Project to sequence human genome moves on to the starting blocks

London. Almost exactly ten years after the idea was first suggested, funding agencies in the United States and Britain have given the green light to pilot projects designed — if successful — to lead to the first full-length sequence of the human genome.

Next Saturday (13 May), John Sulston, director of the Sanger Centre in Britain, and Bob Waterston, director of the Genome Sequencing Center at Washington University, St Louis, plan to outline a joint proposal for an international effort to achieve this objective to a meeting at Cold Spring Harbor, New York.

Research teams at various centres would be responsible for helping to sequence a set of cosmids covering each chromosome in turn. These sequences would then be combined, and the complete sequence data would eventually be placed in the public domain.

Although the specific strategy being suggested by Sulston and Waterston would, for technical reasons, result in a sequence that still contained some gaps, many feel that its cost — estimated at between $300 and $500 million — is sufficiently lower than earlier estimates to compensate for any omissions.

Sulston is also confident that a fully mapped sequence, even if only 99.9 per cent accurate, would still be invaluable to the whole research community. He says that the sequence could be established in as little as five years, and that the task is now timely and urgent. "Every year that goes by without embarking on this, we are wasting both money and time."

Agencies such as the National Institutes of Health (NIH) and the Department of Energy in the United States, and the Medical Research Centre (MRC) and the Wellcome Trust — the main supporter of the Sanger Centre — in Britain, appear to be in agreement.

The NIH, for example, has issued a request for applications for pilot large-scale sequencing projects, and is prepared to give applications for any such project "expedited treatment" in order for it to start by the end of the year. It has earmarked $20 million to be distributed between such projects and work on advanced sequencing techniques.

Similarly the MRC has decided to make available some of the extra money it won from the government in this year's science budget for 'strategic' objectives.

Both Sulston and Waterston say they intend to apply for the funds being made available. Whether their specific proposals receive approval will depend on how they fare during the peer review process.

But, given the strong track record of both groups in sequencing the genome of the nematode worm *Caenorhabditis elegans*, there will be little surprise if their applications are successful.

Whatever the outcome, there is little doubt that the joint proposal has helped to crystallize a growing consensus that, after ten years of preliminary work, the time has arrived for a preliminary assault on the 3-billion bases of the complete human genome.

"There is a lot of interest in moving on to sequencing more quickly," says Elke Jordan of the NIH's National Center for Human Genome Research (NCHGR). "In the past few months, we have gone from a feeling that we were not quite ready to a feeling that 'yes, we should go'."

There is similar enthusiasm at the Wellcome Trust. "We have not yet agreed to fund the project; that will depend on the outcome of the normal peer review processes," says Michael Morgan, programme director at Wellcome. "But we would not be in this position [of accepting proposals] if we did not accept that it is a sensible thing to do."

Up to now, support for genome research by the NCGHR and other bodies has been distributed relatively widely. One priority, for example, has been novel sequencing technologies, based on the argument that large-scale sequencing should wait until the costs become significantly lower.

Some, indeed, continue to support this point of view, and advise caution before rushing ahead with the strategy that Sulston and Waterston are suggesting. Others, however, point out that there have been few major technical advances in sequencing in the past couple of years — and that waiting longer could delay the project unnecessarily.

In contrast, the proposal being put forward by Sulston and Waterston is built largely on techniques already in use in both laboratories to sequence the nematode as well as parts of the human genome.

The approach they suggest would be based on the so-called 'shotgun' sequencing of a minimal set of cosmids covering each human chromosome in turn. The resulting sequences will then be put together, with additional effort applied to 'finishing' the data where necessary.

Sulston says that, on the basis of their previous experience and simulations with available data, he and Waterston believe that the resulting sequences will contain no more than 3 gaps per cosmid — caused chiefly by long tandem repeats, inverted repeats and homopolymeric stretches. Nevertheless the sequence would still contain sufficient information to act as a fine-grain map containing, for example, the accurate location of the entire set of human genes.

"The general idea is to make a thorough attempt at maximal representation of the genome both at the clone level and at the sequencing level, and to tie the whole thing together by identifying the landmarks," says Sulston. "This will provide a means of organizing as much as possible of the information that is already being developed."

Sulston and Waterston claim that, if carried out on a large enough scale, the cost of sequencing could be reduced to about 10 cents a base (for comparison, the nematode ▶

King Holmes/SPL

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Worm's eye view: sequencing techniques for the nematode genome may now be applied to humans

Technicians die in Ariane V accident

Paris. **The European Space Agency (ESA) has set up an inquiry into the deaths last week of two technicians from France's national space agency (CNES) during pre-test verification checks on the cryogenic engines for the launcher Ariane V.**

The technicians died by asphyxiation following the leak of liquid nitrogen. The inquiry is expected to provide details on the origins of the leak, which took place during inspection of the umbilical mast used to fuel the launcher on the launch-pad.

A spokesman for ESA said that the accident is unlikely to delay the first launch of Ariane V, still scheduled for the end of November, with the scientific mission Cluster as its pay-load. □

sequence, which is now 15 per cent complete, and is expected to be finished by the end of 1998, costs about 50 cents a base.)

Some remain sceptical about whether the figure can indeed be reduced this low. "We have yet to see anything on paper to allow us to judge whether it is feasible," says one sequencer while admitting that, in principle, he is strongly supportive of the approach that Sulston and Waterston are taking.

Even at this low price, however, Sulston is keen to emphasize that, for scientific and other reasons, he would like to see the project carried out collaboratively between a number of centres, with different research groups agreeing to concentrate on different chromosomes.

Such an approach is likely to go some way to meeting fears that arose earlier this year, when the proposals were first being discussed by genome research centres, that the whole project would be dominated by the Sanger Centre and the St Louis group. "We have each said that we would be willing to do one third of the operation, but we do not insist on doing it all; the intention is to have a fully cooperating network."

Cooperation — not competition (apart from that for funding, with peer review ensuring the quality of the data produced) — will be the key. "This is a place where the market does not work well," says Sulston. "It is inappropriate when you are trying to establish a framework where there is only one solution; you need a strategic plan."

David Bentley, head of human genetics at the Sanger Centre, suggests as a possible model the consortium of laboratories set up in Europe, with the backing of the European Commission in Brussels, to sequence the yeast genome. "It has been put together in a very informal and flexible way, and people have shifted their aspirations according to what they have been able to achieve," says Bentley. "We will try to work in the same way with the human genome."

If the consensus behind the overall strategy is still growing, it is not yet complete. Some still argue, for example, that a more effective approach would be to build on the knowledge of detailed sequences being established through cDNA libraries, in other words starting with a focus on functioning genes. Others continue to prefer a strategy known as 'skimming' the genome, carrying out the complete sequencing of randomly selected sections.

The Cold Spring Harbor meeting is expected to provide a forum for vigorous debate between those supporting rival strategies. But for Sulston, Waterston and their colleagues, the funding agencies prepared to back their approach, and industrial researchers keen to exploit their results, the main question now is not whether, but when, to start work. If the proposed pilot projects succeed, the main sequencing would be ready to start by the end of the next year. And the whole project could be finished by the year 2001 — five years earlier than many had previously estimated.

US court rules discovery of gene sequence 'not obvious'

London. In a major boost to the biotechnology industry, a US appeals court has ruled that the discovery of a novel gene sequence cannot be described as 'obvious' — and that the sequence in question can therefore be legitimately included in a patent — even if the existence of the gene was previously known, as were techniques for obtaining its complementary DNA (cDNA).

The ruling is based on the fact that, even though the gene can be deduced from the protein that it expresses, so-called 'redundancy' in the genetic code means that its precise nucleotide sequence can be considered unknown until it has been 'discovered'. (The redundancy results from the fact that more than one nucleotide base triplet in a DNA sequence can code for the same amino acid in the protein, and that a wide range of sequences can therefore theoretically code for the same protein.)

The court's decision is a direct challenge to those who have been arguing that the controversy over patents on human and animal genes is likely to blow over relatively quickly as sequencing and cloning techniques become routine laboratory practice, and the discovery of new gene sequences, therefore, a relatively unremarkable affair.

Indeed, one of the main implications of the ruling is to limit the range of techniques whose existence can be invoked to dismiss a discovery as 'obvious', and therefore to increase the degree of predictability in the US patents system.

At the same time, the ruling is being seen in the biotechnology industry as putting gene sequences on the same level — at least as far as patent protection is concerned — as chemical molecules, whose eligibility for patent protection (provided they fulfil criteria of novelty and usefulness) is clear.

"The implication is that new technologies in the biotechnology and genomics area can be covered by existing patent laws in the same way as other areas such as chemistry and physics," says Robert Benson, attorney for Human Genome Sciences (HGS) Inc. in Rockville, Maryland. HGS has over 70 patent applications on partial and full gene sequences awaiting a ruling from the US Patents and Trademark Office (PTO), and hopes the court ruling will increase their chances of being approved.

The decisions concern the application by Thomas F. Deuel and three fellow scientists at Washington University in St Louis, Missouri, for a patent on isolated and purified DNA and cDNA molecules encoding heparin-binding growth factors (HBGFs), proteins that stimulate cell division and therefore facilitate the repair or replacement of damaged or diseased tissue.

A PTO examiner had initially rejected the patent application — which includes coverage on all possible DNA molecules coding for the disclosed proteins — on the grounds that cloning the gene would have been "obvious" to anyone "of ordinary skill in the art at the time of the invention".

The examiner quoted two instances of 'prior art' to support this view. One was that a heparin-binding protein had already been identified and isolated from human and bovine brain tissue, and both proteins had been found to share the first 19 amino acids in their N-terminal sequences. The second was that methods for isolating either DNAs or cDNAs by screening sequence libraries had been initially described by Tom Maniatis of Harvard University in 1982.

The examiner argued that it would have been relatively straightforward to use the known N-terminal sequence, and the techniques described by Maniatis, to design a probe to isolate a gene encoding the protein from a cDNA library. Although challenged by the applicants, the ruling was subsequently upheld in November 1993 by the PTO's own appeals board.

Last month, however, in response to a challenge mounted with the backing of both the Biotechnology Industry Association and the Bay Area Science Centre in San Francisco, as well as biotechnology companies such as Chiron Inc., the US Federal Circuit court overturned the PTO's objection.

The court ruled that, even though knowledge about the proteins identified earlier meant the general chemical nature of the cDNA molecules in the patent claims "may have been obvious" — and the knowledge that "some gene existed" may have been clear — redundancy in the genetic code meant that the precise cDNA molecules isolated from the library "would not have been obvious".

"Until the claimed molecules were actually isolated and purified, it would have been highly unlikely to one of ordinary skill in the art to contemplate what was ultimately obtained," said the court. "What cannot be contemplated or conceived cannot be obvious."

More broadly, the court stated that a general motivation to search for some gene that exists "does not necessarily make obvious a specifically-defined gene that is subsequently obtained as a result of that search".

It added: "Even if, [. . .] the existence of general cloning techniques, coupled with knowledge of a protein's structure. might have provided motivation to prepare *a* cDNA or made it obvious to prepare *a* cDNA, that does not necessarily make obvious a particular claimed cDNA". **D. D.**