IN THE NEWS

No Mickey Mouse science

"For hundreds of years we have watched curiously as mice run round wheels, press levers or navigate mazes. Finally we have the genetic blueprint that will unveil the mysteries of the mouse" (*The Times*). This is how Simon Festing of the Association of Medical Charities described the achievement published in the 5 December issue of *Nature*. "The mouse sequence is the phrasebook that is transforming our ability to translate the human book of life." (*The Financial Times*) said Jane Rodgers of the Sanger Centre.

"Scientists have attached such an importance to the murine genome that its [draft] sequence has been completed three years after the project was launched" (*The Times*) and the finished sequence is expected in 2005. The international consortium used the shotgun method, previously used by Celera to sequence the human and mouse genomes (both of which have been available to subscribers for 18 months). Craig Venter, the former director of Celera, said he felt "wonderfully vindicated that they [the consortium] have seen the power of a whole genome shotgun" (*The New York Times*).

"Even the genomes for small creatures are huge", says *The Guardian*, for the mouse genome is 2.9 billion bp long. Its analysis revealed that mice and humans "share the same genes for blood pressure, temperature regulation, bone manufacture, cell division, tissue growth and so on." (*The Guardian*) The sequence is complicated — "it's rather like being dropped into the middle of Tokyo with no knowledge of Japanese, and being asked to find your way around using a local newspaper" says Ewan Birney (*The Times*), so his task will be to develop "a guidebook or phrasebook, something that tells us what's good, what's bad and what's boring about this genome" (*The Times*).

Magdalena Skipper

MOUSE GENOME

The mighty mouse

The strains of mice that are widely used in laboratories today originated mainly from mouse fanciers who bred hybrids of *Mus mus musculus* and *M. m. domesticus*, and who have inbred the mouse lines extensively to create ~50 commonly used strains. Now, the mighty mouse has come to the rescue again with its ultimate contribution: the sequences of both its genome and its transcriptome, published in the December 5 issue of *Nature*. Through analysis of the initial sequences, we will learn much more about mammalian life, and about ourselves.

The 2.5-Gb mouse genome sequence, from the C57BL/6J strain, covers 96% of the euchromatic sequence. According to the Mouse Genome Sequencing Consortium, 99% of mouse genes are directly homologous to human genes, even though the mouse genome is 14% smaller than the human one. This further strengthens the idea that mammalian life can be built on only ~30,000 genes, as novel gene-prediction programs developed for this work suggest that the mouse too has approximately that number.

To understand further the mouse transcriptome, an international consortium sequenced more than 1.4 million expressed sequences, from which they described 37,806 individual 'transcriptional units'. In a boost to research, each of the sequences reported in the transcriptome paper is backed up by a freely available physical clone, generating a valuable resource for the community. As ~4,000 of these transcripts are probably not translated, this too suggests that both mice and humans are built from at least 30,000 genes, although the presence of additional end sequences suggests there might be more genes not represented in the consortium's clone collection at present. So, we seem to have about the same number of genes, 99% of these genes have similar sequence, and we both like cheese. Then why aren't mice more like us? The answer probably lies in the regulation of

MOUSE GENOME

A high-resolution atlas

Down syndrome is the most common cause of mental retardation in humans. Through a yet unknown mechanism, the presence of an extra copy of human chromosome 21 (HSA21) causes developmental defects in many organs, most notably the brain. HSA21 was the second human chromosome to be fully mapped and sequenced, and now two groups report the first comprehensive analysis of the expression patterns of HSA21 genes in the mouse. These 'atlases' promise to be a valuable resource for identifying the genes that underlie Down syndrome.

The two groups — Reymond *et al*. and The HSA21 Expression Map Initiative — identified the mouse orthologues of the 213 genes on HSA21, of which 178 are confirmed, and isolated cDNA fragments corresponding to these genes. Then, using complementary approaches, they looked at their expression in different tissues and at different developmental stages of mouse embryogenesis using reverse transcriptase PCR and mRNA *in situ* hybridization. In addition to this 'wet bench' approach, The HSA21 Expression Map Initiative relied heavily on informatics to measure their frequency in publicly available mouse EST libraries and to identify genes with similar expression profiles. They also conducted mRNA *in situ* hybridization, focusing on whole embryos and on the neonatal brain. Both groups have deposited their data in a web interactive database.

Together, these studies highlighted several HSA21 genes that, because of their expression in the tissues and organs that are most severely affected by Down syndrome, are good candidates for further investigation. Interestingly, the expression of many HSA21 genes is ubiquitous in early embryogenesis, but becomes more restricted as development proceeds. This is consistent with findings by The HSA21 Expression Map Initiative, who showed that genes found exclusively in multicellular organisms are more likely to be expressed in a spatially or time-restricted pattern, compared with the ubiquitous expression of those genes that are also found in yeast. Both groups also observed genes with similar expression patterns that cluster on particular

these genes and in many interactions between RNA and protein molecules in the cell.

In a third paper, Dermitzakis *et al.* directly compared the sequences on most of human chromosome 21 to their counterparts in mice. Even the gene-poor region of the chromosome shows extensive similarity between the two organisms, and conserved sequences seem to be under selective constraint in the mouse, despite the fact that they seem not to be encoding transcripts. Further exploration is needed into the non-coding regions of both genomes to explain the differences between us.

As the sequence has all been deposited in public databases, the mouse genome should provide a huge boost to genetic researchers. For example, crosses between the C57BL/6J strain and others are often used to map mutations that mimic human disease. In another paper in the issue, Wade *et al.* compare the C57BL/6J genome with 17 Mb of finished sequence from strain 129/Sv as well as a genome-wide sample of reads from multiple strains, finding regions of high and low polymorphism in single nucleotides between the strains. Extensive mapping of high-polymorphism areas will be extremely useful for mapping quantitative trait loci, and therefore possibly the genes that contribute to the inheritance of complex traits. Given the bonanza of information they've just received, scientists might now consider the laboratory mouse to be "man's best friend".

Chris Gunter,
Associate Editor, Nature

References and links

ORIGINAL RESEARCH PAPERS Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002) | The FANTOM Consortium and the RIKEN Genome Exploration Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **420**, 563–573 (2002) | Dermitzakis, E. T. *et al*. Numerous potentially functional non-genic conserved sequences on human chromosome 21. *Nature* **420**, 578–582 (2002) | Wade, C. M. *et al*. The mosaic structure of variation in the laboratory mouse genome. *Nature* **420**, 574–590 (2002)

WEB SITES

Ensembl mouse genome browser: http://www.ensembl.org/Mus_musculus

***Nature* mouse web focus:** http://www.nature.com/nature/mousegenome

UCSC Genome Bioinformatics site: http://www.genome.ucsc.edu

NIH Mouse Genome Resources: http://www.ncbi.nlm.nih.gov/genome/guide/mouse

chromosome regions, as recently reported in worms and flies, which hints at how chromatin structure might govern gene expression.

Mapping the gene expression of an entire chromosome at high resolution provides the next level of genome annotation and brings researchers closer to identifying the function of every human gene — the most important issue in genomic biology.

Natalie DeWitt,
Senior Editor, Nature

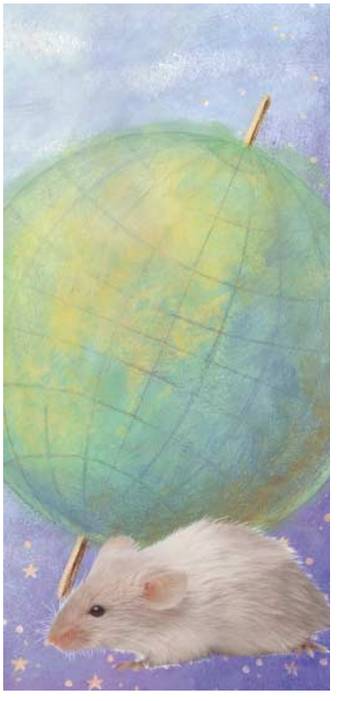
References and links

ORIGINAL RESEARCH PAPERS Reymond, A. *et al*. Human chromosome 21 gene expression atlas in the mouse. *Nature* **420**, 582–586 (2002) | The HSA21 Expression Map Initiative. A gene expression map of human chromosome 21 orthologues in mouse. *Nature* **420**, 586–590 (2002)

WEB SITES

Supplementary information for Reymond, A. *et al.*: http://www.tigem.it/ch21exp

Supplementary information for The HSA21 Expression Map Initiative: http://chr21.molgen.mpg.de/data



IN BRIEF

DEVELOPMENTAL BIOLOGY

BMP4 initiates human embryonic stem cell differentiation to trophoblast.

Xu, R.-H. *et al. Nature Biotechnol.* **20**, 1261–1264 (2002)

Although embryonic stem (ES) cells are well known for their controversial application in transplantation therapy, they are also important for understanding the development and function of different cell types. This paper shows that human ES cells can be induced to differentiate into trophoblast — which develops before the three embryonic germ layers — by bone morphogenetic protein 4. This result reinforces the value of human ES-cell research as mouse and human trophoblast are significantly different and mouse ES cells differentiate poorly into trophoblast.

DISEASE GENETICS

A defect in a novel Nek-family kinase causes cystic kidney disease in the mouse and in zebrafish.

Liu, S. *et al. Development* **129**, 5839–5846 (2002)

Mice with the juvenile cystic kidney (*jck*) mutation are good models for the human disorder polycystic kidney disease (PKD). The authors have positionally cloned the *jck* mutation, which corresponds to a point mutation in the kinase-encoding gene *Nek8*. The role of Nek8 in PKD was confirmed by *in vitro* overexpression studies and the polycystic phenotype of zebrafish in which the Nek8 orthologue was inactivated using antisense oligonucleotides. Gene function can therefore be successfully investigated using comparative analyses.

FUNCTIONAL GENOMICS

Integrating interactome, phenome, and transcriptome mapping data for the *C. elegans* germline.

Walhout, A. J. M. *et al. Curr. Biol.* **12**, 1952–1958 (2002)

Gene clustering based on RNAi phenotypes of ovary-enriched genes in *C. elegans*.

Piano, F. *et al. Curr. Biol.* **12**, 1959–1964 (2002)

Integration of functional genomics approaches and protein–protein interaction data on a global scale provides an opportunity for the thorough testing of biological hypothesis. Such an integrative analysis has already been carried out in yeast, and now Walhout *et al.* have done a similar analysis for the *Caenorhabditis elegans* germline. They have combined a protein–protein interaction map (obtained from yeast two-hybrid data for 600 germline-expressed transcripts) with large-scale phenotypic analysis of germline-specific genes (based on RNA interference (RNAi) data from Piano *et al.*) and transcriptome profiling (obtained by clustering transcripts from several expression profiling experiments). They find that, in a quarter of germline interactions, both partners have RNAi phenotypes that are either embryonic lethal or masculinizing, indicating that these interactions might be most relevant for germline biology.