Introduction
The group has a growing interest in data from post genomic research, including microarray based measurements of gene expression, and, more recently, tissue microarrays. Work is collaborative with local experimental groups who generate data, and we are responsible for three aspects of these projects: (i) appropriate storage and archiving of data according to international standards, and efforts to advance these standards; (ii) data analysis using methods of multivariate statistics; and (iii) the use of private and public data to make predictions and motivate experimental verification or refutation.

Human cancer pathology
Tissue samples are obtained from individuals in a clinical trial (both cancerous and normal tissues from the same person) for biochemical and histopathological analyses. Depending on the nature of the trial, these tissue samples undergo extensive characterisation using a high-throughput molecular biology techniques. The high-throughput techniques most commonly used in cancer research are cDNA microarrays, comparative genomic hybridisation arrays and tissue microarrays. The purpose of the cDNA microarray approach is to gain an insight into the expression levels of all the predicted genes in the human genome with the aim of identifying a set of genes related to a clinical outcome that may be either up or down regulated in tumour verses normal tissue. Comparative genomic hybridisation (CGH-arrays) arrays are used to study chromosomal instability at a genome level within tumour verses normal tissues. TMA is a technique that enables the analysis of a large cohort of clinical specimens in a single experiment thereby studying the molecular alterations (at the DNA, RNA, or protein level) in thousands of tissue specimens in parallel. The aim of cDNA microarray and CGH-array techniques are to either identify biomarkers that can be verified by TMA. Our involvement in this research involves analysis of the cDNA microarray and CGH-array data using statistical approaches, and the development of storage and analysis software for tissue microarray experiments, an area where we are contributing to the development of international standards, and the integration of this data with MIAME compliant microarray databases.

TMAs are used in the laboratory to assess on a large-scale the diagnostic and therapeutic significance of various genes and proteins in colorectal tumour samples. A relational database has been designed and implemented in MySQL. The information stored in the database include TMA design constructs, tissue staining protocols, the results including images scanned from digital slide scanners and the pathology reports associated with each tumour sample. Additional information includes experiment authors, dates of each experiment, quality of cores on each TMA slide and the storage location of each TMA within the laboratory. This database is interfaced with the World Wide Web (WWW) thereby enabling users to query and assimilate their own data into the database.

Collaborators: Professor P. Quirke

Publications