Metagenomics

Maureen A. O'Malley
Department of Philosophy, Quadrangle A14, University of Sydney, Sydney, NSW 2006, Australia
maureen.omalley@sydney.edu.au

Synonyms
Environmental genomics
Community genomics
Systems microbiology

Definition
Metagenomics is an emerging microbial systems science that is based on the large-scale analysis of the DNA of microbial communities in their natural environments. It merges the scope of genomic techniques with the breadth and depth of environmentally oriented approaches.

Characteristics

Studies of metagenomes are revealing the vast array of biodiversity in a wide range of environments, as well as new functional capacities of individual cells and communities, and the complex evolutionary relationships between different lineages. The 'meta' prefix of metagenomics can be given any of three interpretations that were developed after the term was coined in 1998 (Handelsman et al. 1998). The first reading of 'meta' is that the field *transcends* culturing limitations and obstacles to understanding microbial biodiversity. The second is that it generates an *overarching* understanding of genetic diversity from a world that will never be exhaustively sampled. The final interpretation of the meta-nature of metagenomics is that it aims to achieve an *aggregate*-level approach to biology, not an individual organism or single-genome focus (Committee on Metagenomics 2007: 13).

Metagenomics is thus conceived of as an approach through which the limitations of a focus on individual genes and particular species, as well as the separation of organisms from environments, can be transcended. Instead of individual genomes [monogenomes] or single-gene markers, metagenomics starts with large amounts of the DNA collected from microbial communities in their natural environments (Handelsman 2004; Hugenholtz and Tyson 2008). Probably the most developed and celebrated metagenomic discovery is that of a whole new class of genes in the rhodopsin family, called proteorhodopsin (PR) genes (Béjà et al. 2001). This example also provides a useful illustration of how data-driven and hypothesis-driven approaches interact to generate biological insight under the banner of metagenomics.

Metagenomic research programmes, such as those involving proteorhodopsin, examine biodiversity at every level (genetic, metabolic, organismal, communal) and attempt to generate evolutionary and functional explanations of such heterogeneity (Handelsman 2004). The term metagenomics in its largest sense therefore includes metatranscriptomics, metaproteomics and metametabolomics. These approaches culminate in what could be understood as metaorganismal metagenomics. The metagenomic study of the human microbiome is a prime example of such studies. Within this broad conceptualization of metagenomics, there is clearly much more going on than the production of sequence catalogues. Initial gene inventories are the basis of fundamental insights into microbial community structure and biogeography. They enable subtle understandings of eco-physiological characteristics of communities, and how adaptations to different environmental gradients have resulted in various metabolic and morphological strategies (e.g. capacities for movement) that have spread vertically and horizontally through community members (DeLong et al. 2006; Dinsdale et al. 2008). The genomic heterogeneity in environmental samples shows that the genomes of single isolated organisms can no longer be considered as typical of whole populations or species, and that more sophisticated understandings of evolutionary processes and the mechanisms of genetic exchange and recombination are needed (Allen and Banfield 2005).

Metagenomic analysis thus extends earlier findings of comparative microbial genomics, which showed the limitations of an evolutionary paradigm ruled by the tenets of vertical inheritance and mutation-driven species divisions giving rise to a single tree of life. Rather than focusing on individual organismal lineages, such metagenomic studies shift scientific attention to an overall evolutionary process in which diverse and diversifying metagenomes underlie the differentiation of interactions within dynamic ecosystems. Metagenomics thereby aligns itself with an area of investigation that is sometimes called horizontal genomics. Both are concerned with the plethora of mobile genetic elements available to microbial communities and with the ways in which the metagenomic resources they inherit are shared and utilized. From this perspective, the communal gene pool is evolutionarily important, and genetic material can be thought of as the community resource for a superorganism or metaorganism, rather than the exclusive property of individual organisms. Central to the broadly conceived project of metagenomics, therefore, are radically revised ways of thinking about biological entities (Doolittle and Zhaxybayeva 2010).

It is becoming increasingly clear that a range of fundamental questions about life on this planet will find their answers only with advances in system-based understandings of microbial communities in global environments. Whether we want to understand CO_2 levels, ocean acidification, methane release, food production or disease distribution, metagenomics can begin to address these issues and offer real hope of eventually finding solutions (Committee on Metagenomics 2007). Standard ways of doing biology, especially at the molecular level, are unlikely to prove sufficient for understanding and dealing with such complex wide-ranging issues, however. While there is no doubt that mechanistic and reductionist investigations of life processes have generated profound insights, there are limits to how far such investigations can take biology in understanding the dynamic stability of processes organized into hierarchies of interacting levels. More extensive understanding will require models that are able to incorporate both the knowledge of components revealed by conventional molecular approaches as well as the constraints and causal influences imposed by properties of the wider systems of which these constituents are parts. These sorts of insights are what advocates of the rapidly growing project of systems biology and metaorganismal metagenomics are aiming to generate (Vietes et al. 2009; Doolittle and Zhaxybayeva 2010).

Metagenomics broadens the scope of systems biological inquiry to the mutuality of system-environment determination. In system-level understandings of microbial communities, the metaorganism is conceived of as deriving causal powers from the interactions of the individual components that constitute it, but at the same time those components are themselves understood to be controlled and coordinated in various ways by the causal capacities of the metaorganism. Metagenomics also goes beyond standard approaches to the understanding of molecules in environments. As well as examining how genes, for example, achieve cellular effects, metagenomics determines the ways in which environments influence the composition of genomes and metagenomes, and how those entities work as niche-defining indicators of environmental properties. Although metagenomics might be far away from the full implementation of such a multi-level and suitably interdisciplinary research programme, future advances in microbial systems science are anticipated as the results of today's technically restricted but theoretically far-reaching programmes of research.

Although metagenomics may still be limited in regard to the extent to which its copious data can be interpreted, it has already generated a whole new understanding of biodiversity and its distribution. Viral metagenomics, for example, is finally giving an indication of the diversity and activity of the prolific genetic reservoirs constituted by viruses and how they fundamentally shape microbial communities. The scope and adaptive importance of the variability indicated by metagenomic analysis has generated huge ambitions for applications to biotechnology and bioremediation, and these hopes feed back into and re-stimulate metagenomic efforts to generate biological understanding.

Cross-references

Data-driven
Hypothesis-driven
Metaorganism
Superorganism
Microbiome
System
Genomic

Systems biology
Interdisciplinarity
Models
Dynamic
Mechanistic
Reductionist

References

Allen EE, Banfield JF (2005) Community genomics in microbial ecology and evolution. Nat Rev Microbiol 3:489-498

Béjà O, Spudich EN, Spudich JL, Leclerc M, DeLong EF (2001) Proteorhodopsin phototrophy in the ocean. Nature 411:786-789

Committee on Metagenomics: Challenges and Functional Applications, National Research Council (2007) The new science of metagenomics: Revealing the secrets of our microbial planet. National Academies Press, Washington DC

DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard NU, Martinez A, Sullivan MB, Edwards R, Brito BR, Chisholm SW, Karl DM (2006) Community genomics among stratified microbial assemblages in the ocean's interior. Science 311:496-503

Dinsdale EA, Edwards RA, Hall D, Angly F, Breitbart M, Brulc JM, Furlan M, Desnues C, Haynes M, Li L, McDaniel L, et al. (2008). Functional metagenomic profiling of nine biomes. Nature 452:629-633.

Doolittle W, Zhaxybayeva O (2010) Metagenomics and the units of biological organization. BioScience 60:102-112

Handelsman J (2004) Metagenomics: Application of genomics to uncultured microorganisms. Microbiol Mol Biol Rev 6:669-85

Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM (1998) Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products. Chem Biol 5:R245-49

Hugenholtz P, Tyson GW (2008) Metagenomics. Nature 455:481-483

Vietes JM, Guazzaroni M-E, Beloqui A, Golyshin PN, Ferrer M (2009) Metagenomics approaches in systems microbiology. FEMS Microbiol Rev 33:236-255