



G8 on GM
Krebs' panel urges industrial nations to get it together
p112



Waste of space?
Zero-gravity crystals have yet to prove their worth
p114



AIDS therapy
Panel ponders South Africa's policy on AZT and other drugs
p115



Military fallout
India's research budget gains from political tension
p116

Genome leaders told to keep their eyes on the main prize

JAMES D. WILSON/LIAISON

Washington

The Human Genome Project is in danger of being sidetracked from its goal of sequencing our entire genetic blueprint by 2003. That is the fear of several prominent advisers to the project, who are worried that enthusiasm for turning the emerging 'rough draft' of the human genome into a finished product is waning.

Their concern emerged last week at a meeting of the US National Advisory Council for Human Genome Research. It was sparked by a surge of interest in initiatives designed to make sense of the draft human genome. These include efforts to sequence the mouse genome and to develop computational tools and databases.

The Human Genome Project has sequenced about 1.7 billion unique base pairs so far. By June, the international coalition of publicly funded and charity-supported labs should have sequenced another billion. This will give a draft covering 90 per cent of the genome, sequenced five times over.

Finishing the job means filling hard-to-sequence gaps and ensuring that 99 per cent of the genome has been sequenced ten times over — the 'gold standard' to ensure that the sequence is accurate. But once geneticists have their hands on the draft, this goal may lose its appeal. At least in the short term, many are more interested in partial information on single genes than a complete genome.

This is where the other initiatives come in. Because mouse genes can be experimentally 'knocked out', the mouse genome can help to define the function of genes common to both mouse and human. Creating a library of mammalian complementary DNA (cDNA), representing the coding sequences of genes expressed in particular tissues, should help to elucidate gene functions.

Computational strategies will also contribute by annotating the draft with functional information. David Lipman, director of the National Center for Biotechnology Information in Bethesda, Maryland, says that algorithms developed by his centre and others are able to trawl the draft genome and find homologues to known genes "amazingly well". The forthcoming relaunch of the main



Collins: each new genomics initiative is valuable, but the challenge lies in juggling them all.

sequence database (see below) will let annotators navigate the genome with ease.

Francis Collins, director of the National Human Genome Research Institute, stresses that each of these initiatives is valuable. "Every full-length cDNA will add value," he says. "So will every mouse sequence read."

But at last week's meeting he warned of the problem of finding resources to support these initiatives without delaying the production of the finished sequence: "How we juggle all these things is challenging."

Some of Collins's advisers fear that the juggling will compromise the project's main

goal. "Somewhere in here we need to do something well," advisory council member Maynard Olson, of the University of Washington in Seattle, told the meeting.

Michael Morgan, chief executive of the Wellcome Trust's genome campus, which leads the British branch of the project, says the US's international partners are determined not to miss the 2003 target. "We must not lift our eyes from the goal of completing the genome," he says.

Originally, the Human Genome Project was intended to produce finished sequences, chromosome by chromosome. That resulted in the project's biggest milestone so far, the virtually complete sequence of chromosome 22 (see *Nature* 402, 447; 1999). However, the project shifted strategies towards an interim goal of producing a draft of the entire genome after Celera Genomics, a sequencing company in Rockville, Maryland, joined the race.

Although the strategy has changed, the final goal should not, argues genome advisory council member Aravinda Chakravarti of Case Western University in Cleveland.

"It's not a question of either/or," says Robert Horvitz of the Massachusetts Institute of Technology, a member of the genome advisory council. "It's a question of balance." But that leaves Collins with a difficult balancing act.

Paul Smaglik

Revamped GenBank offers extra data links

Washington

GenBank, the leading human genome database, is about to undergo a facelift.

The National Center for Biotechnology Information (NCBI) in Bethesda, Maryland, will relaunch GenBank next month in a form that will be easier for users to navigate and will include a wealth of new 'annotation', such as information on genetic markers and gene and protein functions.

Users will be able to click on

part of an onscreen chromosome and retrieve information such as the predicted function of a partially sequenced gene.

Until now, this information has only been available for chromosome regions for which a contiguous DNA sequence is available. Much of this annotation will be generated automatically by computer algorithms.

Users will also be able to navigate more easily from various genetic markers in the database,

such as sequence tagged sites and expressed sequence tags. And they will find it easier to zoom in and out on specific chromosome regions, says David Lipman, director of the NCBI.

The new GenBank will also allow more links with the database of single nucleotide polymorphisms. This separate public database details the subtle variations in sequence found between individuals for certain genes.

P. S