diversity of amplified sequences is simply resolved by differential electrophoretic migration on agarose or polyacrylamide gels, which depend on their size (ARDRA, t-RFLP, RISA, RAPD) or sequence (DGGE, TGGE). Complex banding patterns were generated, representing the genetic structure of the bacterial community. Since a single bacterial species may have

several different rRNA sequences and different bacterial species may have closely related rDNA sequences, this approach seems to be quite useful for molecular taxonomic studies<sup>9</sup>. The profile data can then be analyzed in terms of similarities and relationships can be represented by a dendrogram of similarities.

After the introduction of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbiology diversity these techniques are now routinely used in many laboratories worldwide as molecular tools to compare the diversity of microbial communities. However, the main limitation for both the methods is again the inability to use large parts of DNA. Sequences up to 500 bp are usually used. In addition, the resolution of the profiles obtained, in terms of band numbers, is not always sufficient to illustrate the considerable bacterial diversity in indigenous soil communities, and some studies have shown that fragments of different sequences might migrate at the same position<sup>10</sup>.

Amplified ribosomal DNA restriction analysis (ARDRA) is the extension of the technique of RFLP (restriction fragment length polymorphism), where the restriction analysis of amplified rDNA to the gene encoding the small (16S) ribosomal subunit of bacteria is carried out. The technique involves PCR based amplification of the ends of 16S rDNA gene using primers, followed by digestion using tetra cutter restriction enzymes. The banding Patterns obtained from three or more restriction enzymes can be used to phylogenetically characterize the cultured isolates. However, in this method, the complexity of the profile is great and no information can be deduced in detecting and quantifying a specific ribosomal pattern. Terminal RFLP (T-RFLP) takes advantage of the high resolution and throughput of automated sequencing technologies to split the polymorphic terminal fragments after restriction digestion. Because the polymorphism is based solely on the length of the fragment, direct reference can be made to the sequence database

Ribosomal intergenic spacer analysis (RISA) involves PCR amplification of a region of the rRNA gene operon between the small (16S) and large (23S) subunits called the intergenic spacer region ISR, i.e. 50 bp to more than 1.5 kb depending on the species. The various amplified sequences are directly separated on polyacrylamide gels on the basis of size. RISA has proven to be a reliable and relatively inexpensive method that offers a higher resolution when compared to 16S rDNA-based approaches.

DNA sequencing is the most prominent and reliable procedure for identification of most microorganisms.

Bacterial identification based on 16S rDNA gene sequencing, since the gene is universally found in all bacteria and is considered as evolution chronometer for use in new species identification. The biggest disadvantage is the inability to sequence fragments greater than 1000 bp; it is caused by the deficiencies in the distribution of higher molecular weight fragments. Other disadvantage is, of course, the high price of DNA analysis and requirements for the high purity of DNA. New sequencing technologies such as pyrosequencing (so called 454 technology) have been successfully used as rapid and efficient tools to enable in-depth analysis of bacterial composition and diversity of consortia of environmental microbes<sup>11</sup>.

In community analysis by fatty acid profiling (phospholipid fatty acids, PLFAs) are also widely accepted as biomarkers, indicates viable microbial biomass and provide a microbial community 'fingerprint'. Several PLFAs are found to be most prominent for the detection of specific groups and whole cell fatty acid analysis is even used for the routine identification of microbial species.

It was reported that the *in situ* evaluation of distribution and function of microorganisms through fluorescence *in situ* hybridization (FISH) where oligonucleotide are used as hybridization probes complementary to regions of the 16S rRNA gene. Similar to the PCR-based methods, this technique can be customized to target specific groups of organisms<sup>12</sup>.

Nowadays microarray have proven to be one of the most useful and high-throughput method to provide targeted DNA sequence information for up to many thousands of specific genetic regions in a single test. A microarray consists of multiple DNA oligonucleotide probes that, under high stringency conditions, hybridize only to specific complementary nucleic acid sequences (targets). A phylogenetic microarray, developed by the Andersen Laboratory, the Phylo Chip, is an example of a microarray that targets the known diversity within the 16S rDNA gene to determine microbial community composition. Using multiple, confirmatory probes, the PhyloChip is able to simultaneously identify any of thousands of taxa present in an environmental sample.

### ***Bacterial Diversity in CSIR-NEIST and North-East India***

North-East India is best known for its rich biodiversity and its untapped bioresources has been identified as the Indo-Burma Mega Hot Spot by Conservation International<sup>13</sup>. Researchers are now busy to explore the rich microbial biota from this area. North East

Institute of Science and Technology (CSIR-NEIST), Jorhat, Assam is presently engaged in exploration of rich microbial diversity as well as their potential application in industrial, pharmaceutical and agricultural field from this untouched region. *Aquimonas voraii*, *Microbacterium assamensis* and *Kocaria assamensis* are some of the novel strains identified by NEIST, Jorhat collaboration with IMTECH, Chandigarh from this region. Recently, NEIST has reported potential biocontrol agent, *Bravibacillus laterosporus* (BPM3) isolated from 'Borpung' a natural hot water spring of Assam was found to be most active against different phytopathogenic fungi and gram positive bacteria<sup>14</sup>. Researchers of this institute reported genetic diversity of antagonistic potential fluorescent pseudomonads from tea rhizospheric soil of Assam<sup>15</sup>. The microbial diversity exploration group of Biotechnology Division, NEIST, Jorhat have submitted about 100 bacterial sequence to NCBI-Gene Bank, which includes potential antibiotic producing *Streptomyces* spp., extremophilic cold tolerant bacteria isolated from Tawang. Recently, the group has submitted 24 numbers of bacterial sequences from rhino dung of Kaziranga National Park, Assam to NCBI Gene Bank. The group also identified some prominent strains of protease producing *Pseudomonas aeruginosa* from soil of Kaziranga National Park of Assam. Another group of the institute is presently working on phylogenetic analysis of hydrocarbon degrading bacteria and their potential application in revegetation in crude oil contaminated soil. The Biotechnology Division of the Institute has also developed PGPR consortia and their technologies have already been transferred.

Institute of Bioresource and Sustainable Development (IBSD), Imphal, Manipur is presently engaged on characterization and diversity analysis of cyanobacterial microflora. The Energy Research Institute (TERI) has reported the isolation of a new strain of yeast (*Candida digboiensis* TERI ASN6) from crude oil contaminated soil of Digboi, Assam which was found to be most effective in hydrocarbon degradation.

### Conclusion and Future Prospects

Thus a number of modern techniques are currently available for studies on soil microbial communities. In this new decade successful assessment of bacterial diversity is possible with the consensus approach of utilizing the molecular genetic techniques along with the conventional methods. The ribosomal database, which now contains almost 10,000 complete and partial nucleotide sequences of collection cultures and clones is being constantly enlarged and in the near future will allow a more adequate

interpretation of 16S rRNA sequences of the total DNA extracted from environment. The metagenome analysis involving the cloning of the large fragments of soil bacterial DNA and express cloning may be the further progress in the study of viable. The application of these approaches not only limited to the analysis of 16S rRNA genes, but also structural genes, such as the nitrate reductase, ribulose 1,5-diphosphate carboxylase and methanol dehydrogenase genes, which will allow the relationship between the taxonomic structure of soil bacterial communities and their functioning to be studied. The continual and dynamical development of faster sequencing techniques, together with the advancement of methods to deal with the exponentially increasing amount of data generated, are expanding our capacity for the analysis of bacterial communities from an unlimited variety of habitats and environments. The synergism with the new emerging "omics" approaches is showing the path to functional metagenomics and to adopting integrative, wider viewpoints like systems biology. Another promising aspect suggests that, the clamping of the largest PCR product sequences with specific peptides, will allow establishment of rare and new phylogenetic group.

### Acknowledgments

The work is supported by ICAR-Network Project (AMAAS) funded by ICAR, Govt. India, New Delhi and the authors are thankful to the funding agency. □

### References

1. D.E. Dukhuizen, Santa Rosalia revisited: why are there so many species of bacteria? *Antonie van Leeuwenhoek*, **73**: 25–33 (1998).
2. J.D. Van Elsas, K. Smalla, Methods for sampling soil microbes. In: Hurst, C.J., Knudsen, G.R., McInerney, M.J., Stetzenbach, L. D., Walter, M.V. (Eds.), *Manual of Environmental Microbiology*. ASM Press, Washington, D.C, pp. 383–391 (1997).
3. J.T. Trevors., Bacterial biodiversity in soil with an emphasis on chemically-contaminated soils. *Water Air Soil Pollut.* **101**: 45–67 (1998b).
4. M. Blumer, Organic compounds in nature: limits of our knowledge. *Angew Chem Int Ed Eng.* **14**:507-514 (1975).
5. L.R. Lynd, M.S. Laser, D. Bransby, B. E. Dale, B. Davison, R. Hamilton, M. Himmel, M. Keller, J.D. McMillan, J. Sheehan, C.E. Wyman, How biotech can transform biofuels. *Nat. Biotechnol.***26**:169-172 (2008).
6. L. Hilterhaus, A. Liese, *Adv. Biochem. Engen. Biotechnol.* **105**:133-173(2007).
7. N. J. Clegg, B.M. Honda, I.P. Whitehead, T. A. Grigliatti, B. Wakimoto, H. W. Brock, V. K. Lloyd, D. A. R. Sinclair, Suppressors of position-effect variegation in *Drosophila melanogaster* affect expression of the heterochromatic gene light in the absence of a chromosome rearrangement. *Genome* **41**(4): 495–503 (1998).

8. J.M. Tiedje, S. Asuming-Brempong, K. Nusslein, T.L.Marsh, S.J. Flynn, Opening the black box of soil microbial diversity. *Appl. Soil Ecol.* **13**: 109– 122. (1999)
9. V. Gurtler, V. A. Stanisich, New approaches to typing and identification of bacteria using the 16S-23S-rDNA-spacer region. *Microbiology*,**142**: 3-16 (1996).
10. T. Vallaey, E. Topp, G. Muyzer, Evaluation of denaturing gradient gel electrophoresis in the detection of 16S rDNA sequence variation in rhizobia and methanotrophs. *FEMS Microbiol. Ecol.* **24**: 279–285 (1997).
11. J. Kling, Ultrafast DNA sequencing. *Nat. Biotechnol.* **21**: 1425–1427 (2003).
12. Z. Sarrate, F. Vidal, J. Blanco, Role of sperm fluorescent in situ hybridization studies in infertile patients: indications, study approach, and clinical relevance. *Fertil. Steril.* **93** (6):1892–902 (2010).
13. N. Myers, A. M. Russel, G. Cristina, A. B. Gustavo-Foneca, J. Kent, Biodiversity hotspots for conservation priorities. *Nature* **403**:853–858 (2000).
14. R. Saikia, D. K. Gogoi, S. Mazumdar, A. Yadav, R. K. Sarma, T. C. Bora, B. K. Gogoi, *Brevibacillus laterosporus* BPM3, a potential biocontrol agent isolated from a natural hot water spring of Assam, India. *Microbiol. Res.* **166** (3): 216-225(2010)
15. R. Saikia, R. K. Sarma, A. Yadav, T. C. Bora, Genetic and functional diversity among the antagonistic potential fluorescent pseudomonads isolated from tea rhizosphere. *Curr. Microbiol.* **62**(2):434-44 (2011).