DDT Vol. 9, No. 6 March 2004

Metabolic profiling: pathways in discovery

Royston Goodacre, Department of Chemistry, UMIST, PO Box 88, Sackville Street, Manchester, UK M60 1QD, tel: +44 161 200 4480, fax: +44 161 200 4519, e-mail: R.Goodacre@umist.ac.uk

Whole genome sequencing projects have been producing vast amounts of potentially informative data for about eight years (http://www.tigr.org/tdb) and the number of available whole genomes is increasing almost exponentially. However, this is only the beginning of the story, and we soon began to realize the paucity of our knowledge with respect to the function of novel genes.

In fact, sequencing the microbiologist's pet organism *Escherichia coli* revealed that a staggering 38% of the total 4288 ORFs had never been observed or studied before [1]. Recently, the completion of the human genome [2] 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. Functional analyses have thus emphasized analyses at the level of gene expression (transcriptomics), protein translation (proteomics) and the metabolite network (metabolomics), with a view within a systems biology approach of defining the phenotype and bridging the genotype-to-phenotype gap [3].

Metabolome analysis

It was the analysis of the metabolome that drew many industrial and academic scientists together on 8–9 December 2003 at the Hyatt Regency Princeton, NJ, USA. Despite one of the worst snow storms to hit the East coast of North America for over 50 years, which had temporarily closed a few of the local airports, the Cambridge Healthtech Institute 3rd annual *Metabolic Profiling: Pathways in Discovery* conference went ahead unhindered.

The 'metabolome' can be defined as the quantitative complement of all the low molecular weight molecules present in cells in a particular physiological or developmental state [4]. The measurement of the totality of the metabolic pool is the ultimate goal for metabolomics; however, it is obvious that with vastly different chemical species that need to be measured quantitatively over large dynamic ranges, there is still a long way to go. Thus, there are opportunities for instrument manufacturers to develop new and better measurement strategies and one of the new approaches adopted by several speakers, and detailed by Paul Gamache (Director, Applications Development, ESA Inc; http://www.esainc.com), was a coulometric electrochemical array detector which measures redox active small molecules.

HPLC is used as the chromatographic separation technique and the eluent is split to a mass spectrometer and to a parallel array of up to 16 coulometric EC array cells operating at increasing mV. Although this approach does not give structural information on metabolites, it does have a low limit of detection, comparable to MS, and has potential applicability to studying pathological processes because redox active molecules are commonly associated with, for example, oxidative stress, drug metabolism and xenobiotic toxicity.

Metabolic networks

We were all taught in biochemistry lectures that metabolism is a linear process; A goes to B, B onto C, and so on. However, there has recently been a shift from mental constructs involving metabolic pathways to those based on metabolic neighbourhoods, and many would argue that the 'Boehringer' metabolic pathways map needs to be updated both radically and conceptually. This was the topic of an excellent presentation by Zoltan Oltvai from the Department of Pathology, Nortwestern University (http://www.nortwestern.edu) entitled *Flux organisation in metabolic networks*.

Protein complexes have been shown recently to exist [5,6] and so it is not surprising to imagine metabolic networks being at the heart of these complexes, however, how they are arranged and how metabolite fluxes operate need to be understood. Zoltan discussed the theoretical representation of substrate connections in metabolic networks and how these can be considered as fundamental building blocks of cellular organization [7]. These modules can be combined in a hierarchical way to form some metabolic function and data were presented that showed this with *E. coli*. Global flux organization in the metabolic network of *E. coli* was also discussed with supporting data. This is an area of immense importance as we strive to accomplish global integrative analyses at all the 'omic' levels and will

1359-6446/04/$ – see front matter PII: S1359-6446(04)03027-2

be essential within systems biology where the aim is to model the genetic, macromolecular and metabolic networks.

In silico gene-expression analysis

Masaru Tomito from the Institute for Advanced Biosciences, Keio University in Japan (http://www.keio.ac.jp), highlighted the many impressive activities that his research group and others in Keio are undertaking. One of these was the 'E-cell project' (http://www.e-cell.org/) which is an *in silico* analysis of the expression level of genes and the metabolic network within bacteria. He reminded us that 'biology is quantitative', and my own feeling is that this has been somewhat 'forgotten' by many biologists but is being re-recognized as being paramount to understanding biological systems.

Thus, accurate reliable quantitative data are needed to define kinetics and steady-states, and Tomito's research highlighted this for the analysis of the *E. coli* metabolome where many analytical approaches (mostly MS-based) were used to make the analyses as comprehensive as possible. Of course there are still a few gaps in predicting metabolic networks from the available genomes but these will ultimately be plugged.

Metabolomics: more to come

Metabolomics as a discipline is on the late lag-phase of the growth curve and we can expect a great deal more interest from pharmaceutical companies as they tap into the role that metabolic profiling will have in the discovery of new biomarkers or the toxicological effect of drugs on the mammalian system. Even at this early stage, if an approach can be labelled 'traditional' it probably belongs to NMR, which has been widely used to try to understand how the metabolic response of an organism changes to pathophysiological stimuli [8]. However, NMR is not very sensitive to low amounts of metabolites and so this method is being supplemented and replaced by MS-based methods, which have lower limits of detection, and in tandem MS can give some structural information on the fly.

This was exemplified by John Haselden (Head of UKs GSK Metabonomic group; http://www.gsk.com) who gave a fascinating account of the activities of GSK in this area and posed the question; 'Is metabolic profiling value added or a distraction to pharma?' Their approach has been to combine NMR and HPLC–ESI–MS and investigate preclinical drug discovery particularly with respect to toxicity of potential drugs. Haselden showed lots of good data and convinced us that this approach does indeed have added value.

Outside of the main arena there were some discussions on where metabolomics as a field might be going and how the many interesting broad ranging activities in this area can be properly publicised and brought together. The first in this conference series has lead to the publication of a book on *Metabolic Profiling* [9]. Overall, this was an interesting conference and, because metabolic profiling should have moved into the early exponential phase of the growth curve, the 4th *Metabolic Profiling: Pathways in Discovery* conference should be an even more interesting and worthwhile trip.

References

1 Blattner, F.R. *et al.* (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–1474

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

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

4 Oliver, S.G. *et al.* (1998) Systematic functional analysis of the yeast genome. *Trends Biotechnol.* 16, 373–378

5 Ho, Y. *et al.* (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415, 180–183

6 Gavin, A.C. *et al.* (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141–147

7 Ravasz, E. *et al.* (2002) Hierarchical organisation of modularity in metabolic networks. *Science* 297, 1551–1555

8 Nicholson, J.K. *et al.* (1999) 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 29, 1181–1189

9 Harrigan, G.G. and Goodacre, R., eds (2003) *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis*, Kluwer Academic Publishers