

## System Level Metagenomics and Metatranscriptomics in Quantitative Site Assessment and Bioremediation

**Gary S Sayler** (Sayler@utk.edu) (University of Tennessee, Knoxville, TN, USA)

**Background/Objectives.** Predictable and successful bioremediation is improved by a quantitative knowledge of the biodegrading microbial community, its activity and dynamics within the context of the geohydrology of the contaminated site and available engineering interventions of the remediation process. For over two decades it has been possible to effectively quantify these community dynamics from the perspective of a community genome. This is accomplished by directly extracting the nucleic acids (NA), both DNA and messenger RNA (mRNA) (the transcribed product of the expressed genes), directly from contaminated site samples. This allowing monitoring the genes and their activity in mediating the bioremediation process as well as discovery of new genes and biodegradative mechanisms. However the practical development and application of this knowledge in high-resolution bioremediation assessment has been limited to a relatively small set of genes for critical sub populations of the community. Recent technology advances are revolutionizing the ability to systematically reduce this ability to standard practice and rapidly advance our process understanding.

**Approach/Activities.** The analysis of the community genome at the level of DNA and RNA has become known as metagenomics and metatranscriptomics respectively, and seeks to supplement and overcome many limitations of community analysis associated with lab isolation and cultivation of environmental microorganisms. In a half decade the vast improvements and cost reductions in high through put (HTP) NA sequencing, quantitative PCR (polymerase chain reaction) technology, gene chip arrays, and bioinformatics software is offering an unprecedented ability to quantify the vast diversity of microbial populations existing in contaminated site matrices. In addition, the array of genes representing the complete genetic diversity of a site and whether those specific genes that are functionally active (expressed) in bioremediation can be accomplished. An analogy is the human liver in which all tissue cells contain the complete human genome, yet only a small set of those genes are actively involved in normal detoxification process and an even smaller set responds to specific toxic agents such as PCB or dioxin. The precise situation exists for the community genome of contaminated soil. Critical omic knowledge derived from site assessment and monitoring along with historic lab information is providing new insight into underlying mechanisms of bioremediation success or failure, and ultimately designs and operating strategies to achieve optimal and predictive bioremediation.

**Results/Lessons Learned.** Replicated NA extraction can be achieved for virtually all sample types and is the hi fidelity representation of the community genome suitable for HTP sequencing, bioinformatics sequences assembly to genes, and gene annotation to function and discovery, and activity. These analyses can be efficiently and systematically employed to establish base case in site assessment, need if any for bioaugmentation and stimulation of syntrophic processes, efficacy of process design and operation, and restoration and recovery.