Lawrence Berkeley National Laboratory
Lawrence Berkeley National Laboratory

Title
Omics for understanding microbial functional dynamics

Permalink
https://escholarship.org/uc/item/9gr429g3

Author
Jansson, J.

Publication Date
2012-03-01

DOI
DOI:10.1111/j.1462-2920.2011.02518.x

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Omics for understanding microbial functional dynamics

Janet K. Jansson[1]*, Josh D. Neufeld[2], Mary Ann Moran[3], Jack A. Gilbert[4,5]

[1]*Lawrence Berkeley National Laboratory, Earth Sciences Division Berkeley, CA, USA.*
[2]*Department of Biology, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.*
[3]*Department of Marine Sciences, University of Georgia, Athens, GA 30602, USA.*
[4]*Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, USA.*
[5]*Department of Ecology and Evolution, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA.*

As a discipline, the field of microbial ecology has generally been limited by access to tools that have sufficient depth of resolution to enable exploration of complex microbial communities to determine 'who is there?' and 'what they are doing?'. Until the mid 1980s most of our knowledge was based on the study of cultured microbial isolates or net enzymatic activities measured from laboratory enrichment experiments or environmental samples, without knowledge of the particular members of the community that were contributing to those functions. Since then, microbiologists have added a variety of molecular tools to our toolbox, allowing us to directly assess the identities and functions of microbial communities in a variety of environments without the necessity of prior cultivation.

Most of the advances in molecular tools for assessing microbial community function have focused on the sequencing of nucleic acids. Sequencing of genomic or metagenomic DNA provides phylogenetic and functional gene information about specific organisms or communities respectively. These types of data have proven to be highly valuable for characterizing microbes that inhabit the various habitats on Earth, and importantly, their possible roles in the function of that ecosystem. However, one should keep in mind that DNA sequence data only represents the potential metabolic capacity of a system. In many environments, DNA that is sequenced may originate from dead or dormant cells. Further, even actively growing microbes express only a fraction of their genes at any given time. For example, *Escherichia coli* genes are tightly regulated depending on, for example, the concentration of nutrients in their media (Tao *et al*., 1999). Only a third to a half of the genes in a given organism are expressed at any given time (Passalacqua *et al*., 2009).

Methodology to address the longstanding dilemma of 'which organisms are active?' include the use of viability stains to first sort (e.g. with fluorescently activated cell sorting; FACS) and then sequence genomes from active populations, using stable isotopes to detect community members that are metabolically active (e.g. stable-isotope probing; SIP; Neufeld *et al*., 2007) and the use of bromodeoxyuridine (BrdU), a thymidine analogue, to specifically label DNA from growing cells that can then be isolated and sequenced (Edlund and Jansson, 2008; Mou *et al*., 2008). Alternatively, one can directly extract and sequence RNA to determine which microbes are active and which genes are transcribed. When applied to an entire microbial community, this analysis is referred to as 'metatranscriptomics'. The metatranscriptomic approach has proven to be more challenging than metagenomics because RNA is less stable than DNA and because most of the RNA that is extracted is ribosomal RNA, usually representing over 90% of the total RNA (Urich *et al*., 2008). Ribosomal RNA sequences can inform about active community members due to a correlation between metabolic activity and ribosome abundance (Rehman *et al*., 2010), but do not reveal the genes and pathways expressed under a given set of conditions. In order to specifically focus on mRNA transcripts, several methods have been developed to enrich the mRNA fraction of the total RNA pool, usually by subtraction of the ribosomal RNA prior to sequencing (Fig. 1). The problem with loss of mRNA during the multiple processing steps is particularly

E-mail jrjansson@lbl.gov; Tel. (+1) 510 486 7487; Fax (+1) 510 486 7152. The submitted manuscript has been created in part by U. Chicago Argonne, LLC, Operator of Argonne National Laboratory ('Argonne'). Argonne, a US Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The US Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.

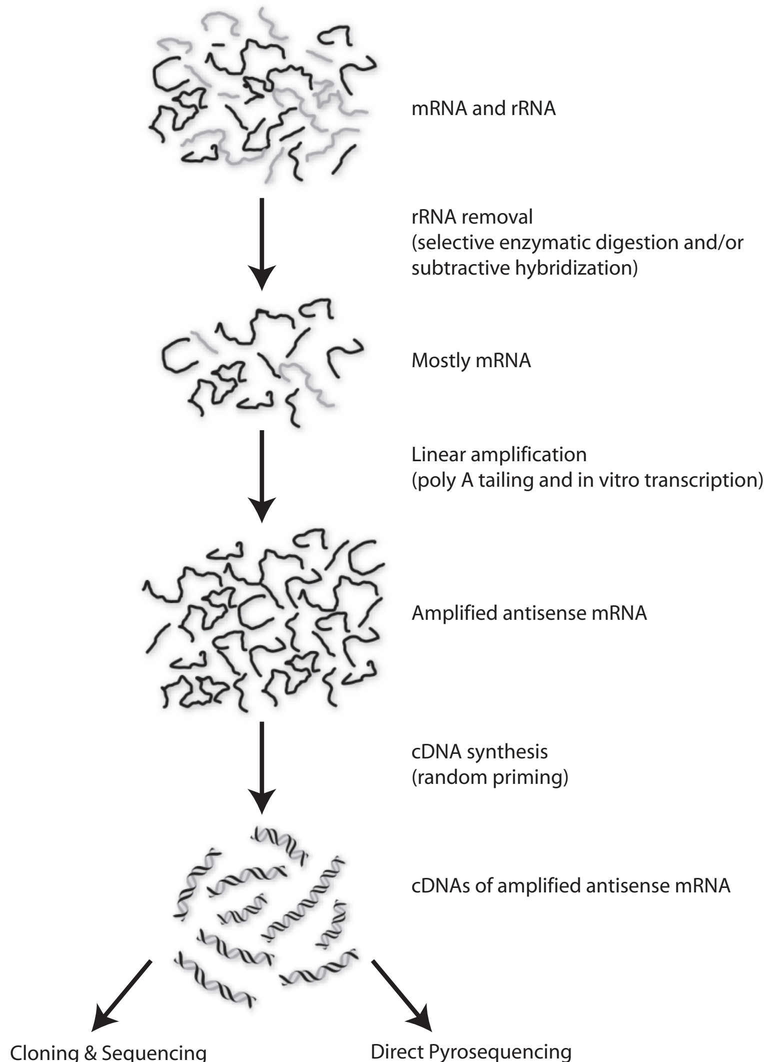

Fig. 1. Outline of a metatranscriptomic approach for an mRNA enrichment.

problematic since current sequencing platforms often require microgram quantities of DNA synthesized from RNA template (cDNA) for sequencing. In addition to direct capture of expressed transcripts or intact ribosomes, the recent advent of mRNA-SIP now provides a direct mechanism to enrich for transcripts that are expressed during the use of particular labelled carbon sources (e.g. Dumont *et al*., 2011).

An emerging step in the 'omics pipeline' is to determine which proteins are produced by microorganisms in particular environments. Proteins are the ultimate product of gene regulation and they provide direct evidence for gene function. The study of proteins in mixed communities is known as 'metaproteomics', first demonstrated for the relatively simple acid mine drainage (AMD) community (Ram *et al*., 2005). Recently, metaproteomics has been used to identify proteins expressed in the human gut (Verberkmoes *et al*., 2009), soil (Chourey *et al*., 2010) and the surface ocean (Morris *et al*., 2010). Although metaproteomics is still a relatively young field, its feasibility is increasingly recognized with advances in gel-free shotgun mass spectrophotometry-based approaches. Currently, it is possible to identify hundreds or thousands of proteins in a complex sample. A major challenge for progress in metaproteomics is increasing the analysis depth to parallel the quantity of genetic information obtained through metagenomics or metatranscriptomic techniques on the same sample.

Important evidence for community function includes the metabolites produced by a microorganism or collectively by a microbial community (i.e. 'metabolomics'). Although metabolites cannot yet be directly linked to a given gene, it is now possible to correlate the presence of metabolites to specific members of the community (Jansson *et al*., 2009). Importantly, recent advances in functional annotation of metagenomic data and computational processing of inferred metabolomic networks enable linkages to be made between the metagenome/metatranscriptome and

the metametabolome (Larsen *et al.*, 2011). Once we can use comparative metagenomics (i.e. across time/space or in response to a known manipulation) to explore the changing metabolic potential of a microbial community, it may be possible to predict relative changes in the turnover of specific metabolites. This can then be related to observed changes in the relative abundance of metabolites, allowing validation studies to be performed using multi-omics approaches.

'Megasequencing' projects, which aim to leverage considerable 'omics' capability against challenges in microbial ecology, will redefine what it means to explore microbial space. One such effort, the Earth Microbiome Project (http://www.earthmicrobiome.org), is currently running a pilot study to sequence the metagenomes of 10 000 environmental samples. The aim of this project is to produce a global 'biomap' of microbial taxonomy and function through space and time to explore the metabolic potential of different ecosystems, and essentially the functional resilience of the microbial world.

Our ability to link microbial communities' composition and function has matured rapidly; methodology and cost are less of an impediment to rapid progress in microbial ecology. But there remains room for transformative growth and the potential to further redefine our scientific discipline. Multi-meta-omics is on the threshold of being able to describe the totality of a microbial ecosystem, which has been the holy grail for microbial ecologists for quite some time. The ability to be able to use omics information to map every trophic interaction and biochemical pathway, and to model those dynamic genetic interactions for descriptive and predictive ends, may still be a long way off, if ever achievable. But along the way, these technological advances for exploration of microbial communities on Earth should result in fundamental contributions towards our understanding of the role of microbes on our planet.

Acknowledgements

This work was supported in part by the US Department of Energy under Contracts DE-AC02-06CH11357 and BER Contract No. DE-SC0004953, NSF grant MCB-0702125, and in part by the US Department of Energy under Contract No. DE-AC02-05CH11231 with the Lawrence Berkeley National Laboratory. Funding for J.D.N. was provided by a Strategic Projects Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K.L., Tom, L.M., *et al.* (2010) Direct cellular lysis/protein extraction protocol for soil metaproteomics. *J Proteome Res* **9:** 6615–6622.

Dumont, M.G., Pommerenke, B., Casper, P., and Conrad, R. (2011) DNA-, rRNA- and mRNA-based stable isotope probing of aerobic methanotrophs in lake sediment. *Environ Microbiol* **13:** 1153–1167.

Edlund, A., and Jansson, J.K. (2008) Use of bromodeoxyuridine immunocapture to identify psychrotolerant phenanthrene-degrading bacteria in phenanthrene-enriched polluted Baltic Sea sediments. *FEMS Microbiol Ecol* **65:** 513–525.

Jansson, J., Willing, B., Lucio, M., Fekete, A., Dicksved, J., Halfvarson, J., *et al.* (2009) Metabolomics reveals metabolic biomarkers of Crohn's disease. *PLoS ONE* **4:** e6386.

Larsen, P.E., Collart, F., Field, D., McGrath, J., Quinn, J., Gilbert, J.A., *et al.* (2011). Predicted Relative Metabolomic Turnover (PRMT): determining metabolic turnover from a coastal marine metagenomic dataset. Conference proceedings for BIOSTEC 2011; Meta2011, Metagenomic Sequence Data Analysis.

Morris, R.M., Nunn, B.L., Frazar, C., Goodlett, D.R., Ting, Y.S., and Rocap, G. (2010) Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J* **4:** 673–685.

Mou, X., Sun, S., *et al.* (2008) Bacterial carbon processing by generalist species in the coastal ocean. *Nature* **451:** 708–711.

Neufeld, J.D., Vohra, J., Dumont, M.G., Lueders, T., Manefield, M., Friedrich, M.W., and Murrell, J.C. (2007) DNA stable-isotope probing. *Nat Protoc* **2:** 860–866.

Passalacqua, K.D., Varadarajan, A., Ondov, B.D., Okou, D.T., Zwick, M.E., and Bergman, N.H. (2009) Structure and complexity of a bacterial transcriptome. *J Bacteriol* **191:** 3203–3211.

Ram, R.J., Verberkmoes, N.C., Thelen, M.P., Tyson, G.W., Baker, B.J., Blake, R.C., 2nd, *et al.* (2005) Community proteomics of a natural microbial biofilm. *Science* **308:** 1915–1920.

Rehman, A., Lepage, P., Nolte, A., Hellmig, S., Schreiber, S., Ott, S.J. (2010) Transcriptional activity of the dominant gut mucosal microbiota in chronic inflammatory bowel disease patients. *J Med Microbiol* **59:** 1114–1122.

Tao, H., Bausch, C., Richmond, C., Blattner, F.R., and Conway, T. (1999) Functional genomics: expression analysis of Escherichia coli growing on minimal and rich media. *J Bacteriol* **181:** 6425–6440.

Urich, T., Lanzen, A., Qi, J., Huson, D.H., Schleper, C., and Schuster, S.C. (2008) Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. *PLoS ONE* **3:** e2527.

Verberkmoes, N.C., Russell, A.L., Shah, M., Godzik, A., Rosenquist, M., Halfvarson, J., *et al.* (2009) Shotgun metaproteomics of the human distal gut microbiota. *ISME J* **3:** 179–189.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.

Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.