Metagenomics: window to the microbial universe
Anne Jurkowski and Jo Handelsman

Metagenomics — the genomic analysis of communities of organisms — is reshaping the landscape of microbiology. Transcending individual genes and genomes, metagenomics offers access to all of the genomes in a community to reveal the secrets of the ‘uncultured world’ — the enormous number of microbial species that cannot currently be isolated into pure culture. Metagenomics will expand our ability to discover and benefit from microbial capabilities, improve our understanding of microbial communities and probably lead to major advances in medicine, agriculture, energy production and bioremediation.

1. Sampling
Samples can be taken from any environment, including invertebrates, the human gut and deep-sea vents.

2. Extraction of DNA
Genomic DNA is extracted directly from the sample to yield DNA fragments from all members of the microbial community. DNA is then analyzed directly or cloned prior to analysis. New methods can extract proteins or RNA directly from the sample. Most metagenomics projects currently focus on organisms with smaller amounts of DNA, such as ancient bacteria or archaea, rather than more complex organisms.

3. Construction of a metagenomic library
Commodity DNA is fragmented and cloned into plasmids for study and preservation. Plasmids are introduced into bacteria to create a metagenomic library — a living repository of all the DNA from the sampled microbial community.

4. Analysis
a. Sequence-based metagenomics: provides information on the distribution of functions in a community, directly links genetic organization and horizontal gene transfer. Approaches typically involve either sequencing of random clones to assemble contigs or clone libraries for detection of sequence information or identification of genes encoded by plasmids that detect a particular sequence. With both of these approaches, phylogenetic markers are sought on the clones of interest to link cloned sequences with the probable origin of the DNA.

b. Function-based analysis: enables identification of enzymes, antibiotics or other reagents in libraries for disease assessments. Approaches include:

   i. Heterologous expression, in which clones that express the desired function are identified. An important limitation to heterologous expression is that the demonstrative host bacteria must be able to express transcribable (translatomic) the genes for the products to be detected.

   ii. Selection, in which the clone expressing the desired function grows and others do not. Selections provide the most powerful approach to finding rare clones. Selectable characteristics include antibiotic resistance and metal resistance. A ‘functional-anchor approach’ involves adding all of the clones that express a certain function and sequencing them completely to determine the diversity of genomic environments from which that function originates.

   iii. Sequencing of random clones to assess the diversity of the whole collection. Selections provide the most powerful approach to finding rare clones. Selectable characteristics include antibiotic resistance and metal resistance. A ‘functional-anchor approach’ involves adding all of the clones that express a certain function and sequencing them completely to determine the diversity of genomic environments from which that function originates.

   iv. Metagenomic analysis: uniquely suited for identifying horizontal gene transfer.

Cross-disciplinary applications
Metagenomics makes possible insights that could help to address some of the most complex medical, environmental, agricultural and economic challenges in today’s world.

Medicine: understanding how the microbial communities that inhabit our bodies affect human health could lead to new strategies for diagnosing, treating and preventing diseases.

Earth sciences: exploring how microbial communities in the soil and ocean affect the atmosphere and environmental conditions could help us understand, predict and address global changes.

Alternative energy: harnessing the power of microbial communities might result in sustainable and eco-friendly bioenergy sources.

Environmental remediation: metagenomics could aid the development of microorganism-based tools for monitoring environmental damage and cleaning up oil spills, groundwater, sewage, nuclear waste and other hazards.

Biotechnology: taking advantage of the functions of microbial communities might lead to the development of new functional food and health products.

Agriculture: understanding the roles of beneficial microorganisms being in, on and around domesticated plants and animals could enable detection of diseases in crops and livestock, and aid the development of improved farming practices.

Biodefence and microbial forensics: metagenomic analysis can yield informative insights. Investigations of communication among bacteria, for example, have found that sublethal concentrations of many antibiotics induce quorum sensing. Using metagenomics to screen for signalling and inhibitory compounds might therefore yield molecules that are quorum-sensing inducers as well as antibiotics.

Metagenomics can also support community-wide assessments of metabolic and geochemical functions.

Illuminating biology
Metagenomics could answer some fundamental biological questions. Microbial communities are composed of thousands or millions of different but interdependent individuals who are closely related enough to be considered the same species, whereas others have few genes in common. Genetic material in these communities is often passed from one species to another, which poses questions such as: what is a genome; what is a species; how diverse is life; how do microbial communities react to change; and how do microorganisms evolve?

Metagenomics is uniquely suited for identifying genes involved in competition or cooperation. Such genes are almost impossible to identify outside of the community context, but metagenomic analysis can yield informative insights. Investigations of communication among bacteria, for example, have found that sublethal concentrations of many antibiotics induce quorum sensing. Using metagenomics to screen for signalling and inhibitory compounds might therefore yield molecules that are quorum-sensing inducers as well as antibiotics.

Metagenomics can also support community-wide assessments of metabolic and geochemical functions.