Tumor 'Omics'

JR Griffiths
Cancer Research UK Cambridge Institute, Cambridge CB2 0RE, UK

The Human Genome Project, completed in 2003, spawned a host of new sciences that have already revolutionised research into the biology and treatment of cancer, and are beginning to influence oncology in the clinic. The main four omic sciences are:

- Genomics – the totality of genes in an organism (note that non-coding and regulatory DNA is important as well as the protein coding genes)
- Transcriptomics (or expressomics) – the mRNA complement of an organism, tissue type, or cell
- Proteomics - the totality of proteins (expressed genes) in an organism, tissue type or cell
- Metabolomics - the totality of small-molecule metabolites in an organism, tissue type or cell (with sub-sets such as glycomics or lipidomics)

Many other names for omics sciences have been coined, such as spliceomics (the totality of the alternative splicing protein isoforms) or interactomics (the study of molecular interactions in a holistic fashion), but they are not yet widely used, so I shall only talk about the four omic sciences listed above, which cover (i) the genes, (ii) their transcription into mRNA, (iii) their translation into proteins (mainly enzymes), and (iv) the small molecule metabolites that these enzymes use to manipulate the reactions of the metabolites.

There are ~25,000 genes in the human (and mouse) genome; because most of them are present in numerous short sections (introns), interspersed by lengths of non-coding DNA (exons), they can be transcribed into a much larger number of mRNAs (the "transcriptome"). These mRNAs in turn can be spliced to make mRNA sequences that can be translated into an even greater number of protein species (the "proteome"), although many of these are "splice variants" of standard proteins. To complicate matters further, however, these proteins can undergo post-translational modifications, e.g. by phosphorylation. Many of the proteins in the proteome are enzymes and they control the interconversion of the small molecule metabolites in the cell; those metabolites constitute the "metabolome".

The omics sciences impinge on the worlds of NMR and MRI in two main ways. Firstly, many preclinical cancer studies now use omics methods, or data obtained from omics databases (e.g. genomes). Transcriptomic methods such as gene expression microarrays or quantitative polymerase chain reactions (qPCR) are now widely used both in preclinical and clinical research, and proteomics is catching up. Furthermore, the next generation of gene sequencing instruments is likely to be so fast that they will take over many of these roles. Thus knowledge of these methods is becoming requisite for any cancer researcher, and they are likely to figure increasingly in MRI and MRS studies.

Metabolomics

The only omic science that can be performed directly by NMR is metabolomics, the study of the metabolome, which is the totality of small-molecule metabolites in an organism, cell or disease state. Unlike the genome (but like the transcriptome or proteome), the metabolome of a cell is state-specific – it can change from minute to minute, depending on factors such as stage in the cell cycle or an organism's environment.

How does metabolomics differ from the closely related technique of metabonomics, which uses the same analytical and statistical tools? Rather than being primarily concerned with the cell or tissue metabolome, metabonomics "broadly aims to measure the global, dynamic response of living systems to biological stimuli or genetic manipulation" (Nicholson & Lindon, 2008); it usually focuses on body fluids rather than cell or tissue metabolites. In practise, despite these basic philosophical differences, the two terms are often used interchangeably.

Metabolomics is usually performed by mass spectrometry which is far more sensitive than NMR and detects a much wider range of metabolites. However, because chemical derivatisation and prior chromatographic separation are required to prepare a sample for mass spectrometry, NMR has compensating advantages, which have been reviewed by Robertson (2005). NMR is more reproducible than mass spectrometry, it is more quantitative, sample preparation is easier, and there is less scope for sample bias due to factors such as ion suppression. Thus from NMR we know with good precision and reproducibility the concentrations of a relatively small number of metabolites in a sample. For these reasons NMR is still a competitive method. In practice, it is only possible to assign about 50 metabolites in a 600 MHz spectrum (though many more excretory products than that can be detected in urine); as that constitutes only a small sub-set of the 2,000-3,000 small molecules in a metabolome, it is more reasonable to talk about "metabolic profiles".

NMR-derived metabolic profiles lend themselves to the monitoring of experiments in which metabolomes are perturbed, often by genetic modification. The classic study was performed by Raamsdonk et al (2001) with the aim of assigning functions to "silent" genes in yeast – i.e. genes whose deletion does not induce an obvious phenotypic change. They obtained metabolic profiles from yeasts in which genes of known functions had been deleted, and then performed cluster analysis on them. Their hypothesis was that metabolic profiles generated by deletion of genes with a similar function, or genes that code for proteins in a single metabolic pathway, will be similar, and that they will therefore cluster together. If the metabolic profile generated by deletion of an unknown, silent gene clusters within this group of genes with a known function, then we can infer that the unknown gene performs a function related to that of the known genes.

Studies on gene function can be performed very simply by NMR metabolic profiling. A typical problem would be comparison of a cell line that has a single gene knocked out against the normal or "wild-type" cells in which that gene was functional. Cells from the two lines are cultured in flasks, in the normal way, and harvested by scraping into perchloric acid, instantly denaturing all the proteins and "freezing" all metabolic interactions. The protein pellet is spun down and usually assayed to provide a quantification standard (alternatively cell number can be used for normalisation), or it can be used for proteomic analysis. The excess perchloric acid is precipitated from the supernatant as potassium perchlorate, which is then spun down. The pH of the

supernatant is neutralised, and it is then freeze dried and re-dissolved in D_2O; a TSP standard can be added at this point for chemical shift standardisation and quantification. Finally the ^1H spectrum is obtained either in 1D or 2D. All the metabolites are extracted and assayed simultaneously, in the same sample, so all errors should be identical; thus their relative concentrations are known with great precision and the metabolic profiles of the genetically-modified and wild-type cells can easily be compared. There are also extraction methods for lipids, or for combined lipid and soluble metabolite preparations. One problem with making extracts of cultured cells is that the plastic on which they are grown will leach chemicals into the extract, and these can introduce unexpected artefacts.

Alternatively, solid tissues from genetically modified model tumours in animals or even biopsies from tumours in patients can be studied; in the latter case one would probably compare the metabolic profile of the tumour biopsy with that of a biopsy from the normal tissue of origin of that tumour (Pollard et al. 2005). Solid biopsies can be studied by perchloric acid extraction or by HR-MAS NMR. Similar studies can be performed on cultured cells or model tumours before and after treatment with anticancer drugs (Griffiths & Stubbs, 2005).

Since metabolic profiling inherently produces patterns of metabolite concentrations it is often necessary to use clustering or pattern recognition methods for analysing it. The FANCY method used principal component analysis (PCA) but many other methods can be applied.

A wide variety of problems has already been tackled by this simple and versatile method (e.g. Porstman et al, 2005; Weeks et al, 2006, Mayr et al., 2008). Metabolic profiling becomes particularly powerful when used in conjunction with proteomic methods (Mayr et al, 2004, 2008, and references cited therein), since one can measure changes in the concentration of both the metabolites themselves and the enzymes that catalyse their formation. Another synergistic method is transcriptomics; again, one can combine metabolomic data with information about the transcription of the relevant genes. Often metabolomic data serves to clarify the results of studies using many other methodologies, e.g. Porstman et al, 2005.

References

Griffiths, JR**,** & Stubbs, M. The cancer metabolic phenotype: exploiting the cancer metabolome in drug discovery.. In: The Oncogenomics Handbook: understanding and treating cancer in the 21st century. The Humana Press. pp135-149 (2005)

Mayr, M, et al. Vascular proteomics: Linking proteomic and metabolomic changes. Proteomics 4: 3751-3761 (2004)

Mayr, M, et al. Combined metabolomic and proteomic analysis of human atrial fibrillation. Journal of the American College of Cardiology 51: 585 – 594 (2008)

Nicholson JK and Lindon JC. Metabonomics. Nature 455: 1054-1056 (2008)

Pollard, PJ, et al. Accumulation of Krebs cycle intermediates and over-expression of HIF-1a in tumours which result from germline FH and SDH mutations. Human Molecular Genetics 14: 2231-2239 (2005)

Porstmann, T, et al. PKB/Akt induces transcription of enzymes involved in cholesterol and fatty acid biosynthesis via activation of SREBP.. Oncogene. 24: 6465-81 (2005).

Raamsdonk, LM, et al. Functional genomics via the metabolome: a FANCY approach to characterising mutations with a silent phenotype. Nature Biotech. 19, 45-50 (2001)

Robertson, D. G. Metabonomics in Toxicology: A Review. Toxicol. Sci., 85, 809-822 (2005)

Weeks ME, et al. A parallel proteomic and metabolomic analysis of the hydrogen peroxide-and Sty 1p-dependent stress response in *Schizosaccharomyces prombe*. Proteomics 6: 2772-2796 (2006).