Strategies and technologies in functional genomics

Although a massive quantity of sequence information is accumulating, the functions of thousands of genes remain undetermined. Functional genomics refers to the methods for assessment of gene function by making use of the information provided by structural genomics. Pharmacogenomics, an offshoot of genomics, refers to the application of genomic technologies in drug discovery and development.

Highlights of the important presentations given at a recent IBC conference (*Functional Genomics: Strategies and Enabling Technologies for Target and Drug Discovery*) held in London, UK, on 4–6 November 1998 are reviewed here, and a more comprehensive review of this subject is published as a special report [Jain, K.K. (1999) *Genomics in Drug Discovery and Development,* Decision Resources, Waltham, MA, USA].

The business of genomics

Progress of the Human Genome Mapping Project (HGMP) was reviewed by Chris Mundy (HGMP Resource Centre, Cambridge, UK) who pointed out that the year 2003 (and not the earlier goal of 2005) was a firm target for a complete 'high quality' human sequence. There are commercial implications of the 'shotgun' sequencing of the human genome as planned by Celera Genomics (Rockville, MD, USA – a joint venture of The Institute for Genomic Research and Perkin-Elmer) for the year 2001. Single nucleotide polymorphisms (SNPs) define most changes in the coding sequence and have functional significance. Mapping of expressed sequence tags (ESTs) in relation to previously mapped molecular markers has proved very useful in providing candidate genes as well as orientation points for genomic sequence. The relevance of genomics to healthcare products are:

- Association of gene products with disease
- Validation of gene products as effective targets for drugs and diagnostics
- Knowledge of SNPs that may influence drug development

Klaus Leadpaintner (Roche Genetics, town, country) reviewed the gene to drug paradigm and described the mission of Roche Genetics which will be to:

- Support the individualized healthcare system
- Integrate the expertise, resources, tools in genetics, genomics and proteomics across pharmaceuticals, diagnostics, drug development and drug delivery
- Develop new concepts in healthcare
- Educate both the public and the medical profession
- Manage the ethical, social and legal issues associated with genomics

Two examples of activities of Roche in these areas are collaboration with deCode Genetics (Rejkavik, Iceland) and integration of HIV diagnostics and therapeutics.

Claire Allan (Stevenage, UK) described the approach of Glaxo Wellcome to link genomics and medical genetics with pharmacogenomics and pharmacogenetics (influence of genetic factors on the action of drugs). The objectives of Glaxo Wellcome Genetics Directorate are to identify susceptibility genes for common diseases with the largest unmet needs, to assist in translating gene discoveries into target selection and to apply genetic methods to the development of the right medicines for the right patients.

There are several important technologies; gene-expression profiling is useful for finding disease genes as potential therapeutic targets, clinical surrogates for diagnostics and monitoring drug efficacy, and as markers for toxicity screening. Among transcriptional assay technologies, closed DNA array enables comparison of thousands of genes monitored in parallel. Data can be accumulated and directly compared over series of experiments. Genomic informatics is also an important element of drug discovery. The tool developed for *in silico* cloning at Glaxo Wellcome is ESTBlast, which has been released for use by the bioinformatics community. The aim is identification of susceptibility genes for common diseases, but this requires large population studies and many external networks, collaborators and clinical trials.

Beyond sequencing, genomics-based drug development will be based on functional studies of genes. Arthur Sands (Lexicon Genetics, Woodlands, TX, USA) described mouse functional genomics and OmniBank – a high-throughput analyser of mutant mice using gene-trap technology. A homologous recombination programme is used for design of 'custom' mice, and phenotypic analysis is used for discovering and validating drug targets. Lexicon aim to create a library of 500,000 expressed sequence (ES) clones, and the company offers collaborative opportunities in phenotype analysis through a programme to screen and study OmniBank mice in various disease categories.

Valuation of market opportunities in functional genomics

Application of functional genomics is anticipated to expand market opportunities, shorten the drug development cycle and improve drug approval rate. It

1359-6446/99/$ – see front matter PII: S1359-6446(98)01295-1

is, however, difficult to evaluate the marketing opportunities in functional genomics, but the net present value (NPV) approach was discussed by Enal Razvi (LJL BioSystems, Sunnyvale, CA, USA). In this approach, the total size of a given biotech space is estimated in terms of future marketing opportunity (in today's dollars) and then the total 'captured' fraction of this opportunity is computed by estimating the 'available' opportunity for entry into a given space. The discount rate reflects the inherent risks in genomic technologies and quality of the specific partner. The time period is dependent on the level of commitment to the project, the current status of the molecular knowledge of a disease and the therapeutic limitations. Market opportunity depends on the nature of the disease pursued, its severity and the patient population affected. The NPV in this context reveals the value of a gene within given parameters. By comparing the NPV for a genomic space with the total value of the partnership, it can tell us whether a strategic collaboration in genomics is correctly valued, over-valued or under-valued, and the difference between the NPV and the value of genomic partnership reflects the competitiveness of given disease class.

Model genetic systems and target validation

Mouse is considered to be the best genetic model organism for human biology because the biological function of any gene can be studied in pre-designed, targeted mouse mutants. The application of transgenic mouse technologies to functional genomics and target identification was discribed by Satbir Kaur (Chrysalis DNX Transgenic Sciences, Princeton, NJ, USA). Gene addition can be carried out by microinjection so that gene function can be evaluated or gene modification performed by homologous recombination. Systematic mutagenesis in the mouse enables generations of animal models that mimic human disease and facilitate identification of disease-related drug targets. Therapeutic targets can be validated by determining the gene function, linking it to a defined pathway and by testing the mechanism of action.

Systematic genetic screens have been used in the zebra fish (*Brachydanio rerio*) to identify novel therapeutic proteins and drug screening targets as described by Ralf Kuehn (Artemis Pharmaceuticals, Cologne, Germany). Inducible knockouts in the mouse can be used to validate genes thus identified. For example, an inducible knockout mouse for a zebra fish gene regulating cartilage formation will develop excess cartilage. Inducible inactivation of the target gene mimics the inhibition of the gene product by the drug. A drug inhibiting the target gene product may be used for cartilage regeneration in man.

Gordon Smith Baxter (Pharmagene, Royston, UK) presented new approaches to human therapeutics through systemic evaluation of the human phenotype. Technologies for data synthesis involve a triad of genomics (mRNA dynamics), pharmacology (tissue responses to drugs) and biochemistry (protein dynamics and signal transduction). 'Level zero' data is obtained from a variety of tissues from donors of diverse background. These data are converted into knowledge in eight levels or more leading to drug lead optimization. For focused evaluation of mRNA expression, the company uses ABI 7700 sequence-detection technology, and other technologies are used to generate 'pathfinder' data and to develop leads.

Determining gene function

Transcript profiling helps in identification of new drug targets and toxicology profiles. Serial analysis of gene expression (SAGE) was presented by Greg Landes (Genzyme Molecular Oncology, Framingham, MA, USA). This technique is a high-throughput, high-efficiency method for simultaneous analysis of thousands of genes at one time and is based on the following two principles:

- Short bits of genetic information (nucleotide sequence tag of 9–10 base pairs) are isolated from the expressed genes in the cells being studied. The SAGE tags are linked together for efficient sequencing. These contain sufficient information for identifying 95% of the genes, provided it is isolated from a defined position within the transcript.
- Serial analysis of sequence data is carried out by computer software to identify each gene expressed in the cell and the level at which the gene is expressed. The information forms a library that can be used to analyse the differences in gene expression between cells.

The potential uses of SAGE in drug discovery are to:

- Identify novel genes for use as drug discovery targets
- Analyse the effects of drugs on tissues and provision of information about disease pathways
- Identify the differences between normal and diseased tissue, particularly between non-cancerous and cancerous tissue. SAGE differences translate directly into RNA differences as assessed by northern blot analysis and alterations identified in a few samples are consistent with data from a larger sample of primary tumour isolates [Bertelsen, A.H. and Velculescu, V.E. (1998) *Drug Discovery Today* 3, 152–159].

A clue to the function of a gene can be obtained by studying the effects of its absence. As an alternative to targeted mutagenesis, antisense oligonucleotides can interrupt the flow of information from the gene to the protein by 'removing'

the mRNA. The use of oligonucleotide arrays and DNA chips for this purpose was described by Kalim Mir (University of Oxford, UK). These can be used for gene expression analysis, genotyping and sequencing by hybridization [Case-Green, S.C. *et al.* (1998) *Curr. Opin. Chem. Biol.* 2, 404–410].

RNA splicing is an essential regulatory mechanism that controls gene expression and nearly 10% of all inherited human genetic mutations affect RNA splicing. Laurent Bracco (Exonhit Therapeutics, Paris, France) described how DATAS technology used for screening mRNA spliced isoforms gives access to qualitative splice-related events that can be missed by existing gene profiling techniques. It has potential application in target identification, toxicology and diagnostics.

It is well recognized that inherited differences between individuals underlie the variable responses to drug treatment. Colin Dykes (Variagenics, Cambridge, MA, USA) described application of genotyping in clinical trials. Variagenics uses several technologies for scanning of genetic variance, including Variance Scanning™, Variance Imaging™ and Variance Typing™. Understanding the variance in the pathways affecting drug action enables safer and more effective therapies, and can be used for increasing market shares and profits.

Microarrays and chips

Array technology has been applied successfully in the miniaturization and automation of many genomic technologies. Sebastian Meier-Ewert (Genome Pharmaceuticals Corporation, Munich, Germany) presented a gene expression technology called Oligofingerprinting, which is based on sequential hybridization of synthetic oligonucleotides to arrayed DNA libraries. The steps involved are library construction, PCR amplification, automated array technology, hybridization with oligonucleotides and capturing of the patterns (fingerprints). To complement this technology, the authors have developed an approach using target micro-arrays in hybridization with total mRNA isolated from the tissues of interest [Maier, E. *et al.* (1997) *Drug Discovery Today* 2, 315–324]. This enables a rapid tissue/sample profiling of potential targets.

Proteomics

A proteome is a set of functional proteins encoded by a gene and is the protein complement of the genome. Proteomics is the systematic analysis of protein profiles of tissues and parallels the related field of genomics. The use of mass spectrometry as a tool to elucidate the structure and function of gene products (proteins) was presented by Darryl Pappin (Imperial Cancer Research Fund, London, UK). Robert Burns (Oxford GlycoSciences, UK) described the integrated technology for proteomics of his company, which starts with control sample and proceeds to 2D-gel, digital imaging, Proteographs™ and ROSETTA™ bioinformatics systems to characterize and analyse protein sequences that are correlated with disease pathways (see also Page, M.J. *et al.*, p. 55). Protein–protein interactions can be used for determining gene function.

Selective infective phage (SIP) technology was described by Corinna Loehning from Morphosys (Munich, Germany). SIP enables efficient screening of large collections of substances that bind tightly to a particular target. An extension of SIP focuses on protein–protein interactions and can reveal how proteins in signalling pathways interact with each other. This technology is used to determine the function of the drug targets, such as cell-surface receptors, signalling molecules and related proteins [Spada, S. and Pluckthun, A. (1997) *Nat. Med.* 3, 694–696].

Bioinformatics

Bioinformatics, with its use of highly sophisticated computer databases to store and analyse biological information, is playing a vital role in genomics and proteomics. Stephan Bryant (National Center for Biotechnology Information, NIH, Bethesda, MD, USA) described protein structural analysis using the Cn3D tool. This enables the users to view protein and nucleic acid structures from the Molecular Modeling Database. Rolf Apweiler (European Bioinformatics Institute, Cambridge, UK) described the SWISS-PROT, a protein sequence database that provides, among other criteria, a description of the function of proteins, their domain structures, post-translational modifications and variations. It has 76,000 entries abstracted from >60,000 references, and TrEMBL – a bioinformatic supplement to SWISS-PROT – speeds up the sequence analysis, information analysis and data integration. Further software is in development that will accelerate the annotation of sequences by automating and combining, for instance, similarity searches and special sequences analysis tools. The InterPro project plays a central role in these developments.

Keith Steward (GeneLogic, Columbia, MD, USA) presented the automation of discovery and data mining for industrial genomics. The company uses gene Flow-thru Gene Chip technology for high-throughput screening and bioinformatics designed to generate and analyse functional gene data. They use BioWizard™, an 'intelligent agent' that automatically contacts and seeks out information from scientists and automates retrieval of candidate gene sequences from databases. It automates the running of BLAST (Basic Local Alignment Search Tool), interprets the results and sends the reports to appropriate scientists; it may possibly be licensed to other companies in the future.

Concluding remarks

There is tremendous activity in functional genomics for drug targeting and

discovery. The major activity is in the industrial sector where >200 companies are involved, and the conference provided a forum for exchange of information between the industry and academia. Of the numerous genomic technologies, only a handful can be presented at a conference, but the conference was successful in providing a fair sampling of the state-of-art in genomics. Sequencing of the human genome is expected to be complete early in the next century. In the post-genomic era, drug discovery and development will be based on functional genomics and this will have an impact on healthcare in the form of reclassification of diseases, personalized rational medicines and the combination of therapeutics with diagnostics.

Kewal K. Jain
Bläsiring 7
CH-4057 Basel, Switzerland
fax: +41 61 692 4461

Paralysis prevention drug

An anticancer compound that blocks blood vessel formation could soon be used to help patients with paralysing spinal injuries to regain feeling and even to walk again.

In a paper dedicated to actor Christopher Reeve who was paralysed in a riding accident, Carl Hellerqvist of the Department of Biochemistry at Vanderbilt University (Nashville, TN, USA) describes how mice paralysed by spinal cord injury (SCI) and then treated with the angiogenesis inhibitor CM101 – a polysaccharide-type molecule composed of GalNAc/GlcNac/Gal/Gly – showed a dramatic recovery in walking ability within 2–12 days [*Proc. Natl. Acad. Sci. U. S. A.* (1998) 95, 13188–13193].

Natural anticancer agent…

CM101 (produced by Carbomed, Brentwood, TN, USA) is an exotoxin produced by group B *Streptococcus* bacteria. The molecule has the ability to disrupt angiogenesis and as such it has potential as a therapy in several disease conditions. It is currently being investigated as an anticancer agent. When the supply of nutrients from the blood is cut off, a tumour shrinks and dies, and in tests with mice, angiogenesis inhibitors have indeed been shown to stop tumour growth.

…Applied to paralysis

Hellerqvist and his team, in work supported by Carbomed, figured that an angiogenesis inhibitor such as CM101 might also prevent the growth of scar tissue (gliosis), which also requires new blood vessel growth, following SCI. It is the growing scar tissue that blocks the path of severed nerve cells as they try to reconnect, so halting scar formation might allow damaged nerves to repair themselves.

The team tested their theory by using intravenous administration of CM101 to treat mice with hind limbs paralysed as a result of SCI. They found that 24 out of 26 mice treated with the drug survived at least 28 days, but importantly all of them recovered their ability to walk. By contrast, in the control group, only three of 14 untreated mice survived for the same period, and none of them regained limb function.

Magnetic resonance imaging and microscopy of spinal-cord samples from the mice showed a significantly reduced area of scarring in the treated mice, according to the researchers. The team also found that CM101 helps restore nerve cell connectivity not only by preventing scarring, but also by protecting damaged nerve cells from degeneration and death. Electrophysiological measurements on isolated CNS and neurones in culture using intracellular microelectrode recordings showed that CM101 protected axons from degeneration and reversed γ-aminobutyrate (GABA)-mediated depolarization of traumatized neurones. An improvement in the recovery of neuronal conductivity of isolated CNS in culture was also observed by the team. 'Restored conductivity would be the equivalent to recovery of walking in adult animals,' Hellerqvist says.

Major clinical challenge

Recovery from SCI in people is 'a major clinical challenge', says Hellerqvist. SCI causes immediate mechanical damage and subsequent tissue degeneration due to ischaemia, haemorrhaging and oedema which all lead to nerve cell death. However, various studies have shown that CNS neurones are capable of regrowth along their length, which means, under certain conditions, severed nerves can regrow – surgical methods are being investigated to open up this possibility.

Interestingly, the team also noted that the mice treated with CM101 in this study experienced rapid healing of their wounds from the surgical incisions and without scar formation. This additional discovery could be useful in controlling scar tissue in facial injuries to lessen the long-term cosmetic impact of a wound.

David Bradley
tel/fax: +44 1954 202218
Web: http://www.camsoft.com/elemental/

 1359-6446/99/$ – see front matter PII: S1359-6446(98)01297-5