ISSN: 0975-833X

Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 4, Issue, 02, pp.017-020, February- 2012

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

RESEARCH ARTICLE

METAGENOMICS: UNRAVELING THE OBSCURE

Naveen Kango[1]* and Swetlana Nagal[2]

[1]Department of Applied Microbiology and Biotechnology, Dr. Hari Singh Gour Vishwavidyalaya, Sagar (M.P.) 470003

[2]Department of Biotechnolgy (LFTS), Lovely Professional University, Phagwara, Punjab (India) - 144402

ARTICLE INFO

Article History:
Received 9th November, 2011
Received in revised form
8th December, 2011
Accepted 10th January, 2011
Published online 29th February, 2012

Key words:
Metagenomics,
Microorganisms,
Diversity,
Uncultivable,
Microbiome.

ABSTRACT

Metagenomics involves isolation, amplification, cloning and screening of DNA extracted from environmental samples. This direct approach that surmounts the need of cultivating microorganisms has opened up novel facets in biological research. In the absence of appropriate cultivation methods, more than 99% of microbial diversity awaits disclosure. Metagenomics transcends conventional genomics by a direct and comprehensive approach of analyzing a collective metagenome rather than individual genomes. Metagenomic analysis of human and insect gut, acid mine drainage, water bodies etc. have indicated possibility of finding novel organisms, bioactive molecules, pathways and has exposed the nature, role and dynamics of microbial communities in these niches.

INTRODUCTION

The very existence of microorganisms was debated before Antonie van Leeuwenhoek in 1673 startled the world by his drawings of the tiny and agile microorganisms. He could see these little animalcules through his crude microscope in almost every sample including that of pond water and scum of his teeth. A plethora of discoveries were to follow this landmark revelation. The presence of microorganisms, which were almost non-existent earlier, was suddenly being felt in every nook and corner of life. Later this was followed by the golden age of microbiology (1857-1914) and as of today we strongly believe that perhaps most path-breaking discoveries in terms of products and services will come through microbial biotechnology. Culture-independent techniques, such as 16S rRNA analysis, have revealed that the prokaryotic diversity is much more than what we are able to cultivate in laboratory. Microbiologists isolate, culture, classify and catalogue all cultivable forms using conventional culture methods. But there are many fastidious microbes having some obligatory requirement that perhaps we are not able to furnish. Woese and Fox in 1977 [1] reported that prokaryotic diversity can be studied using phylogenetic analysis of the short and conserved 16S ribosomal RNA (16S rRNA) molecules. Such analysis of environmental samples revealed that there are many 16S rRNA sequences that do not match with known or reported cultivated species. This fact clearly indicated that there is a large diversity that remains obscure in the absence of proper culturing methods. More than 99% of bacteria from natural environments are uncultivable or can not be cultured *in vitro* using routine techniques. Since most of the recorded diversity of bacteria is that of cultivable ones, the uncultivable bacteria represent new divisions of families of bacteria. A rough estimate of number and types of such bacteria suggests that there is a huge gene pool that remains unexplored. Culture independent methods of metagenome reconstruction and its functional or sequence based screening have made it possible to understand the significance of these obscure forms of life. The prokaryotes include a wide range of forms distributed over two major domains, namely, eubacteria and archaea. These ubiquitous forms have been reported to occur in almost all kinds of environments and ecological niches including the most hostile ones like thermal vents, glaciers, earth's crust and poles. The term 'Metagenomics' was introduced by Handelsman *et al*. [2] to refer studies of a collection of genes isolated from environmental samples and were comparable to single genome. Metagenomics can be defined as the application of modern genomics techniques to the study of communities of microbial organisms in their natural environments, bypassing the need for isolation and cultivation. A large number of explorations resulted into startling revelations that a plethora of hitherto unknown organisms exist in the so called well characterized samples such as garden soils and dental plaques. Total prokaryotic cells present on earth have been estimated to be 4-6 x 10^{30} and more than 99% of these remain uncultured and uncharacterized. This bewildering and yet unexplored biodiversity presents an exciting challenge for finding out novel genera, species, genes, products and pathways. Studies on microbial genomes (microbiomes) from ecologically diverse locations such as hot water springs, garden soils, and insect and human guts have revealed an overabundance of valuable information regarding types of organisms and their interaction. Metagenomics

Corresponding author: nkango@gmail.com

developed as an effort to unravel the hidden treasure of natural environments such as soil, marine water, and gastrointestinal tract of invertebrates. This field involves direct isolation of DNA from one habitat and subsequent cloning of this metagenome in surrogate host. Cloning is followed by either sequence based or function driven analysis. The sequence based approach involves screening of clones for highly conserved 16S rRNA genes for taxonomic and identification purposes. Entire genome is also sequenced in order to fish out phylogenetic anchors. Function based analysis, on the other hand, involves screening of metagenomic libraries for identifying some important phenotypic characters that may be of biotechnological significance (Figure 1). Genes coding for specific characters (e.g. salt tolerance, thermostability) and bioactive molecules of importance and value (e.g. antibiotics, enzymes) can be discovered and their origin can be traced out in the metagenome. Obviously, this will lead to addition of several new genera and species to different domains. Forty such divisions have already been identified. Out of these, thirteen divisions are altogether new and are not represented by any cultured member.

Screening Strategies

After extraction and amplification of DNA it is imperative to screen metagenomic library. This involves complete sequencing of phylogenetic anchors such as 16s rRNA. Sequence based screening allows us to screen the library without any compulsion of expression of foreign genes in the heterologous host. Function based screening on the other hand involves expression of a particular function of the isolated metagenomic DNA in the surrogate host. This helps in identification of novel genes with known functions. Such analyses have lead to the identification of genes expressing novel antibiotics, ion transporters, hyrdrolase enzymes etc. However, many a times the results of functional screening are misleading due to inefficient expression of the gene and improper secretion of the protein in heterologous host. Efficient screening systems such as metabolite related expression (MRE), where the cell containing metagenomic DNA has an intracellular sensor for gene product and substrate induced gene expression (SIGEX), where the intracellular substrate induces expression of gene in cloned metagenomic DNA have been developed for comprehensive screening of the cloned metagenome. Among several success stories, metagenomics has been used for obtaining novel hydrolases and other biocatalysts (e.g. proteases, lipases, amylases cellulases, agarases and decarboxylase, alkane hydroxylase etc.). Novel antibiotics such as terragine, violacein and palmitoylputrescine have been produced in heterologous hosts. Metagenomic approach to environmental samples has also revealed novel pathways that play significant role in biogeochemical cycles. For example, genes encoding ammonia monooxygenase, a major enzyme of nitrifying or ammonia oxidizing bacteria (AOB), was identified on a metagenomic soil clone of archaea (crenarchaeota). Xenobiotic degrading bacteria have been identified from samples from black sea using similar approaches. The following section describes the results of metagenomic screening of four habitats in detail.

(1) Human Gut and Skin Microbiome

The human intestinal microbiota is composed of 10^{13} to 10^{14} microorganisms whose collective genome ("microbiome") contains at least 100 times as many genes as human genome. The human gut microbiota represents an extremely complex microbial community, and its microbiome encodes functions that are believed to have a significant impact on human physiology. Gill *et al*. [3] analyzed about 78 million base pairs of unique DNA sequences and 2062 PCR amplified 16S ribosomal DNA sequences obtained from the fecal DNA of two healthy adults. Interestingly, 13 divisions of the primitive archaea are also reported among 70 divisions of bacteria found in the human gut but their role is yet to be determined. Metabolic function analysis of identified genes has confirmed that microbial genome plays important role in metabolism of glycans, amino acids, xenobiotics, methanogenesis, biosynthesis of vitamins and isoprenoids. Co-colonization of the gut commensals *Bifidobacterium longum* and *Bacteroides thetaiotaomicron* in a murine model system revealed that the presence of bifidobacteria induces an expansion in the types of polysaccharides targeted for degradation by Bacteroides and also activates host genes involved in innate immunity. In addition, comparative analysis of individual human gut microbiomes also has revealed various approaches that the microbiota adapts itself to the intestinal environment. Thus, it can be envisaged that humans are super organisms whose metabolism represents an amalgamation of microbial and human attributes.

In another study, researchers have generated a diversity profile of human skin microbiota by sequencing 16S rRNA on the inner elbow to study the extremely common inflammatory skin disorder, atopic dermatitis, which affects this area of the skin and is associated with *Staphylococcus* infections. They found that Proteobacteria (including *Pseudomonas* and *Janthinobacterium*) predominated in all the healthy subjects tested. They also found that the microbiota of mouse skin is comparable to that of the human inner elbow, suggesting a potential model for human skin disorders related to microbiota.

(2) Marine Microbiome (Sargasso Sea)

Direct DNA sequencing of entire microbial population of Sargasso sea, a nutrient limited region of Atlantic ocean near Bermuda was carried out by Venter *et al*. [4]. This survey yielded more than 1.6 billion base pair of DNA data from more than 1800 bacterial species that includes 148 novel phylotypes. Different phylogenetic biomarker genes were used to elucidate the composition of the bacterial community on surface waters of the Sargasso Sea, resulting in the identification of nine major bacterial phyla (Proteobacteria, Actinobacteria, Cyanobacteria, Firmicutes, Bacteroidetes, Chloroflexi, Spirochaetes, Fusobacteria and Deinococcus-Thermus) and two archaeal phyla (Crenarchaeota and Euryarchaeota). The vast majority of all these sequences clustered within the α- and γ-divisions of the phylum Proteobacteria. The Sargasso Sea surface water samples were also found to harbor the fastest-evolving, smallest genomes with the lowest GC content. This study also revealed 782 new proteorhodopsin genes (Proteorhodopsin are light driven proton pumps for generation of ATP in phototrophs) with entirely different spectral properties. This increases the number of known proteorhodopsin genes by almost 10-fold. These data also revealed presence of monooxygenase gene (for oxidation of ammonia to nitrite) in archaea associated forms that were formerly supposed to be exclusively present in

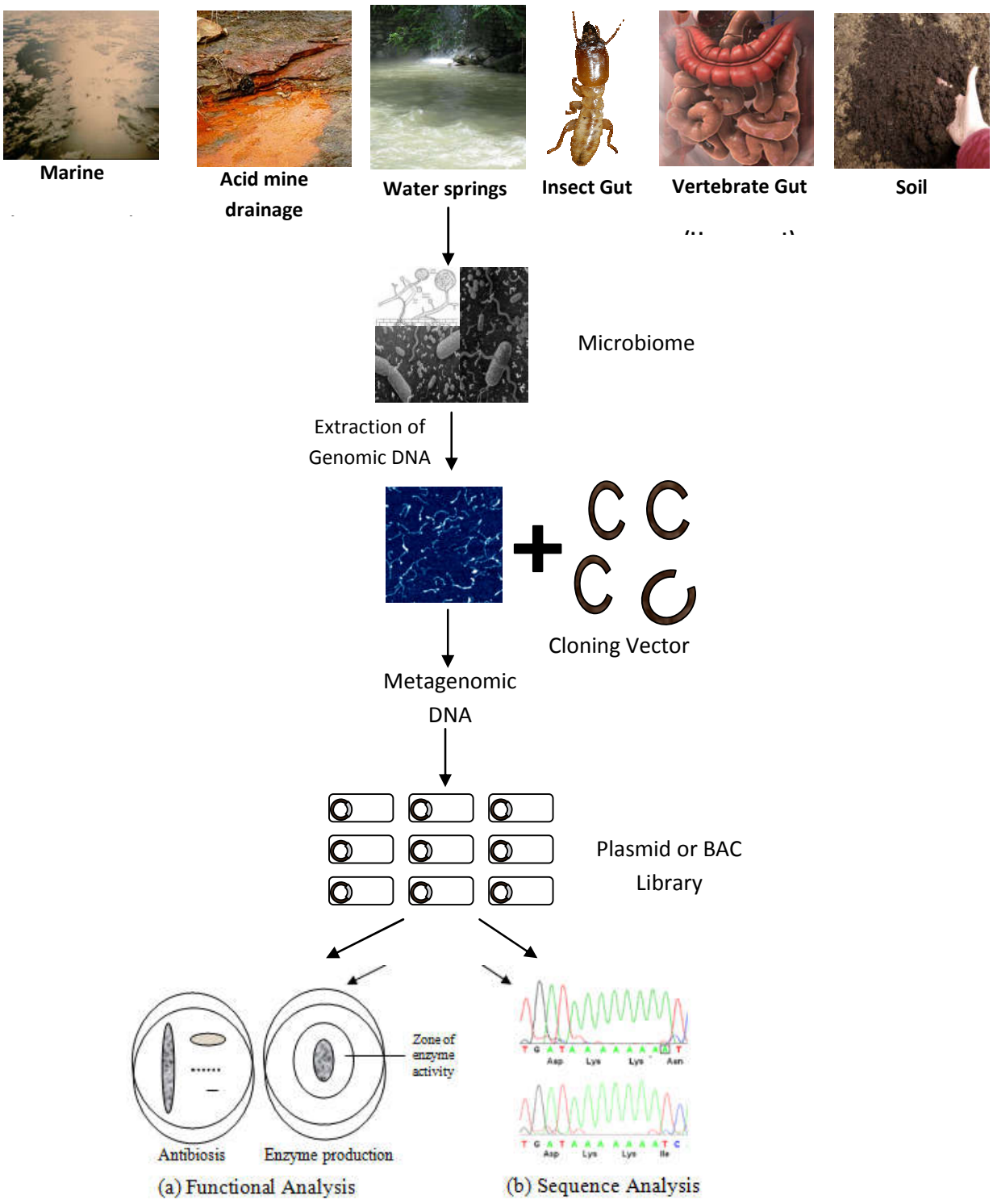

Figure 1 Making and Screening of metagenomic libraries of environmental samples. (a) Specific bioactive molecules (e.g. antibiotics, enzymes) can be identified using functional analysis. (b) Sequence analysis is used for comparison of metagenomic DNA with the known sequences to establish identity.

eubacterial population. Another important observation was presence of gene similar to those involved in phosphonate uptake or utilization of polyphosphates and pyrophosphates, which are present in this extremely phosphate limited Sargasso biome. This collection of genomes and the function based analysis will provide a new route for discovery of the mechanisms of phosphorus acquisition and transformation and better understanding of nutrient cycling.

(3) Acid Mine Drainage Microbiome

The hydrological, microbiological, and geological conditions promote rapid oxidation of large amounts of pyrite ores in mines which results in the formation of extremely acidic, warm, metal-rich solutions, referred to as acid mine drainage (AMD). The bacteria and archaea living in the mine form a several millimeters thick pink biofilm that floats on the

surface of the hot acidic water. These communities have been extensively characterized using culture-independent 16S rRNA gene surveys, sequencing of random shotgun genomic libraries and proteomics. The microbial biofilm community from acid mine drainage in Richmond Iron mountain mine, Northern California was found dominated by *Leptospirillum* and *Sulfobacillus*, although presence of *Ferroplasma* and *Thermoplasmataleare* was also reported. Thorough sequencing and metabolic analysis also provided the starting point for a 'proteogenomic' analysis where tryptic digests of proteins extracted from AMD biofilms were applied to shotgun mass spectrometry. The most significant outcome of this analysis was the identification sequences of a novel acid stable iron-oxidizing c-type cytochrome (Cyt 579) from the *Leptospirillum* group II. Low but distinct eukaryotic community based on 18S rRNA gene library in AMD at the Iron Mountain has also been noticed [5]. The diversity was reported to be limited to ascomycetous fungi (68% sequence) belonging to the *Dothideomycetes*, the *Eurotiomycetes* and protists falling into the deeply branching lineage named the acidophilic protist clade (APC) and *Heterolobosea*. The APC group represents kingdom-level novelty, with less than 76% sequence similarity to 18S rRNA gene sequences of organisms from other environments. The eukaryotic lineages (except APC) were found closely related to neutrophiles, suggesting their recent adaptation to this extreme environment. Protists play a predatory role (engulf acidophilic bacteria) and influence the AMD community members. α-proteobacterial endosymbionts have also been discovered in some protists. On the basis of these studies it was concluded that eukaryotes play a large role in the cycling of carbon within AMD communities [5].

(4) Termite Gut Microbiome

Termite gut carries a complex microbial community within their hind guts which provide them the capability to digest wood. They also help to compensate scarcity of nutrients in woods. Warneck and coworkers [6] sampled luminal component of hindgut segment of 165 adult worker termites and generated about 70 Mbp of shotgun sequences for taxanomic and phylogenetic study. They discovered 216 novel and distinct phylotype of bacteria. They also reported genes for large number of glycoside hydrolases; however, no gene sequences for lignin degrading enzymes were obtained from hind gut suggesting that these may be located in other segment of termite gut as the process requires molecular oxygen. They also carried out a quantitative analysis of functional genes in the termite gut samples and reported a number of glycoside hydrolases (e.g. cellulases, xylanases etc.). They detected 45 distinct groups and predicted *Treponema* (a genus of Spirochaetes) as the most likely source for the majority of these enzymes. In addition, a number of gene families related with such hydrolase domains were found, including carbohydrate-binding domains and other functional domains. Analysis of a particular subset of the proteome (the secreted extracellular proteins) in luminal fluid using mass spectrometry confirmed, for the first time, that bacterial glycosidases are indeed produced in the termite gut. More than 40 of the glycosidase genes were individually cloned, expressed heterologously and tested on acid-solubilized and microcrystalline cellulose to confirm that termite guts harbor functional glycosidase enzymes.

Conclusion

Metagenomics has changed the way the microbial world is viewed and studied by allowing direct investigation of microbes irrespective of their culturability and taxonomic identities. This new field of biology is making important contributions in many areas including ecology, biodiversity, bioremediation, and bioprospective of products in medicines. Inefficiencies in sampling, contamination of DNA, limitation of ideal phylogenetic anchor and inadequacies in analysis of huge data flooding in from recent worldwide initiatives are few challenges to be dealt with. Refinement of methods and interdisciplinary coordination of experts from different fields is required in this regard. Ultimately, integrated analysis of metagenomes, metatranscriptomes, metaproteomes and metabolomes will be needed to study biological networking across multiple hierarchical level *in situ* which will help in future in understanding of any global problem or situation such as climate change, emergence of antibiotic resistance and aspects of bioremediation.

REFERENCES

1. Woese, C. and Fox, G. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci USA* 74 (11): 5088–90.
2. Handelsman, J., Rondon, M. R., Brady, S.F., Clardy, J. and Goodman, R.M.(1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 5: R245-249.
3. Gill S.R, Pop M, Deboy R.T, Eckburg P.B, Turnbaugh P.J, Samuel B.S., *et al*. (2006). Metagenomic analysis of the human distal gut microbiome. *Science*312: 1355–1359.
4. Venter, J.C.K. Remington, J.F. Heidelberg, A.L. Halpern, D. Rusch,J.A. *et al*. (2004). Environmental genome shortgun sequencing of the Sargasso sea. *Science* 304: 66-74.
5. Baker, B.J. and Banfield, J.F. (2003). Microbial communities in acid mine drainage. *FEMS Microbiol. Ecol.* 44:139-152.
6. Warnecke, F. *et al*. (2007). Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 450:560–565.

****