HUMAN GENOME RESEARCH

POST 142

Report Summary June 2000

On June 26th 2000, scientists in the UK and US announced the forthcoming publication of a 'first draft' of the sequence of the human genome. While this is a significant milestone, it is by no means the end of the quest to crack the secrets of the human genetic code. Indeed, it is perhaps best thought of as the 'end of the beginning' of such endeavours. In the coming years the knowledge gleaned from such research will have profound implications for the way doctors classify, diagnose treat and prevent disease.

This note is a summary of a POST report[1] that outlines the main developments, examines the implications and discusses the issues that arise.

SEQUENCING THE HUMAN GENOME

DNA is made up of 4 building blocks, the bases A, C, G and T (**Box 1**). The genetic code that gives each person his or her unique characteristics is defined by the sequence of these bases along the DNA molecule. Even the best current techniques for reading DNA sequences are capable of decoding only relatively short stretches at a time. This means that a '**shotgun**' approach must be used, where the long lengths of DNA that make up chromosomes (Box 1) are broken down into large numbers of much shorter fragments. The problem facing researchers is to deduce the correct order of these fragments. Two approaches have evolved: a publicly funded Human Genome Project (HGP) and a privately funded approach.

The Publicly Funded HGP

The HGP approach involved first mapping 'land-marks' at intervals along each of the chromosomes. Successively smaller 'libraries' of overlapping fragments were generated, and the sequence read in machines. The detailed maps assist reassembling the sequences into the correct order. Large-scale sequencing efforts are divided between:

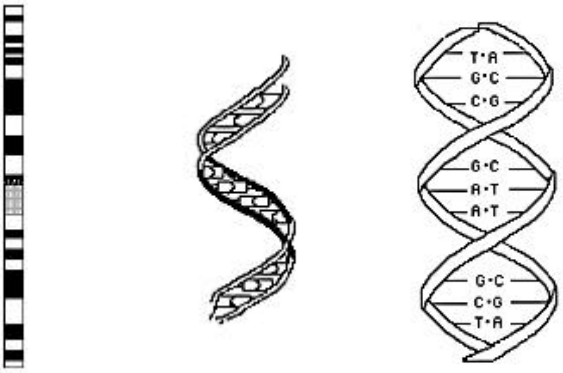
- Sanger Centre, in Cambridge, UK;
- Washington University Genome Sequencing Center in St. Louis, US;
- Whitehead Institute in Cambridge, US;
- Baylor College of Medicine in Houston, US;
- and the US Joint Genome Institute.

UK involvement is through the Sanger Centre, jointly set up and funded by the Medical Research Council (MRC) and the Wellcome Trust. It is funded to sequence around one third of the 3 billion bases in the human genome. All laboratories collaborating in the HGP deposit finished sequences in a public database within 24 hours of obtaining them.

BOX 1 AN INTRODUCTION TO THE HUMAN GENOME

Human Cell 23 pairs of chromosomes

The human body consists of billions of cells. Each of these contains a copy of the human genome within the cell nucleus. It consists of 23 pairs of chromosomes; 22 matching pairs and 1 pair of sex chromosomes (XX in females, XY in males).



Chromosome Gene Base sequence

Each chromosome is a long chain of DNA wrapped in protein containing discrete sites (genes). Each gene (there are up to 140,000 spread among the 24 different chromosomes) codes for a single product (protein). The code is carried by the sequence of the 4 nucleotide bases that make up DNA (ACG&T). This base sequence determines the order of building blocks (amino acids) in the protein coded for by the gene. In all, the human genome consists of some 3 billion bases, only about 5% of which are gene sequences.

Privately Funded Genome Project

An American company (Celera) is employing 'whole genome shotgun sequencing' to obtain the sequence of the human genome. This was originally developed to obtain the sequence of smaller, less complex, bacterial genomes. It effectively involves sequencing millions of overlapping fragments 'blind', relying on supercomputers to reassemble the fragments in the correct order without recourse to the painstaking mapping used in the HGP. Celera started full-scale human genome sequencing in September 1999. By April 2000, the company announced it had completed the sequencing phase of one individual's genome and had started the computational assembly of the sequence.

Celera plans to publish its sequence in full, but will first seek patents on any 'medically interesting' genes; it has already made provisional patent applications on some 6,500 gene sequences. Other

[1] **Human Genome Research, POST, June 2000.**

companies seeking to 'stake a claim' to parts of the human genome include Incyte (which has already filed patents on some 1.3 million partial gene sequences) and Human Genome Services Inc. (which has applied for patents on some 7,400 genes).

THE ANNOUNCEMENT

Among the key features of the June 26th announcement were:

- 85% of the human genome has been sequenced as a first draft.
- Converting the draft sequence into finished sequence could take up to 2 years (~23% of the human genome is currently available in finished form). This will involve refining the draft and filling the gaps in areas of the chromosomes that are difficult to sequence.
- The draft sequence includes virtually all the 'interesting bits' – the genes that code for the proteins that control biological processes.
- It was a simultaneous announcement from the HGP and Celera – the 'race' between the public and private sectors was agreed to be a draw.

WHAT HAPPENS NEXT?

As outlined in more detail in the full POST report, the human genome sequence will greatly increase understanding of the role genes play in fundamental biological processes. But obtaining the sequence is by no means the end of the story. Indeed, it is merely the first step towards understanding how genes work, their complex interactions with each other and with the environment, their role in diseases, etc. Much more research will be needed to achieve such understanding. Among the underlying technologies and research areas examined in the full report are:

- Bioinformatics – using information technology to store, retrieve and process the huge quantities of genome data.
- Functional genomics – a name applied to studies aimed at understanding all aspects of genome function.
- Comparative genomics – identifying the functions of human sequences by comparison with sequences of known function from other species.
- Population studies - linking genetic data to patterns of disease/environmental factors. In the UK, the Wellcome Trust and MRC are exploring the possibility of collecting data on up to 500,000 people for a period of several years (perhaps even decades).
- Variations in sequence between individuals (polymorphisms) which may help explain why people are more or less susceptible to certain diseases, why some drugs work better for some people than for others, etc.
- Proteomics – the study of all the proteins produced in a cell. Matching this with genetic data reveals information on the regulation of genes, the roles played by proteins in disease states, etc.

GENES AND DISEASE

Genome research will greatly increase knowledge of the role of genetics in a wide range of diseases. In the first instance, the main 'product' of a better understanding of the genetic basis of disease is likely to be improved diagnostic tests. Among the main implications discussed in the full report are;

- Diagnosis and classification of disease will increasingly be based on genetic information rather than a somewhat crude assessment of physical signs and symptoms.
- This will not only embrace the 'classic' genetic diseases – i.e. single gene disorders such as cystic fibrosis[2], caused by a fault in a single gene.
- It will also encompass many of the common 'killer' diseases such as cancer, diabetes and heart disease, where genetics is only one of the factors involved.
- Diagnosis of other diseases may also be affected; for instance, there is evidence that a person's susceptibility to some infectious diseases may be at least partly determined by genetic factors.

GENES AND TREATMENTS

Improved understanding of how genes and the proteins they code for are implicated in disease will also allow the development of better treatments (although this will take at least 10 years of further research). The report discusses likely developments in three main areas:

- Gene therapy – treatments involving transfer of beneficial genetic material to human cells.
- Pharmacogenetics – genetic information about how rapidly a person breaks down a drug should allow existing drugs to be used more effectively. For instance, doctors will increasingly be able to assess whether a person will react badly to a particular drug, the most appropriate dose, and whether the drug will be efficacious.
- New drugs – improved understanding of disease mechanisms, basic cell processes, etc. will provide new targets for the development of novel treatments.

[2] See POSTnote 124, 'Cystic Fibrosis', March 1999

ISSUES

Intellectual Property Rights (IPR)

Disputes over IPR have been an on-going feature of the HGP since its inception. The debate originally focused on whether patents should be awarded on gene sequences at all. In practice however, agencies throughout the world have been granting patents on inventions involving human gene sequences for more than a decade. Debate has thus shifted towards a more detailed examination of the technical aspects of the patenting process itself. Among the issues discussed in the main report are:

- Usefulness – there is now a consensus that simply disclosing the sequence of a gene and knowing its function is not enough. Researchers must also submit evidence demonstrating an application (e.g. therapeutic or diagnostic) for this knowledge.
- Scope - applicants may seek to maximise their patent portfolio by claiming the widest possible rights for their invention. This means that claims are often framed in very broad terms – for instance claiming that the gene sequence can be used for therapeutic and/or diagnostic purposes in humans and other species. If granted, such claims effectively award the patent holder a monopoly on all possible future uses of the gene sequence in question.

The issues of scope and usefulness raise the question of at what stage in the process it is most appropriate to patent. There is near universal agreement that raw sequence data should not be patentable *per se*. This position was reiterated by the recent joint statement[3] by President Clinton and Prime Minister Blair, which stated that: "*We applaud the decision by scientists working on the Human Genome Project to release raw fundamental information about the human DNA sequence and its variants rapidly into the public domain, and we commend other scientists around the world to adopt this policy*".

On the other hand, there is also a consensus that patents have a role to play at some stage in the development of gene based inventions. For instance, the joint statement also noted that:"*Intellectual property protection for gene-based inventions will also play an important role in stimulating the development of important new healthcare products*".

In practice however, many IPR issues will only be resolved in time, as a jurisprudence emerges from the application of patent law and challenges to it through the courts.

[3] White House Press Release, March 14th 2000

Impact on the NHS

The report discusses a number of reasons for assuming that genome research will have an increasing impact on the NHS in the next few years. Key factors driving trends will be:

- The development of new genetic tests. In the first instance these will largely be new tests for single gene disorders. But as genome research informs a shift towards the use of genetic information in classifying and staging disease, so the demand for tests for a range of common complex disorders will rise.
- New technology - the development of 'gene chip' technology promises to deliver faster, cheaper, more comprehensive and accurate genetic tests.
- New treatments – although these will take longer to materialise, the identification of new drug targets will lead to the development of more effective drugs.
- Personalisation of medicine – using gene tests to target drugs at those who will benefit from them.
- Public demand – media influence could fuel (unrealistic) expectations among the general public over what the new genetics can deliver.

Such factors are likely to lead to new genome-derived tests and treatments for a wide range of conditions, with implications for the NHS, medical profession and policy makers alike. Among the main issues dealt with in the report are:

- Organisation of genetic services. Current services are geared at providing tests for rare inherited disorders and are thus organised on a regional basis. However, the development of new, genetic tests for diagnosing/assessing susceptibility to common diseases would require the provision of services at the primary care level.
- Counselling – specialists advice is offered to individuals/families affected by or identified as being at increased risk of a genetic disorder, with the aim of allowing fully informed choices. The report discusses the likely implications for current counselling services of the development of new types of (e.g. susceptibility) tests for common diseases for use in primary care.
- Education and training – relatively few doctors in current practice have received much formal training in clinical genetics. This will need to be addressed through undergraduate education and continued professional training.
- Uptake of products by the NHS. The report discusses whether there is a need for more research to establish better methods for assessing the health benefits and cost-effectiveness of genome derived tests and treatments.

Uses of Genetic Tests

Genome research will lead to new genetic tests being developed for a wide range of purposes (**Figure 1**). These can be split into two main categories:

- Disease genetics – these include diagnostic tests for rare inherited disorders and diagnostic and susceptibility tests for common diseases.
- Pharmacogenetics – tests to assess whether an individual will react adversely to a drug, whether the drug will be efficacious and what dose is appropriate.

Such tests can be used in a variety of ways, each of which raises slightly different issues. Among the issues discussed in the main report and in a recent POSTnote[4] are:

- genetic testing in a clinical context (e.g. for screening, pre-natal diagnosis, pre-implantation genetic diagnosis, and presymptomatic testing);
- genetic tests supplied directly to the public;
- whether third parties such as insurers and employers should have access to test results and if so under what circumstances;
- the appropriateness of subjecting routine pharmacogenetic tests (which reveal nothing about a person's state of health) to the same regulations as tests for genetic diseases (which do).

Research Issues

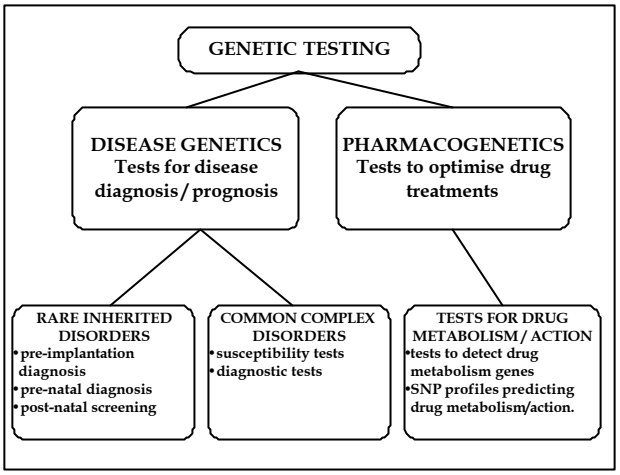
The report also covers a number of research issues in the rapidly developing area of human genetics. Among the main issues covered are:

- Gene therapy. The report discusses issues such as the possibility of gene therapy causing inadvertent germ line modification, the use of stem cell transplants[5] in utero, and recent concerns over the safety of certain vectors used to transfer genes in gene therapy.
- Pharmacogenetics. While genetic testing will allow drugs to be matched to individual patients, such developments raise issues for the design and regulation of clinical trials and post-marketing surveillance systems.

Social and Ethical Aspects

The potential benefits stemming from genome research will not appear tomorrow – new diagnostics may emerge over the next few years, but new treatments could take 10 or more years. In the absence of immediately obvious clinical benefits, the public may become increasingly sceptical about the 'new genetics'. This would be unfortunate, in view

FIGURE 1 DIFFERENT TYPES OF GENETIC TESTING



Source:Adapted from Roses AD, "Pharmacogenetics and Future Drug Development and Delivery", The Lancet, 355, 1358-61

of the fact that the public's co-operation is required to turn the mass of code into clinically useful diagnostics and treatments. Researchers need to link genetic data (the sequence code) to data from medical records about illnesses, environmental factors, etc.

Initiatives such as the MRC/Wellcome Trust UK-based population biomedical collection require the recruitment of large numbers of people to donate samples, divulge information, etc. – their success will thus depend to some extent on public attitudes to such exercises. In this context, the MRC and Wellcome Trust are currently undertaking a public consultation exercise to ensure that the public's views are taken into account when details of the biomedical collection are finalised.

The Wellcome Trust also funds other projects through its Medicine in Society Programme, which is funded to the tune of £15M over five years. This encompasses a Biomedical Ethics Programme focused on genetics and neuroscience issues. The overall aims of the programme are to:

- support research into the social, ethical and other consequences of advances in biomedicine;
- build/enhance national capacity in this area;
- ensure that the research is relevant to public policy and that the results are communicated effectively to those making such policy.

[4] **see POSTnote 139, 'Genetic Testing', May 2000**
[5] **see POSTnote 141, 'Stem Cell Research', June 2000**