The Parliamentary and Scientific Committee

INFORMATION FOR MEMBERS

From the

Parliamentary Office of Science and Technology

Briefing Note

37

November 1992

PATENTING HUMAN DNA

The UK is active in a large international programme to decipher the human genetic code. A split has developed between those (led by the US Government but followed reluctantly by the UK) who are attempting to patent basic genetic information as soon as it is decoded, and those who believe that patents should only be considered after further inventive steps. This undermines the international cooperation on which the programme relies.

This briefing note considers the arguments surrounding the patenting of human DNA and the issues raised.

THE HUMAN GENOME PROJECT

The questions on intellectual property rights (IPR) arise from the effort to decode, or sequence, the human genome under the Human Genome Project - HGP (see the Box for an explanation of the technical steps and terms involved). The HGP comprises a number of national genome programmes, primarily in the USA (spend £83M p.a.), France (£26M p.a.), Germany (£32M p.a.), Japan (£11M p.a.) and the UK. In the UK, funding comes from the MRC (£15M p.a. of which £5M is specifically for human genome mapping), the EC (which will make 27.5M ECU available to European genome researchers over the next 18 months) and from a number of health charities.

Coordination between funding organisations is assisted by the Human Genome Organisation (HUGO), which was established in 1989 and consists of 500 research groups from 23 countries. UK scientists are well-represented in both the HUGO membership and on its governing council. HUGO's role is to encourage collaboration and act as a central depository for data, to minimise duplication between national efforts and ensure the effective use of resources through information exchange, uniform standards for data etc.

In the last three years, scientists have focused upon the estimated 2-5% of the genome related to genes which are active in human cells, by isolating 'complementary' (c)DNA. The procedures used (see Box) reveal the sequences of pieces of DNA of around 300 or so bases long. Thousands of these **partial gene sequences** have

BOX DECODING THE HUMAN GENOME

The **human genome** is the entire genetic material of a person, and is contained in the 23 pairs of **chromosomes** present in most body cells. The essential genetic material in the chromosomes is a chemical molecule called **DNA** which is made up of a series of units called **bases**. There are only four kinds of base, usually referred to by the letters **A**, **T**, **G** and **C**. The sequence in which they occur along the length of a chromosome constitutes the **code** which provides the genetic instructions for the cell to produce the many proteins necessary for life. The code is thus like an instruction book in a language of just four letters. Each set of instructions responsible for a discrete cellular activity or product is called a gene - there are thought to be between 60,000 and 100,000 genes in the human genome.

Sequencing the genome discovers the order of the bases in each gene. It employs instruments which can automatically 'read' the base sequence in a fragment of DNA presented to them. Machines can handle up to a few hundred bases, thus the DNA has first to be cut up into handleable lengths before it can be sequenced. Since the genome has some 3 billion bases, to examine every part would be a mammoth task even with international cooperation, so scientists have attacked the task in an economical way. Current approaches exploit the fact that only around 2-5% of the genome appears to be comprised of genes (the function of the rest is unknown).

The 'active' genes can be detected by looking for some of the intermediate stages that the cells use to transcribe the information from the DNA and turn it into proteins (see POST Briefing Note 15). The cell makes copies from the 'master' DNA which act as 'messengers' to other parts of the cell, each messenger taking the genetic instructions for only one gene. The messengers can be isolated from cells and DNA copies prepared in the laboratory. The resulting DNA is known as **complementary DNA (cDNA)**, and a (purified) collection of cDNA copies of all the active genes found in a particular human cell type is known as a **'cDNA library'**. The genes contained in such a library are then cut into smaller fragments which can be sequenced. These **partial gene sequences** are around 300 bases long (typical genes contain thousands of bases), and give researchers sufficient sequence data to differentiate between new genes and those which are already known.

Because cDNA libraries from various cell types are commercially available and the sequencing techniques are automated, obtaining such partial sequences is a routine laboratory procedure. Although whole gene sequences can be deciphered by this approach (by sequencing several overlapping fragments) this is time-consuming - hence the present emphasis on describing partial sequences from many genes rather than whole sequences from fewer genes.

FIGURE SOME OF THE STEPS FROM PARTIAL GENE SEQUENCES TO USEFUL APPLICATIONS

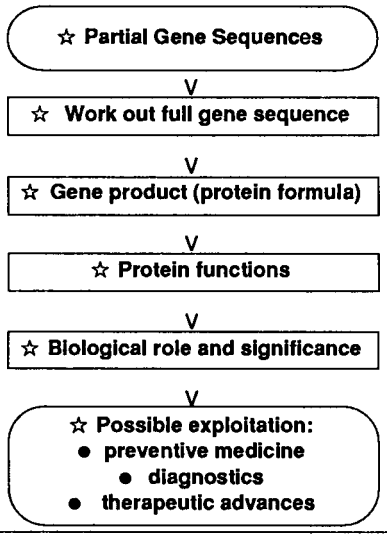

Table EXAMPLES OF DISEASES WITH A GENETIC BASIS

Single gene disorder	**Gene product (protein) involved**
Sickle cell anaemia	ß-globin
Thalassaemia	α and ß-globin
Haemophilia A and B	Factor VII and IX
Cystic Fibrosis (CF)	CF transmembrane regulator
Duchenne Muscular Dystrophy	dystrophin
Chromosomal disorder	**Chromosomes involved**
Down's syndrome	Extra copy of chromosome 21 caused by faulty chromosome separation after fertilization
Chronic Myelogenous Leukaemia	Hybrid gene on chromosome 22 caused by recombination of DNA from chromosomes 9 and 22
Diseases with a probable genetic component	
Cancer	
Some forms of Heart Disease	
Late onset Diabetes	
Rheumatism	
(Possibly) Schizophrenia	

been obtained routinely and checked against computer databases to see if they come from genes which are already known, or are from previously uncharacterised genes. It is these partial gene sequences that are the subject of the current patent dispute.

Collecting this raw data is however only the first of many steps towards a scientific understanding of the gene's function and possible commercial developments (see the Figure for a summary of the many and complex steps involved). The most obvious candidates for new therapies arising from knowledge of gene sequences are diseases known to have a genetic basis. These include around 5,000 or so diseases thought to be caused by a fault in a single gene (**single gene disorders**). **Chromosomal disorders** are less common and arise as a result of gross changes in chromosome structure. Other diseases are thought to have a significant **genetic component** (**Table 1**). In addition, the increased understanding of basic biological processes (e.g. the immune system) arising from genome research also holds out the prospect of therapeutic advance in other areas.

THE PATENTING DISPUTE

In view of the possible commercial applications, there is a potential conflict between advancing scientific knowledge through international collaboration and protecting national economic interests. This has been a particular issue in the USA, where policy-makers are wary of projects which will produce knowledge that may be commercialised by other countries - they wish to see the investment of public funds lead to new products and services to the benefit of the USA. As a result, the US National Institutes of Health (NIH) decided in 1991 to lodge patent applications for the basic data from their cDNA sequencing programme.

The NIH justified its decision on the grounds that patenting would protect an investment of publicly-funded research. They argued that companies would be unlikely to invest money to explore potential applications of the new knowledge unless it was adequately protected. The NIH saw the previous practice of publishing partial sequence data as undermining the patentability of whole genes in the future.

This stance is opposed by scientists outside the USA and by many inside; in addition, all the national scientific associations which belong to the International Council of Scientific Unions (ICSU) are opposed to such patenting. The MRC also believes that patents should not be granted at such an early stage in discovery but, after consultation with the UK Government, filed a similar application to protect any potential UK interests. All other participants in the HGP have however decided to continue to make their partial sequence data available without attempts at patent protection. This has led to a dispute amongst HGP collaborators and governments about when to seek patent protection for human gene sequences, which is tending to inhibit the international cooperation on which the HGP is based.

To date, the NIH have filed applications for over 6,700 different (cDNA) partial gene sequences with the US Patent and Trademarks Office (USPTO); the MRC applications concern about 1,100 sequences and have been filed with both the UK Patent Office and the USPTO. In contrast, 8,500 partial gene sequences have been published by French, German and Italian scientists without seeking patent protection; however Japan has yet to publish its estimated 5,000 sequences.

ISSUES

Gene sequences and patent law

Although national patent laws vary, all require that an invention be **novel** (i.e. must not previously exist in the public domain), **non-obvious** or **inventive** (i.e. cannot be readily surmised from publicly available material)

and **useful** (i.e. have demonstrable utility). Natural materials are not normally patentable, except where there is **human intervention** which is inventive and supplies an element of novelty. Thus patents have been granted for the **isolation** and **purification** of natural substances (e.g. steroids) and the **genetic manipulation** of organisms (see POST Briefing Note 26). Patents have also been granted on genes *per se* on the basis of partial gene sequences (e.g. the gene coding for a blood clotting agent Factor IX). However, in all cases to date, the function of the gene has been known.

The NIH applications seek protection for any DNA (however long) which contains the partial gene sequence concerned or any part thereof of 15 bases or more. This approach is unprecedented because it attempts to patent sequences of genes or parts of genes, whose functions have not been identified. The applications thus represent important test cases that will determine how the basic tenets of US patent law are applied to genome research, with implications for the biotechnology and pharmaceutical industries.

In a decision in August 1992, the USPTO rejected the NIH applications. One apparent ground for the rejection was that the sequences lacked legal novelty because some of them originated from publicly available cDNA libraries (see Box). Another was because the sequences were 'obvious' in so far as some of the claimed sequences of only 15 bases long were found to be in genes whose sequence was already known and available in the open literature.

This decision is however by no means the end of the matter. An initial rejection is very much the norm in new areas of patenting; the ultimate granting of a patent often follows a long dialogue between the applicant and the USPTO to refine the original claims made. The objection concerning novelty might not be sustained on appeal, since many of the base sequences deciphered by the NIH were previously unknown and hence not in the public domain. Equally, the 'obviousness' objection could be circumvented if the NIH rewrote the claims to refer to longer portions of the sequences. The NIH have several months to submit modified claims if it decides to appeal.

The MRC applications are not at such an advanced stage, having been submitted as a response to the NIH applications. The MRC and NIH claims are not carbon copies, however any final decision by the USPTO on the NIH claims will clearly be very relevant to the MRC's applications. UK patent procedures are unlikely to lead to a decision on the MRC applications before 1994.

Objections to patenting these sequences have been many. Some objections are on ethical grounds on the basis that no-one should 'own' the basic genetic code of a human being, and that it is unfair and inappropriate to reward what is largely an automated laboratory procedure which can be being applied worldwide by anyone with a sequencing machine. Traditionally, patents have only been considered after more of the inventive steps in the Figure have been completed.

On strictly legal grounds, most observers believe that any NIH appeal should fail because the biological function of the genes is unknown. The original NIH strategy was to claim specific uses for the sequences (as genetic markers, for forensic identification and for tissue typing), while at the same time claiming property rights for all future uses of the sequences and any genes in which they occur. The specific uses suggested all involve using the sequences as **probes**; i.e. as a means of finding out if or where that particular sequence occurs in a sample of DNA. However, **any** sequence has the potential to be used in such a way - it is an inherent property of the DNA molecule. Most thus believe that the NIH applications should not be allowed because they do not specify **what** genes the sequences could be used to probe for; neither have NIH established other more specific uses of the DNA sequences they are seeking to protect.

Although the NIH applications have been widely criticised by researchers, industry and patent experts alike, there is a view that current US patent law may be sufficiently generous to allow at least some elements of the claims on appeal, especially in view of the difficulty which courts can have in adjudicating on such technical matters. Some consideration of the possible consequences of patents being granted is thus warranted.

If patents were awarded on the DNA sequences *per se*, and on all future uses thereof, any company wishing to develop products based on knowledge of the sequences would need a licence from the NIH. Biotechnology and pharmaceutical companies view the award of such monopolies as being a disproportionate reward for routine laboratory work and suggest that it would precipitate a rush to identify cDNA sequences by researchers in both private and public sectors, in an attempt to claim as much of the genome as possible. There would be considerable confusion if different groups gained patents that covered different sections of the same gene or if patents were awarded on genes, parts of which were already in the public domain. The ensuing legal actions about who owns what, would be unlikely to help the commercial development that patenting is designed to facilitate.

UK policy

The dispute has already undermined the spirit of international co-operation upon which the HGP depends. Other European countries are opposed to such patents and have not sought them. Indeed, French scientists have recently given their information to UNESCO, on

the grounds that the basic sequence of human DNA is part of the 'scientific heritage' of mankind, and have called on scientists in other countries to do likewise.

The MRC points out that its applications should be seen as a defensive measure, designed to protect UK interests in the event of the US applications proving successful. Indeed, the Government has stated that the MRC would waive any rights arising from the applications in the event of an international agreement that no country should seek patents on genome sequences of unknown utility identified as a result of publicly funded research. The Government also said in March that it intended to pursue such an agreement through both bilateral and multilateral (e.g. OECD) channels. Despite this, some European researchers are refusing to deposit sequence data in the MRC's database as a result of the UK action.

While Government policy continues to emphasise the need for an international understanding, no steps have been taken towards this end, largely because it was expected that the USPTO would reject the NIH applications. However, the grounds for rejection do not appear to have settled the matter, and many fear that a further period of uncertainty may follow, inhibiting both the exchange of information and the search for useful applications. Hence, the UK Government and MRC are coming under increasing pressure to reconsider the position rather than to rely on the USPTO to resolve the situation.

The MRC has two options; to pursue existing policy or to withdraw the patent applications. On the first, most see the negotiation of a formal agreement through an international body such as the OECD as potentially very time-consuming. This has led to the suggestion that an informal agreement between HGP collaborators would be more appropriate, and should be pursued with the new US administration. UNESCO or HUGO could also provide a forum for such efforts; another possibility would be for the MRC or the Office of Science and Technology (OST) to host an international meeting on genome patenting. At present, the OST is consulting interested parties within the UK (e.g. MRC, health charities, industry) to discuss related policy.

Withdrawing the MRC patent applications would have an immediate (beneficial) effect on international collaboration within the HGP and bring the UK more into line with other European thinking on genome patents. Such a move is also likely to result in European genome researchers resuming the deposit of new sequence data in the MRC database, and make the data more readily accessible to UK companies with an interest in developing genome-based products (currently they have to access such data through the European Molecular Biology Laboratory database). There are also questions over whether the costs of patenting are justified in view of the recent escalation in the release of sequences into the public domain, since publishing a sequence precludes subsequent patent protection by other groups. On the other hand, unilateral withdrawal would weaken the UK's negotiating position *vis a vis* the USA, and risk the loss of IPR in the event of the USPTO approving patents in this field. This might in turn undermine the biotechnology industry outside the USA.

However, neither of the options outlined above are seen as a long-term solution, since they do not address the issue of how companies seeking to develop genome-based products can best protect their research investment. Additionally, any informal agreement not to pursue patents on cDNA sequences of unknown biological function will only apply to publicly funded organisations, not to individuals or commercial concerns which are now active in the USA.

This has led some to argue for more effort to harmonise international patent law. Currently, the fora through which such initiatives occur are the tripartite talks between the European, Japanese and US Patent Offices (in which the UK is only indirectly represented) and through the Uruguay round of the GATT talks (where the UK exerts rather more leverage). Some progress has been made; for example, it has recently been agreed in principle that Europe and Japan would adopt the US system giving inventors one year's grace to apply for patents following publication, if the US reciprocates by adopting the European system of first to file (rather than the current first to invent). However, progress is slow and the US Congress is not expected to consider this matter in detail until 1994.

UK and European patent laws also require an applicant to demonstrate significantly greater 'industrial' and 'commercial' utility than US law. Some patent experts argue that international patent law should seek a common understanding of industrial or commercial utility. Other problems lie in the present ability to claim uses much wider than those which can be demonstrated at the time of application. In the specific case of DNA patenting, if applicants had to know the biological function and significance of the sequence they were trying to patent, many of the difficulties described in this note would be avoided.

FURTHER READING

Additional details and background information are available from POST, 2 Little Smith St., London SW1P 3DL, tel: (071)-222-2688.