Metabolomics Vol. 1, No. 1, January 2005 (© 2005)
DOI: 10.1007/s11306-005-1111-7

Metabolomics – the way forward

Royston Goodacre

School of Chemistry, The University of Manchester, PO Box 88, Sackville Street, Manchester, M60 1QD, UK

Welcome to the inaugural issue of *Metabolomics*, a broad-based journal that aims to publish the most significant current research within the development of metabolomics.

In 1999, the physicist Freeman Dyson said that "Surprises in science often arise from new tools rather than from new concepts" (Madou, 2002), and today whole genome sequencing projects (http://www.tigr.org/tdb), of both prokaryotes and eukaryotes, continue to remind us that our knowledge of how an organism functions at the molecular level is really very poor. Typically 20–40% of the open reading frames found in these genomes have no know function. These so called "orphan genes" must have some cellular role else they would have been lost during evolution. Therefore one of the big "games" in the life sciences is to find out what their purpose is, and both new tools and novel concepts will be necessary to achieve this. Recently, the completion of the human genome (Sanger Institute Press Release, 2003) has accelerated demand for determining the biochemical function of orphan genes and for validating them as molecular targets for therapeutic intervention. The search for biomarkers that can serve as indicators of disease progression or response to therapeutic intervention has also increased; as has the "holy grail" search for prognostic markers. One of the big hopes within clinical microbiology is that novel drug targets will be found allowing the production of new pharmacophores since pathogens are increasingly becoming resistant, at alarming rates, to most currently prescribed antimicrobials.

Functional analyses (Figure 1a) have thus emphasised analyses at the level of gene expression (transcriptomics), protein translation (proteomics; including post-translational modifications) and more recently the metabolite network (metabolomics), with a view within a systems biology approach of defining the phenotype and finally bridging the genotype-to-phenotype gap (Fiehn, 2002). Even the representation in Figure 1a is simplistic, since whilst in our linear conception of the cell the general flow of information goes from gene to transcript to protein to metabolite to phenotype, there are multiple feedback loops from metabolites to proteins and/or transcripts, as well as others. Indeed our traditional view of metabolism (Figure 1b) is no longer true and the cellular processes are in reality networked and should be represented as dynamic protein complexes interacting with neighbourhoods of metabolites (Figure 1b). The construction and visualisation of the metabolic network (Barabasi and Oltvai, 2004) is certainly a big challenge for the future, as is a full understanding of the fluxes through them and their control (Fell, 1996).

Another consideration is that whilst the tools for transcriptomics and proteomics are relatively simple, the same is not true for metabolites which may have incredibly short ½ lives making analysis even more complicated. A nucleic acid (micro-) array needs only to recognise the sequence of 4–5 nucleotide building blocks, and for proteomics the ordering (plus any modifications) of 20 primary amino acids. By contrast, the major challenges to overcome when measuring the metabolome are its chemical complexity and heterogeneity of metabolites, and the wide dynamic range of these biochemical species. There is no Star Trek "Tricorder" and so the need and development of parallel, high throughput analyses is considerable, and will be a major focus for comprehensive metabolome analysis.

Once the metabolome data are generated something has to be done with them. A typical metabolomics experiment is likely to generate huge data floods, or more likely torrents, avalanches or tsunami. These descriptions are often thought of in terms of natural disasters and experiments must be carefully thought out since hiring a statistician after the data have been collected is like hiring a physician when the patient is in the morgue. He might be able to tell you what went wrong, but is unlikely to be able to fix it (Anon).

Even when experiments are designed correctly, the simple "stare and compare" approaches are completely inadequate and alternative, multivariate statistics, chemometric and machine learning-based analyses are desperately needed to turn our data into knowledge.

"Progress in science depends on new techniques, new discoveries and new ideas, probably in that order"
Sydney Brenner, *Nature*, 5 June 1980

*To whom correspondence should be addressed.
E-mail: Roy.Goodacre@manchester.ac.uk

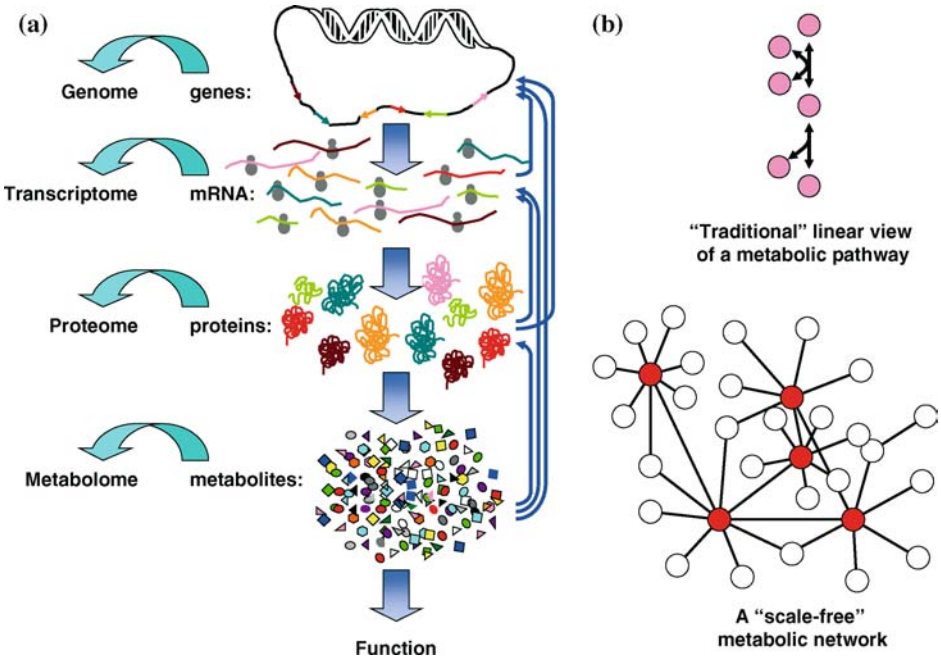

Figure 1. Figure (a) General schematic of the omic organisation. The general flow of information is from genes to transcripts to proteins to metabolites to function (or phenotype); whilst blue vertical arrows indicate interactions regulating respective omic expression. (b) Our "traditional" linear view of a metabolic pathway and "scale free" connections in a metabolite neighbourhood.

Metabolomics is a vibrant diverse field on the exponential part of the growth curve. There are many global metabolomics-based research initiatives on going. An international society has been formed – The Metabolomics Society (metabolomicssociety.org) – the mission of which is to promote the growth and development of the field of metabolomics internationally. *Metabolomics* will be the official journal of this society.

Metabolomics aims to publish the most significant current research in the areas of: the development of various technology platforms for metabolomics, metabolite target analysis, metabolic profiling, and metabolic fingerprinting; improvements in data preparation, storage, curation and analyses; comparative integrated studies with transcriptomics and proteomics including within a systems biology context; and the application of metabolomics within man, animals, plants and microbes. *Metabolomics* will be a broad-based journal that is indispensable to those whose work has implications of near-term practical benefit.

I look forward to receiving your work in this exciting area.

References

Barabasi, A.L. and Oltvai, Z.N. (2004). Network biology: understanding the cell's functional organisation. *Nature Reviews Genetics* **5**, 101–113.

Fell, D.A. (1996). *Understanding the control of metabolism*. Portland Press, London.

Fiehn, O. (2002). Metabolomics – the link between genotypes and phenotypes. *Plant Molecular Biology* **48**, 155–171.

Madou, M.J. (2002). *Fundamentals of Microfabrication: The Science of Minaturisation, 2nd edition*. CRC Press, 752pp.

Sanger Institute Press Release (2003) The finished Human genome – Welcome to the Genomic Age. 14 April 2003 http://www.sanger.ac.uk/Info/Press/2003/030414.shtml.