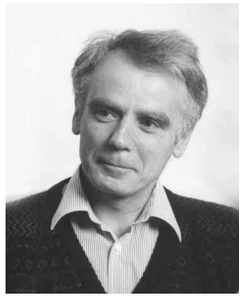
The Human Genome Project – obstacles versus opportunities



'Truly preventative medicine beckons… Unfortunately… the new genetics… is on a collision course with public suspicion.'

As the first public draft sequence of the human genome nears completion, it is a good time to reflect on how well prepared we are for the accompanying tidal wave of information and opportunity. New technical and commercial perspectives have already been widely adopted. The 'thorny' issue of public acceptance of technological innovation, which has so bedevilled the genetically modified (GM) food sector, also has significant implications for the pharmaceutical industry. In this article, I will try to draw together both the technical issues and the need, as I see it, for the industry to take the leading role in building public confidence. I will start by describing the projected outcomes of the Human Genome Project, and the forces that have driven it towards completion at an ever-accelerating rate.

Outcomes and opportunities

The human genome, containing perhaps up to 140,000 genes, is an extensive pool of potential therapeutic targets. In 1992, there was a general consensus that the 3% of the human genome that encoded proteins was the best goal for large-scale DNA sequencing. The other 97% of the genome was thought to be so-called 'junk DNA'. Private and public efforts, both funded in part by pharmaceutical companies, have yielded very large databases of expressed sequence tags (ESTs). Sequences within the public database Genbank, gained mostly from IMAGE libraries, have been clustered to identify up to 60% of human genes. The rest will be revealed by the complete genomic sequence.

Under the threat of Intellectual Property Rights (IPR) barriers, there has been an astounding acceleration in the pace of public genomic sequencing, funded principally by the US government and the UK Wellcome Trust. Just over a year ago, the goal for completion was 2005. Now the goal is 2003, with a first draft, comprising 99% sequence accuracy for 95% of the human genome, publicly available from Spring 2000. One of the major goals is the identification of single nucleotide polymorphisms (SNPs). The public domain SNP Consortium has set a target of 300,000 SNPs within the next two years, of a total of perhaps 3 million. These SNPs will be used to identify disease-associated genes, and eventually indicators of pharmacological effectiveness. Biologists and developers of healthcare products will be engaged for much of the new century in assigning function and interactions to gene products predicted from the new sequence data, and developing products based on that knowledge.

Mining data for gold in the 21st century

The main challenge for scientists is the distillation of knowledge from data, this being especially true for the Human Genome Project. A new generation of computational tools will be required, not only to make use of the ever-expanding data set, but also to enable modelling of complex biological systems and accurate prediction of cause and effect. Hypotheses will be based on simulation in virtual organisms, and confirmed by experiment. This new biological knowledge will join the now 'not-so-new' physics in a virtual world.

Association of gene and function by genetic mapping and mutation analysis, and by gene expression studies, enables prediction of the relevance and value of novel genes. Mutations in some genes contribute to the probability of contracting most of the common causes of death and morbidity in the Western world (such as heart disease, cancer, diabetes and Alzheimer's disease), as well as to the susceptibility to infectious disease. Such mutations are often taken as *a priori* evidence that a particular gene is involved in a given disease process, and such genes become candidate targets for drug development. Similarly, altered expression of a gene in diseased cells is used to infer relevance of that gene.

Chris R. Mundy, Biology Services Manager, UK HGMP Resource Centre, Hinxton, Cambridge, UK CB10 1SB. tel: +44 1223 494515, fax: +44 1223 494512, e-mail: cmundy@hgmp.mrc.ac.uk, URL: http://www.hgmp.mrc.ac.uk

 1359-6446/00/$ – see front matter PII: S1359-6446(00)01503-8

Whatever the limitations of these inferences, focussed biological studies (often in model systems) can enable validation of that new 'seam'. After establishing a basis for value of the gene, the 'seam' can be extended. Sequence similarity analyses, including amino acid class homologies at the level of the protein sequence, enables extension of useful gene families. The seven-transmembrane receptors are a good example. Approaches such as clustering of expression profiles and yeast two-hybrid analyses enable the building of networks of molecules that interact or act in concert. Extensive data sets, particularly in human tissues, can be of great value, although the collation of all these different types of data and their effective use in modelling will require massive investment in software development.

The marriage of diagnostics and therapeutics

How should the pharmaceutical industry respond to the challenge and opportunity of human genetic variation? This is perhaps the major question facing the industry at the beginning of the new millennium. There are the obvious advantages of a modest and measured response, for example to determine polymorphisms in some cytochrome P450 genes responsible for drug metabolism, to ensure that patients are protected from adverse reactions. These analyses can be closely targeted and are unlikely to be particularly controversial.

However, there have been well-publicized examples of treatments that are effective for only a proportion of the patient population, and are actually dangerous for others. Many potentially useful and valuable drugs, which have been expensive to develop, could be rescued if they could be focussed on a responding sub-population. Financial models have been developed for targeted therapies based on an increased compliance, which highlights both the benefits to the industry and to the public. Preventative medicine is particularly important for reducing healthcare costs in the context of increasing intervention, and genetic profiling will be crucial if preventative therapies are to be developed and applied. This is truly a great opportunity for all of us.

It will take many years to unravel the limits of drug effectiveness resulting from human polymorphisms and the interventions that could be possible to prevent disease in different people. The tools currently available only enable examination of a very limited number of possibilities. However, the first steps are being taken. The technical barriers will fall eventually, but unless social issues are addressed, the effort will have been wasted and the opportunity to increase vastly the cost-effectiveness of healthcare will have been lost.

Whose genome is it anyway?

The GM food debacle has provided an object lesson for which we should be grateful. For issues such as *in vitro* fertilization (IVF), it has worked quite well for learned bodies to debate the issues, air their views on television, and make recommendations for legislation. That is because relatively few people are affected directly by the outcomes. When that does not apply, and the same rather patronizing approach is used, there is an uprising of strong opposition. Broad public cooperation necessitates broad consensus.

There are genuine dilemmas for the person in the street. The same information that might enable him or her to have effective preventative treatment for cancer or heart disease, could enable an employer or an insurance company to reject them. Of course in the end, the person, the insurance company and the employer will all benefit, but there is first a PR 'mountain' that the pharmaceutical industry must climb. Furthermore, each person regards his or her genetic make-up as deeply personal. The industry must be seen to be focussed on consumer needs and the debate must be broad, involving all of society, and not evading the key issues of civil rights.

Conclusions

The global public domain Human Genome Project is about to deliver its promises more than a decade ahead of the first expectation, and costing certainly no more than 50% of the original budget. Large-scale initiatives in functional genomics are poised to create a knowledge base of ever-greater depth and utility, aided by developments in bioinformatics. Pharmaceutical companies have contributed significantly, and much of the industry is positioning itself to turn this data and knowledge into products to benefit society. Truly preventative medicine beckons, promising huge benefits to all. Unfortunately, however, the complexity of the new genetics, in the context of the habitually closed culture of the industry, is on a collision course with public suspicion. There is a grave risk that many of the healthcare benefits of the Human Genome Project will be undeliverable. A much more extensive dialogue is required, addressing the real and reasonable fears of the public on the one hand, and the real and valuable benefits of the new genetics on the other. Time is running out. Who is going to take responsibility for driving the debate?

Chris R. Mundy

Have you seen the following related articles, published in *Drug Discovery Today*?

- Lennon, G.G. (2000) High-throughput gene expression analysis for drug discovery. *Drug Discovery Today* 5, 59–66
- Searls, D.B. *et al.* (2000) Using bioinformatics in gene and drug discovery. *Drug Discovery Today* 5, 135–144
- Lenz, G.R. *et al.* (2000) Chemical ligands, genomics and