



Embargoed until Wednesday, December 8, 2004: 18:00 GMT

Chicken genome gives insights into the human genome

Draft sequence of chicken genome completed



Photo by Cath Brooksbank

Some of the Ensembl team members who were involved in the analysis of the chicken genome. From left to right: Jessica Severin (EBI), Stephen Searle (Sanger), Ewan Birney (EBI), Abel Ureta-Vidal (EBI) and Fiona Cunningham (Sanger).

Hinxton/Heidelberg, December 8, 2004 – The draft sequence of the wild chicken, *Gallus gallus*, is published today in the journal *Nature*. The analysis of this genome is not just about getting bigger eggs and tastier chicken – it's giving scientists surprising insights into the human genome. Researchers can use these new data as a tool to identify similar sequences in humans – regions previously thought to be 'junk' DNA in the human. These sequences must have an important role if they have been conserved over the 310 million years since the two lineages diverged.

Ewan Birney of the European Bioinformatics Institute (EMBL-EBI), whose Ensembl team performed much of the computational analysis of the genome, explains the importance of the chicken: "We needed to study a genome that is at the right evolutionary distance from humans. Until now, all the fully sequenced genomes have been too closely related or too distant to get the information that we wanted. For example, the mouse genome gave us lots of useful information about coding regions but we were surprised at how much of the junk DNA was almost identical in mouse and human. This is because not enough time has passed since humans and mice diverged from their common ancestor. As a result, we gained very little new insight into the non-coding regions of the human genome."

The chicken genome, on the other hand, turns out to be an invaluable tool for studying the human genome. It is at the ideal evolutionary distance from the human. During the 310 million years since their paths diverged, sequences with important functions, which include genes and their regulatory motifs, have been under pressure not to mutate because changes to the sequence would have detrimental consequences. By contrast, areas of the genome with no function – bona fide 'junk' DNA – have been under no such pressure, and sufficient time has lapsed for these regions to become quite distinct in the two genomes.

"We found strong conservation in regions of humans that were previously thought to be junk DNA. They don't code for RNA or for protein and their function is still a complete mystery, but they must have been conserved for an important reason," Birney says. "Without the chicken genome, we would have continued to wrongly believe that these regions in humans were unimportant."

Scientists have also used the new genome to find a set of genes that is notoriously hard to pin down. These 'non-

coding RNA (ncRNA) genes' code for functional RNA molecules rather than proteins. Protein-coding genes are relatively easy to identify because their sequences contain characteristic signals that say 'start here' and 'stop here'. Unfortunately ncRNA genes are less straightforward to pick out. This problem is confounded in humans because we have a lot of 'pseudogenes', sequences that evolved from functional genes that were copied and then fell into disrepair. Chickens, however, contain very few pseudogenes so their set of ncRNA genes is likely to be a functional set whose counterparts can now be identified in humans. It's also been easier to find the control switches within protein-coding genes. "When you align the chicken and human sequences, the conserved regions really leap out at you. This doesn't happen when you do the same with the mouse and human because the unimportant bits of sequence in between didn't have time to diverge," explains Birney.

Most of the functional differences and commonalities between chickens and humans can be revealed by studying their protein-coding genes. Peer Bork's team at EMBL-Heidelberg, together with Chris Ponting's team at The Medical Research Council Functional Genetics Unit based at Oxford University, approached this task by looking first for shared genes and then for shared sequences that define gene families.

"We are more similar to birds than you would think," explains Bork. "About 60% of the chicken protein-coding genes have human equivalents. But when you look at the entire gene families they have many more families in common: for example, mammals have expanded a certain family of keratins for hair production, whereas birds use a different set for the formation of feathers and claws. We also found some chicken genes involved in immune function that previously were believed to be unique to humans."

An international consortium of over 170 people from 49 different institutes across the globe assembled and analysed this genome. All of the data has been deposited into freely available public databases (see notes).

"Having the chicken genome sequence is like being armed with an antiques guide at a flea market: suddenly you have a tool that allows you to recognize which pieces are valuable," concludes Birney. We're all sure to reap the rewards. ●

18 EMBL Member States: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom



About EMBL:

The European Molecular Biology Laboratory is a basic research institute funded by public research monies from 18 member states, including most of the EU, Switzerland and Israel. Research at EMBL is conducted by approximately 80 independent groups covering the spectrum of molecular biology. The Laboratory has five units: the main Laboratory in Heidelberg, and Outstations in Hinxton (the European Bioinformatics Institute), Grenoble, Hamburg, and Monterotondo near Rome. The cornerstones of EMBL's mission are: to perform basic research in molecular biology, to train scientists, students and visitors at all levels, to offer vital services to scientists in the member states, and to develop new instruments and methods in the life sciences. EMBL's international PhD Programme has a student body of about 170. The Laboratory also sponsors an active Science and Society programme. Visitors from the press and public are welcome.

About the EBI:

The European Bioinformatics Institute (EBI) is part of the European Molecular Biology Laboratory (EMBL) and is located on the Wellcome Trust Genome Campus in Hinxton near Cambridge (UK). The EBI grew out of EMBL's pioneering work in providing public biological databases to the research community. It hosts some of the world's most important collections of biological data, including DNA sequences (EMBL-Bank), protein sequences (UniProt), animal genomes (Ensembl), three-dimensional structures (the Macromolecular Structure Database) and data from microarray experiments (ArrayExpress). The EBI hosts several research groups and its scientists continually develop new tools for the biocomputing community.

To view the draft assembly:

The draft assembly is based on 6.6-fold coverage of the *Gallus gallus* genome.

Researchers can access or view the sequence data (project accession number AADN00000000) through the following:

- EMBL-Bank (www.ebi.ac.uk/embl/) at the EMBL-EBI
- Ensembl Genome Browser (www.ensembl.org) at the EMBL-EBI and the Wellcome Trust Sanger Institute
- GenBank (www.ncbi.nih.gov/Genbank) at the US National Center for Biotechnology Information (NCBI)
- DNA Data Bank of Japan (www.ddbj.nig.ac.jp).
- UCSC Genome Browser (www.genome.ucsc.edu/) at the University of California, Santa Cruz
- NCBI's Map Viewer (www.ncbi.nlm.nih.gov/mapview/).

Press Contacts:

Cath Brooksbank PhD, EMBL-EBI Scientific Outreach Officer

Wellcome Trust Genome Campus, Hinxton, Cambridge, Tel: +44 (0)1223 492525, www.ebi.ac.uk, cath@ebi.ac.uk

Trista Dawson, EMBL Press Officer

Meyerhofstrasse 1, Heidelberg, Germany, Tel: +49 6221 387 452, www.embl.org, dawson@embl.de

Policy regarding use

EMBL press releases may be freely reprinted and distributed via print and electronic media. Text, photographs & graphics are copyrighted by EMBL. They may be freely reprinted and distributed in conjunction with this news story, provided that proper attribution to authors, photographers and designers is made. High-resolution copies of the images can be downloaded from the EMBL web site: www.embl.org.