

Metatranscriptomic study of microbes in environment: a community function based approach

Abstract

Research on microbial ecology in recent years had shifted methodologically from traditional culture based techniques to culture independent molecular techniques. This is due to the fact that it has already been established that less than 1% of microbes in environmental samples are cultivable. In recent years, metagenomics research has developed into analyzing partial or complete genomes as well as other types of molecules in environmental samples, including RNA (transcripts), proteins (translated transcripts), and metabolites (metabolic intermediates and final products). While the analyses of DNA sequences in environmental samples can provide evidence about the diversity and function of specific groups of organisms in the analyzed ecological niches, the relative importance of those genes in environmental functions cannot be determined from DNA sequences alone. To understand which genes are expressed under a specific ecological condition, the analysis of transcripts from the communities is often needed.

Keywords: soil, RNA, mRNA, meta transcriptomics, soil fertility

Volume 4 Issue 1 - 2017

Subhankar Barua, Anushree Hait, Shilajit Barua

Department of Microbiology, Vijaygarh Jyotish Ray College, India

Correspondence: Shilajit Barua, Department of Microbiology, Vijaygarh Jyotish Ray College, 8/2 Bejoygarh, Jadavpur, Kolkata, India, Tel +919432166493, Email shilajit_26@yahoo.co.in

Received: December 04, 2016 | **Published:** January 11, 2017

Introduction

Since the advent of pure culture technique innovated by Robert Koch in 1882, microbiologists were attracted towards the robust and selective study of microorganisms in pure culture. Most of the knowledge that fills modern microbiology textbooks is derived from the organisms maintained in pure culture. Much of the knowledge in microbial physiology and genetics in the last century were largely based upon readily cultured model organisms. The fact that microbes and their products had been found to be useful for mankind attracted industrialists to set up several microbes based industries throughout the world. Microbes had been isolated from diverse ecological niche/climatic zone and various culture media were successfully developed for selective isolation, maintenance and overproduction of desired product. However, due to inappropriate simulation of growth conditions like, lack of necessary symbionts, nutrients, surfaces, excess inhibitory compounds, incorrect combinations of temperature, pressure, or atmospheric gas composition, accumulation of toxic waste products from their own metabolism, etc, majority of the community members were unable to grow under laboratory conditions and were referred to as unculturable microbes. Hence several culture independent studies were undertaken in the late twentieth century. The most common culture independent studies were nucleic-acid based, which includes analyses of whole genome or selected genes like 16S rRNA. Different culture-independent techniques have been developed after Pace et al. (1985) who proposed direct cloning of environmental DNA. These techniques can be used both for partial and whole community analysis. For the partial community analysis, PCR-based methods are used where the total DNA/RNA extracted from the environmental sample is used as a template for the characterization of the microorganisms. The whole community analysis involves non-PCR based techniques like hybridization, G+C content estimation, whole genome sequencing etc.

Microbes strongly interacting with each other in a microenvironment comprise a local community.¹ This complex interaction among

themselves and with their abiotic environment is key to several significantly important activities like biogeochemical cycling of elements² and cycling of organic compounds. This influences above-ground ecosystems by contributing to plant nutrition,³ plant health,⁴ soil structure⁵ and soil fertility.⁶

The fact that less than 1% of the total microbes in the environment are cultivable lead to the development of culture independent molecular study commonly referred to as metagenomics. As defined by Schloss and Handelsman in 2003, metagenomics refers to meta-analysis of genetic composition of microbial community in environmental samples. There are other terms used for the same approach as community genomics, environmental genomics, and population genomics. In the last decade of the twentieth century and the beginning of twenty first century, microbial ecologists, shifted their research towards culture independent molecular techniques.

For molecular based analysis, genomic and transcriptomic materials are required to be isolated by *in situ* lysis of microbes followed by extraction. It suffers from several challenges like difficulty in separating DNA/RNA released from the cell from soil matrix⁷ and co-extraction of humic contaminants that might inhibit downstream processing of DNA/RNA like Taq polymerase/reverse transcriptase catalyzed PCR amplification, restriction digestion or cloning.^{8,9}

Robust research work has been carried out in the field of metagenomics in last two decades, that has changed our knowledge base and outlook to a more community based study of microbial ecology. Studies carried out on diverse ecosystem from the extreme environments to agricultural soils, from wastewater to compost had enlightened us about the importance of microbial interaction at community level. Effect of biotic and abiotic factors on the community structure hence function has also been extensively elucidated.

For last few years, microbial ecological research has shifted towards molecular elucidation of functional characterization of microbial communities based on RNA analysis. Metagenomic DNA

based studies provided information about the genotypic composition of the microbial community, intern depicts the functional potentiality of a community. But actual soil functioning depends on the expression of those genes. Mere presence of a gene in a community doesn't implies its expression since it is majorly governed by several biotic and abiotic factors. Overall activity of a specific soil population should not be regarded as an indication of a specific function. But, overall activity of a specific soil population should not be regarded as an indication of a specific function. Analysis of mRNA transcripts could be one of the approach for determining specific microbial activities in soil. Due to the short half-life of mRNA transcripts, regulation of their synthesis can subsequently affect synthesis of specific protein and the function it carries out. Therefore, analysis of gene expression would help us in investigating specific microbial process in real time in situ.

Meta transcriptomics involves high-throughput detection and analysis of sequence diversity and associated functions of the transcripts (RNA molecules) extracted from samples with heterogenous microbial genotypic composition. Transcriptomic profiling of samples containing such multiple cell types, species or operational taxonomic units (OTUs), complements the metagenomic approach by focusing on the expressed subset of genes (meta transcriptome), thus reducing the complexity of the data to be analysed.¹⁰

Meta transcriptomics involves study of all the RNA transcripts (rRNA, mRNA, miRNA, and tRNA) produced by the biota in a sampled soil. Microbial response to soil management can be studied by analyzing RNA which reflects portion of the soil microbial community that is active at the time of sampling.¹¹ RNA based soil community analysis account way back to the early nineties.^{12,13} Aneja et al.,¹⁴ carried out random arbitrarily primed PCR (RAP-PCR) on the meta transcriptome of plant litter degrading microbial community and differential expression was successfully observed at various stages of degradation. Active populations of both eukaryotes and prokaryotes had been successfully identified by meta transcriptome based reverse transcription and sequencing of rRNA (RNA-Seq).¹⁵ Meta transcriptomics based study allows detection of sequences associated with a particular environmental condition that may not be so readily identified in metagenomic studies and increases the chance of detecting ecologically relevant active functions. The discovery of functions being induced in a sample as a response to a certain environmental condition (exerted pressure) also gives insight into processes of adaptation and enriches our understanding of communities.¹⁰ Since majority of bacteria exhibit transcriptional gene control that enables rapid adaption to altered environmental conditions, meta transcriptome based approaches would be useful in deciphering immediate microbial gene regulatory responses to changed environment at the time of sampling.¹⁶ Composition of microbial communities as determined by RNA based studies differed from that obtained by DNA based studies, indicating that the active microbial community is just a subset of the potentially active microbial community.^{17,18} Similar observations were also observed in metagenomes and meta transcriptomes of marine or soil environments.^{19,20}

Technically, meta transcriptomic approach of studying microbial ecology involves extraction, reverse transcription, amplification, and sequencing of prokaryotic mRNA transcripts, which was once thought to be impossible. A major drawback of carrying out meta transcriptomic based study is the inherent instability of RNA molecules that have short cellular half life (seconds to minutes) when compared

to more stable DNA (used in metagenomics) or proteins (used in meta proteomics).¹⁰ Also, very low quantitative recovery of transcripts from environmental samples and co-extraction of humic inhibitors are critical methodological considerations required to be addressed for meta transcriptomic studies.

However, due to extensive research on meta transcripts and technological advancements, it has now been possible to successfully isolate RNA from ecological samples.^{21,22} Several commercial kits have been made available for extracting RNA from diverse ecological samples as summarized by Wang et al.¹⁵ Of these, MoBio RNA Power Soil Total RNA Isolation Kit has been widely used.²³⁻²⁵ Functional mRNA accounts to only 5% of the total RNA extracted from an ecological sample.²⁶ Hence analysis of the functional activity of a gene in a community requires to enrich mRNA from the total extracted RNA samples. Presence of 3' Poly(A) tail in eukaryotic mRNA favors direct cDNA preparation by using Oligo(T) primers. Prokaryotic mRNA lacking such Poly(A) tail requires alternative approach. He et al.,²⁶ compared two methods of removing rRNA, i) rRNA-specific exonuclease treatment and ii) rRNA hybridization capture-based protocols for removal of rRNA as an approach of enriching mRNA prior to cDNA preparation. Comparison of two commercial kits, based on i) subtractive hybridization of rRNA (Ambion MICROB Express bacterial mRNA enrichment kit) and ii) rRNA specific exonuclease (Epicenter mRNA-ONLY prokaryotic mRNA isolation kit) were also compared.²² Both of them observed and concluded that rRNA removal by subtractive hybridization yielded a better result with respect to degree of rRNA removal and quantitative mRNA yield.

Both PCR and non-PCR based approaches has been used to describe in-situ transcription profiling of natural microbial communities. PCR based methods are random primer based (RAP-PCR) or gene specific RT-qPCR²⁷ both of which are performed on cDNA obtained from transcripts by reverse transcription. However, using this approach, only a small number of genes can be targeted at a time, and it difficult to design primers that bind only to orthologs of a given gene because, vast nucleotide sequence diversity is observed for functional genes in nature.¹⁶ Environmental microarray can be used to overcome gene number constraint by measuring expression levels of hundreds to thousands of genes at a time, but it also suffers from several methodological constraints like difficulty in designing of microarray probes to address the full diversity of ortholog sequences encountered in nature and difficulties in obtaining sufficient mRNA for replicated microarray studies with environmental samples.²⁸

Primer or probe independent meta transcriptomic approaches of studying community involves random sequencing of microbial community mRNA. In this approach, community gene expression in an environmental sample can be monitored as a function of differential biotic or abiotic parameter which is useful in understanding the timing and regulation of complex microbial processes within communities and consortia, as well as microbial adaptability in response to changing conditions. Direct sequencing based individual meta transcriptomic studies contribute unbiasedly to a growing community database unlike RT-qPCR or microarray, that are influenced by parameters such as array composition, primer design and hybridization conditions. Overcoming the low-throughput (<400 sequences), low coverage, and biasness of priming methods encountered during the early stages of meta transcriptomic studies, coupling of high-throughput sequencing technologies like 454 pyro²⁹ or Illumina platform based

next generation sequencing technologies contributed to a growing community database that can be used to address intractable or unanticipated questions.

Real time qPCR data were either analyzed by absolute quantization,³⁰ which involves study of community transcripts when compared to known quantity of mRNA, or by relative quantization,³¹ wherein a crossing threshold (CT) is established at the earliest cycle possible, which is subsequently used to generate the relative expression level. Functional gene array (FGA), that contains probes from genes with known biological functions, efficiently links microbial community composition to ecosystem functions. For example, FGA termed GeoChip contains >24,000 probes from all known metabolic genes involved in various biogeochemical, ecological, and environmental processes such as ammonia oxidation, methane oxidation, and nitrogen fixation.³² Meta transcriptomic data processing firstly involves removal of short or poor quality sequences and error (like exclusion of rRNA sequences by comparison to available databases and removal of artificial poly (A) reads), correction.^{19,33} Next, the metagenomic sequences are compared to NCBI databases and the relative frequencies of genes could be mapped by comparing reads with known sequences using the Burrow-Wheeler Aligner (BWA, <http://bio-bwa.sourceforge.net/>) or the Blast-Like Alignment Tool.³⁴ Assemblers like Velvet,³⁵ Newbler³⁶ and Genovo.³⁷ Kyoto Encyclopedia of Genes and Genomes,³⁸ the Clusters of Orthologous Groups (COGs),³⁹ and the evolutionary genealogy of genes: Non-supervised Orthologous Groups (egg NOG);⁴⁰ databases could be used for functional categorization of transcripts by comparing with groups of genes that have been assigned to different functional pathways (e.g. denitrification or nitrogen fixation) based on the similarity of protein orthologs from sequenced isolate microbial genomes.

Meta transcriptomic approach is now widely used to study microbial nutrient cycling in nature. In a recent study on deep sea at Gulf of California meta transcriptome assembly of carbon and nitrogen cycling microbial community revealed transcriptionally active populations of diverse and novel microbial communities at high resolution (to the strain level). It also enabled the reconstruction and functional characterization of transcripts that would have otherwise been overlooked by mapping to reference genomic databases.⁴¹ Differential expression of 81 different genes at various stages of plant growth (seedling, vegetative stage, bolting and flowering) in their rhizospheric community was observed by meta transcriptomic analysis.⁴² Žifčáková et al.,⁴³ observed change in soil community transcription profile of fungi between seasons and observed a decrease in fungal contribution from 33% in summer to 16% in winter in total community transcription. Meta transcriptomic studies on paddy field soils revealed the complex and functionally co-ordinated assemblage of microbiomes as a function of the nature of habitat and hierarchical importance for community succession.²⁹

Hence it can be concluded that meta transcriptomics would be a new approach to study complex functioning of microbial community in ecosystems and their response to change in ecological and seasonal parameters.

Acknowledgements

The authors are in debt to Prof. Debasish Chatterjee, Department of Chemistry, University of Kalyani, India for his indepth knowledge sharing related to this work. Financial support by SERB, Govt. of India in carrying out allied research is also deeply acknowledged.

Infrastructural support to the author is provided by host institution, Vijaygarh Jyotish Ray College needs special mention.

Conflict of interest

The author declares no conflict of interest.

References

- Konopka A. What is microbial community ecology? *The ISME journal*. 2009;3(11):1223–1230.
- Wall DH, Virginia RA. Controls on soil biodiversity: insights from extreme environments. *Applied Soil Ecology*. 1999;13(2):137–150.
- Timonen S, Finlay RD, Olsson S, et al. Dynamics of phosphorus translocation in intact ectomycorrhizal systems: non-destructive monitoring using a \square -scanner. *FEMS Microbiology Ecology*. 1996;19(3):171–180.
- Smith KP, Goodman RM. Host variation for interactions with beneficial plant-associated microbes. *Annu Rev Phytopathol*. 1999;37(1):473–491.
- Dodd GD, Soulen MC, Kane RA, et al. Minimally invasive treatment of malignant hepatic tumors: at the threshold of a major breakthrough. *Radiographics*. 2000;20(1):9–27.
- O'Donnell DE, Revill SM, Webb KA. Dynamic hyperinflation and exercise intolerance in chronic obstructive pulmonary disease. *AJRCCM*. 2001;164(5):770–777.
- Feinstein LM, Sul WJ, Blackwood CB. Assessment of bias associated with incomplete extraction of microbial DNA from soil. *Appl Environ Microbiol*. 2009;75(16):5428–5433.
- Tebbe CC, Vahjen W. Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and a yeast. *Appl Environ Microbiol*. 1993;59(8):2657–2665.
- Sharma S, Aneja MK, Schloter M. Functional characterization of soil microbial communities by messenger RNA Analysis. In: Van Elsas JD, Trevors JT, editors. *Modern soil microbiology*. USA: CRC Press; 2006. 342 p.
- Cabellos Ruiz A, Janca H, Baena S, et al. Beyond metagenomics: integration of complementary approaches for the study of microbial communities. *Metagenomics: theory, methods and applications*. UK: Caister Academic Press; 2010 p. 15–37.
- Blagodatskaya E, Kuzyakov Y. Active microorganisms in soil: critical review of estimation criteria and approaches. *Soil Biology and Biochemistry*. 2013;67:192–211.
- Hahn D, Kester R, Starrenburg MJ, et al. Extraction of ribosomal RNA from soil for detection of *Frankia* with oligonucleotide probes. *Arch Microbiol*. 1990;154(4):329–335.
- Selenska S, Klingmüller W. Direct recovery and molecular analysis of DNA and RNA from soil. *Microbial releases*. 1992;1(1):41–46.
- Aneja MK, Sharma S, Munch JC, et al. RNA finger printing—a new method to screen for differences in plant litter degrading microbial communities. *J Microbiol Methods*. 2004;59(2):223–231.
- Wang Y, Hayatsu M, Fujii T. Extraction of bacterial RNA from soil: challenges and solutions. *Microbes Environ*. 2012;27(2):111–121.
- Moran MA. Meta transcriptomics: eavesdropping on complex microbial communities—large-scale sequencing of mRNAs retrieved from natural communities provides insights into microbial activities and how they are regulated. *Microbe*. 2009;4(7):329.
- Anderson IC, Parkin PI. Detection of active soil fungi by RT-PCR amplification of precursor rRNA molecules. *J Microbiol Methods*. 2007;68(2):248–253.

18. Baldrian P, Kolařík M, Stursová M, et al. Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. *ISME J.* 2012;6(2):248–258.
19. Frias Lopez J, Shi Y, Tyson GW, et al. Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci USA.* 2008;105(10):3805–3810.
20. Urich T, Lanzén A, Qi J, et al. Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. *PLoS One.* 2008;3(6):e2527.
21. McGrath KC, Thomas-Hall SR, Cheng CT, et al. Isolation and analysis of mRNA from environmental microbial communities. *J Microbiol Methods.* 2008;75(2):172–176.
22. Mettel C, Kim Y, Shrestha PM, et al. Extraction of mRNA from soil. *Applied and environmental microbiology.* 2010;76(17):5995–6000.
23. Yergeau E, Sanschagrin S, Maynard C, et al. Microbial expression profiles in the rhizosphere of willows depend on soil contamination. *ISME J.* 2014;8(2):344–358.
24. Wei W, Isobe K, Nishizawa T, et al. Higher diversity and abundance of denitrifying microorganisms in environments than considered previously. *ISME J.* 2015;9(9):1954–1965.
25. Rosselli R, Romoli O, Vitulo N, et al. Direct 16S rRNA-seq from bacterial communities: a PCR-independent approach to simultaneously assess microbial diversity and functional activity potential of each taxon. *Scientific Reports* 6. 2016.
26. He S, Wurtzel O, Singh K, et al. Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. *Nature methods.* 2010;7(10):807–812.
27. Steunou AS, Jensen S, Brecht E, et al. Regulation of nif gene expression and the energetics of N₂ fixation over the diel cycle in a hot spring microbial mat. *ISME J.* 2008;2(4):364–378.
28. Zhou J, Thompson DK. Challenges in applying microarrays to environmental studies. *Curr Opin Biotechnol.* 2002;13(3):204–207.
29. Kim Y, Liesack W. Differential assemblage of functional units in paddy soil microbiomes. *PLoS one.* 2015;10(4):e0122221.
30. Bustin SA, Benes V, Nolan T, et al. Quantitative real-time RT-PCR—a perspective. *Journal of molecular endocrinology.* 2005;34(3):597–601.
31. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods.* 2001;25(4):402–408.
32. He Z, Gentry TJ, Schadt CW, et al. GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J.* 2007;1(1):67–77.
33. Stewart FJ, Ottesen EA, DeLong EF. Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME J.* 2010;4(7):896–907.
34. Kent WJ. BLAT—the BLAST-like alignment tool. *Genome research.* 2002;12(4):656–664.
35. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome research.* 2008;18(5):821–829.
36. Chaisson MJ, Pevzner PA. Short read fragment assembly of bacterial genomes. *Genome research.* 2008;18(2):324–330.
37. Laserson J, Jovic V, Koller D. *Genovo: de novo assembly for metagenomes.* Germany: Springer Berlin Heidelberg; 2010. p. 341–356.
38. Kanehisa M, Goto S, Kawashima S, et al. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 2004;32(suppl 1):D277–D280.
39. Tatusov RL, Fedorova ND, Jackson JD, et al. The COG database: an updated version includes eukaryotes. *BMC bioinformatics.* 2003;4(1):41.
40. Bang Jensen J, Gutin GZ. *Digraphs: theory, algorithms and applications.* Germany: Springer Science & Business Media; 2008. 772 p.
41. Baker BJ, Sheik CS, Taylor CA, et al. Community transcriptomic assembly reveals microbes that contribute to deep-sea carbon and nitrogen cycling. *ISME J.* 2013;7(10):1962–1973.
42. Chaparro JM, Badri DV, Vivanco JM. Rhizosphere microbiome assemblage is affected by plant development. *ISME J.* 2014;8(4):790–803.
43. Žifčáková L, Větrovský T, Howe A, et al. Microbial activity in forest soil reflects the changes in ecosystem properties between summer and winter. *Environ Microbiol.* 2016;18(1):288–301.