



**Moving story**  
Natural history  
archive heads for  
a new home  
p290



**Added values**  
Supporters rally  
behind the standard  
model of physics  
p291



**Down under**  
US physicists want  
to mine a rich seam  
of research  
p292



**Fears of a clone**  
Rome-based  
doctor plans  
human cloning  
p293

# Free access to cDNA provides impetus for gene function work

**Alison Abbott, Munich**

Hot on the heels of the publication of the draft human genome sequence, several efforts are under way to provide all scientists with free, synthetic versions of human genes. Their availability will speed up attempts to understand the function of human genes and their associated proteins.

Each of the synthetic genes, called complementary DNAs or cDNAs, can be induced to synthesize the protein encoded from the real gene from which it was derived.

A German team of researchers, half-way through a three-year pilot study, announces in this month's *Genome Research* (11, 422–435) that it has generated 500 previously unknown full-length human cDNAs, and an additional 1,000 known cDNAs. All the cDNAs have been inserted into transportable packages known as vector clones, and are freely available to academic and industrial scientists.

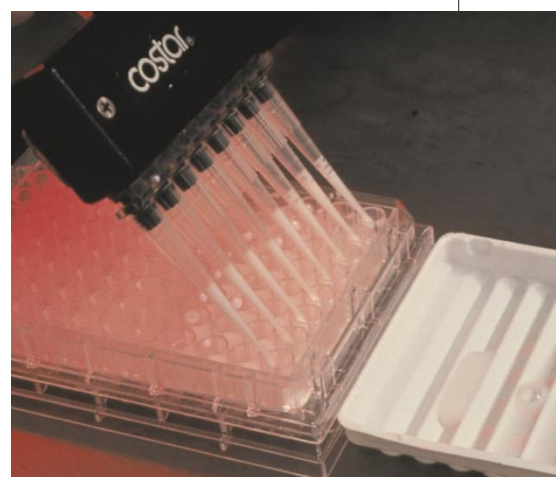
Since these results were submitted, the group, led by Stefan Wiemann of the German Cancer Research Centre in Heidelberg, has

made available over a hundred additional cDNA clones. "Requests for the clones from scientists around the world have increased dramatically this year," says Wiemann.

A typical request, he explains, might come from a cancer researcher who sees very low expression of a particular gene in one type of tumour, and wants to know what it does. If the cDNA for that gene is available as a clone, the researcher can get to work on its function straight away.

The cDNAs are synthesized from 'messenger RNAs'. Each mRNA is a transcript of a gene, and its job is to direct the stringing together of the correct amino acids that constitute the protein encoded by the gene. The enzyme reverse transcriptase is used to direct the synthesis, or reverse transcription, of cDNAs from mRNAs.

The creation of full-length cDNAs, including the entire protein-coding region as well as regulatory regions, is no mean achievement. The enzyme is inefficient *in vitro* — it tends to 'fall off' its substrate before the end of the operation, leaving incomplete cDNAs.



**In sequence: cloning complementary DNAs can expedite studies of gene function.**

Cloning is also inefficient for long cDNA inserts, so it is easier to make short full-length cDNAs than long ones. The average size of the German cDNAs was about 2.5 kilobases (kb), equivalent to a maximum of 500 amino-acid residues.

The work to systematically produce cDNAs was pioneered by the Kazusa cDNA project in Japan, which began in the mid-1990s. This project selects only long cDNAs to synthesize and clone — they are at least 4 kb, equivalent to proteins larger than about 1,000 amino-acid residues. Of the 2,000 cDNAs produced so far, only about half are full-length. "We see large proteins as being the most interesting biologically," says Osamu Ohara, head of the project, which is based at the Kazusa DNA Research Institute in Chiba. The Japanese cDNA clones are freely available to researchers in academia, but not to those in industry.

In the United States, the National Institutes of Health (NIH) started a large cDNA initiative last year, as part of its Cancer Genome Anatomy Project. The NIH project will provide clones freely to all researchers and, like the German project, is creating its library of cDNAs randomly, rather than through size selection. It will be larger than the German project, and has already registered 2,800 full-length sequences with

## Anti-AIDS drugs available 'at cost'

**David Dickson**

As public pressure mounts for AIDS drugs to be made available to poor countries at an affordable cost, a leading drug company has pledged for the first time to sell its protease-inhibitor therapies to them at cost price.

Merck said last week that it will make Crixivan (indinavir sulphate) — a component of the triple-drug anti-AIDS therapy widely used in the West — available at \$600 a year in developing countries. Another antiretroviral agent, Stocrin (efavirenz), is to be sold at \$500. Other major drug companies are expected to follow suit.

The move is an attempt to head off mounting criticism over the cost of AIDS drugs in developing countries (see *Nature* 410, 3; 2001). The announcement came as a court case, brought by 40 drug companies

against the South African government, threatened to turn into a public-relations debacle for the companies.

Merck says that developing countries taking up its offer must introduce procedures to prevent the re-exportation of the drugs to the developed world, where their official price is at least an order of magnitude higher.

"Even at these new lower prices, the issue of access is not resolved," Per Wold-Olsen, president of Merck's human-health business in Europe, the Middle East and Africa, said last week.

Hearings on the legal action, which seeks to block a South African law that would allow the importation of generic copies of patented drugs, opened briefly in South Africa last week, but were adjourned until next month. Kenya is reported to be considering legislation along similar lines. ■

▶ GenBank — although some of these may not be unique.

The ultimate goal of all these projects is to produce a complete set of full-length human cDNAs, corresponding to all the genes in the human genome. As there are now thought to be only about 30,000 genes, and technologies needed to deal with long or rare cDNAs are improving, this may not take very long — “perhaps only a couple of years”, suggests an optimistic Robert Strausberg, head of the NIH project.

With this rapid progress, the need to coordinate these efforts has become more pressing. “It would be nice to have a system whereby different sets of cDNAs were allocated to different groups, so there would not be too much duplication of effort,” says Wiemann. Ohara agrees: “A couple of years ago we were the only group in the game, but now we really need to think of an allocation system,” he says.

There could be complications in organizing such a system, Strausberg points out, because not all groups make their results freely available to all. The issue is to be discussed in the next few months by the various projects. In the meantime, a joint website is planned to help each project keep up with what the others are doing. ■

▶ <http://www.rzpd.de>

▶ <http://cgap.nci.nih.gov>

▶ <http://www.kazusa.or.jp/huge>

## Network aims to link species data from global collections

Georgina Kenyon, London

A global programme has been launched to collate information on the world's most important collections of animal and plant species, in a format that is freely accessible to all.

The Global Biodiversity Information Facility (GBIF), which hopes to provide unrestricted access on its website to species information held in existing museum collections, was officially launched in Brussels last week.

The GBIF has received US\$2 million of initial support from its twelve member countries. Subscriptions to the GBIF differ from country to country, depending on their wealth.

It is hoped that the facility will enable scientists in developing countries to have access to the best information about their local biodiversity, often for the first time.

“The complexity of this initiative is vast,” says Martin Sharman, a spokesman for the GBIF. About 1.75 million species have been scientifically described in different formats in various collections, but between 10 million and 100 million are thought to exist.

Organizers will decide in July where to

place GBIF's secretariat, with Australia, the Netherlands, Denmark and Spain all vying to host it.

“GBIF will assist museums, scientists and educators, and give politicians scientifically factual data,” says Carlos Martinez-Riera, a representative of the European Commission and GBIF's chief scientific officer.

Eventually the programme aims to set up a web page for each species, with a standard set of information including characteristics, taxonomy and geographical distribution, Martinez-Riera says.

Several national programmes are currently under way to develop such databases — including Species 2000 in Britain, which is not a full member of the GBIF, but is involved in steering the initiative. At the moment there is a large amount of ‘double counting’, as scientists have independently discovered and named the same species.

Each member country of the GBIF plans to nominate a scientific institution to head its contribution to the project, which was initiated at the Organisation for Economic Co-operation and Development's Megascience Forum in 1996. ■

▶ <http://www.gbif.org>

## London museum puts its animals on public display

David Adam, London

In a six-month exercise that will provide an important fillip for systemic biology, London's Natural History Museum has started moving its collection of over 20 million zoological samples from storage into a new, purpose-built public archive.

The first specimens being transferred to the £27 million (US\$39 million) archive are the museum's ‘spirit collection’ of fish, reptiles and invertebrates stored in alcohol. Many of these were brought back from Charles Darwin's five-year voyage on the *Beagle* in 1831, and others were collected on Captain James Cook's *Endeavour* expedition in 1763.

The collection includes rays, komodo dragons and tuatara lizards, the last survivors of a group previously considered to have become extinct at the same time as dinosaurs. Many of these specimens will be going on public display for the first time.

The transfer will provide an opportunity for the museum to compile an electronic database of its specimens, in line with efforts to build up a global computerized catalogue of the specimens currently sitting



On show: staff at London's Natural History Museum move a 12-foot oarfish to the Darwin Centre.

in jars and display cases across the world (see story above).

But a global computerized catalogue is some way off, museum officials say. “There is a drive to get some conformity but it will be a long time before those kinds of standards are agreed,” says Paul Henderson, director of science at the museum. “That would be just too labour intensive at the moment because too many institutions have their own systems.”

The London museum is making an effort to standardize species names, however, and is trying to make records about where the specimens came from as accurate as possible.

This first phase of the new archive is expected to open in spring 2002. Phase two will then see the ‘dry’ entomology and botany collections moved to the new facility, which will be known as the Darwin Centre. ■

▶ <http://www.nhm.ac.uk/darwincentre/index.html>