Chapter 16
Metabolomics and Metabonomics

Metabolism is the ensemble of chemical transformations carried out in living tissue (§10.2); operationally it is embodied in the matter and energy fluxes through organisms. Metabolomics is defined as the measurement of the amounts (concentrations) and locations of the all the metabolites in a cell, the metabolites being the small molecules ($M_r \lesssim 1000$; e.g., glucose, cAMP,[1] GMP,[2] glutamate, etc.) transformed in the process of metabolism (i.e., mostly the substrates and products of enzymes).[3] The quantification of the amounts of expressed enzymes is, as we have seen, proteomics; metabolomics is essentially an extension of proteomics to the activities of the expressed enzymes, and it is of major interest to examine correlations between expression data and metabolite data.[4]

Metabonomics is a subset of metabolomics and is defined as the quantitative measurement of the multiparametric metabolic responses of living systems to pathophysiological stimuli or genetic modification, with particular emphasis on the elucidation of differences in population groups due to genetic modification, disease, and environmental (including nutritional) stress. In the numerous cases of diseases not obviously linked to genetic alteration (mutation), metabolites are the most revealing markers of disease or chronic exposure to toxins from the environment and of the effect of drugs. As far as drugs are concerned, metabonomics is effectively a subset of the investigation of the absorption, distribution, metabolism, and excretion (ADME) of drugs.

---

[1] Cyclic adenosine monophosphate.

[2] Guanosine monophosphate.

[3] The official classification of enzyme function is that of the Enzyme Commission (EC), which recognizes six main classes: 1, oxidoreductases; 2, transferases; 3, hydrolases; 4, lyases; 5, isomerases; and 6, ligases. The main class number is followed by three further numbers (separated by points), whose significance depends on the main class. For class 1, the second number denotes the substrate and the third number denotes the acceptor; whereas for class 3, the second number denotes the type of bond cleaved and the third number denotes the molecule in which that bond is embedded. For all classes, the fourth number signifies some specific feature such as a particular cofactor.

[4] These correlations are crucial for understanding the links between genome and epigenetics.

J.J. Ramsden, *Bioinformatics,* Computational Biology 10,
DOI 10.1007/978-1-84800-257-9_16,

Metabonomics usually includes not only intracellular molecules but also the components of extracellular biofluids. Of course, many such molecules have been analysed in clinical practice for centuries; the novelty of metabonomics lies above all in the vast increase of the scale of analysis; high-throughput techniques allow large numbers (hundreds) of metabolites to be analysed simultaneously and repeat measurements can be carried out in rapid succession, enabling the temporal evolution of physiological states to be monitored. The concentrations of a fairly small number of metabolites has been shown in many cases to be so well correlated with a pathological state of the organism that these metabolite concentrations could well serve as the essential variables of the organism, whose physiology is, as we may recall, primarily directed toward maintaining the essential variables within viable limits.

Metabonomics is being integrated with genomics and proteomics in order to create a new systems biology, fully cognizant of the intense interrelationships of genome, proteome, and metabolome; for example, ingestion of a toxin may trigger expression of a certain gene, which is enzymatically involved in a metabolic pathway, thereby changing it, and those changes may, in turn, influence other proteins, and hence (if some of those proteins are transcription factors or cofactors) gene expression.

16.1 Data Collection

The basic principle is the same as in genomics and proteomics: separation of the components followed by their identification. Unlike genomics and transcriptomics, metabonomics has to deal with a diverse set of metabolites even more varied than proteins (which are at least all polypeptides). Typical approaches are to use chromatography to separate the components one is interested in and mass spectrometry to identify them. Alternatively, high-resolution nuclear magnetic resonance spectroscopy can be applied directly to many biofluids and even organ or tissue samples.

Metabolic microarrays operate on the same principle as other kinds of microarrays (§14.1) in which large numbers of small molecules are synthesized, typically using combinatorial or other chemistry for generating high diversity. The array is then exposed to the target, whose components of interest are usually labelled (although their chemical diversity makes this more difficult than in the case of nucleic acids, for example; moreover, the small size of metabolites makes it more likely that the label chemically perturbs them). This technique can be used to answer questions such as "to which metabolite(s) does macromolecule X bind?"

Much ingenuity is currently being applied to determine spatial variations in selected metabolites. An example of a method developed for that purpose is PEBBLES (Probes Encapsulated By Biologically Localized Embedding): fluorescent dyes, entrapped inside larger cage molecules, and which respond (i.e., change their fluorescence) to certain ions or molecules. Their spatial location in the cell can be mapped using fluorescence microscopy. Another example is the development of