

# Three's company

Kerstin Lindblad-Toh

Publication of the rat genome sequence will not only advance physiological studies in this paragon of laboratory animals, but also greatly enhance the power of comparative research into mammalian genomes.

On page 493 of this issue<sup>1</sup> there appears a description of the third mammalian genome to be sequenced and analysed — that of the Brown Norway rat. The excitement generated by this research has a number of roots, not least the fact that the rat is an important model for learning about human physiology and disease; knowing its genome sequence will greatly enhance our ability to associate genes and mutations with traits and diseases. Comparative mammalian genomics will also benefit, because having the rat, mouse<sup>2</sup> and human<sup>3,4</sup> genome sequences will allow us to explore which characteristics are specific to rodents and which are shared by all mammals.

Rats (*Rattus norvegicus*; Fig. 1) were established as a model for learning about human physiology and disease in the early 1800s. In the 1900s, they ceded some of their popularity to mice, which are smaller, quicker to breed and easier to manipulate genetically. Still, the rat never fell out of favour in areas where its larger body size and physiological similarity to humans are important, including pharmacological studies — testing the effects and toxicity of drugs. (The history and importance of rats as lab animals is discussed in the News Feature on page 464 of this issue<sup>5</sup>.)

As with mice, there are several hundred inbred or partially inbred rat strains, which differ in many physiological measures, such as blood pressure or body weight, or in their propensity to diseases, such as diabetes mellitus or arthritis. Studies of these different strains have enabled researchers to associate many such traits with specific large chromosomal regions in the rat genome. Many of the characteristics are quantitative traits — they show a graded spectrum of variation from individual to individual — and have complex inheritance patterns, being controlled by more than one gene; they are detected in the genome as quantitative trait loci (see the rat genome database<sup>6</sup>). But the causative genes at these loci are, in many cases, unknown. Today's publication by the Rat Genome Sequencing Project Consortium of the complete sequence of the Brown Norway rat<sup>1</sup> will accelerate the pace of discovery of genes, alleles (gene variants) and mutations associated with these physiological properties and disorders. In fact, the identification of several disease-associated genes —



Figure 1 Model animal — the lab rat.

including ones involved in a cerebellar defect<sup>7</sup> — has already benefited.

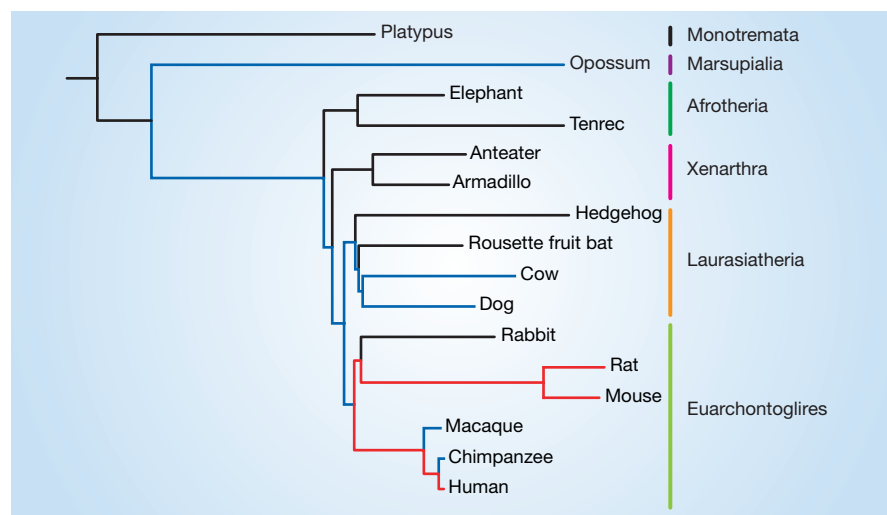
To take full advantage of the wealth of physiological variation among rat strains, researchers also need a map of the genetic variation. The new sequence will serve as the basis for this map, on which researchers will mark the sequence positions that vary between the Brown Norway rat and other strains. As the consortium discusses<sup>1</sup>, a pilot project to identify variations in the coding portions of genes (single nucleotide polymorphisms) is ongoing. A continued effort would generate maps that show which portions of the genome are highly similar and which are quite different between strains — similar to the 'haplotype maps' that are under way for humans<sup>8</sup> and mice<sup>9</sup>.

Beyond its value in unlocking the treasures of rat biology, analysis of the rat genome sequence will also help biologists to identify and understand human genes and gene regulation, and to study mammalian genome evolution. On the evolution front, the consortium's comparison<sup>1</sup> of the rat and human genomes reveals a strikingly large amount of genome shuffling. The previous comparison of the mouse and human genomes<sup>2</sup> noted that most genes have the same immediate neighbours in both species, although blocks of genes have been shuffled considerably. Including the rat genome confirms the suspicion that most of the 300 flips

and swaps between rodents and humans happened in the rodent lineage after it split from the common ancestor with humans, and that 50 of these flips are also different between rats and mice.

Similarly, the rate of change of individual bases seems to be much higher in the rodent than the human lineage, judging by how much variation is seen within sequence repeats that are assumed to have been present in the common ancestor. Together, these findings hint that the rodent genomes are more dynamic than the human genome, evolving roughly three times as fast. This raises the question of whether the mutation rate is different in rodents — or whether it is simply that smaller animals have a shorter generation time and hence go through more mutations in eggs and sperm (mutations that are therefore passed on to the next generation) in the same time frame<sup>10</sup>.

To translate findings from a model organism to human medicine, it is necessary to correlate genes and potential disease-associated mutations across species. So the broad similarity in the number, order and sequence of genes between rodents and humans is reassuring, as is the fact that as many as 90% of rat genes have matches in both humans and mice<sup>1</sup>. This number is higher than the 80% reported when comparing mice and humans<sup>2</sup>. The difference is probably due to a better cataloguing of



**Figure 2 Mammalian evolution and genome sequencing.** This evolutionary tree shows the position of mammals whose genomes have been sequenced and analysed (red) or are being sequenced (blue). Humans, mice and rats come from the same major clade (Euarchontoglires), dogs and cows from a further clade (Laurasiatheria). But two major clades of placental mammals remain unsampled (Xenarthra and Afrotheria). Note also the fast evolutionary rate (represented by a longer branch length) of rats and mice compared with humans.

genes and pseudogenes (copies of genes that are no longer functional) in the rat genome — possibly because of comparison across the three genomes.

The 10% of genes that are not present as one-to-one pairs between species belong to gene families that have expanded in number in one of the species. Many of these gene families are involved in olfaction, immunology or reproduction, and can easily be associated with biological features that make each species unique; for instance, humans rely less than rodents on their sense of smell, and so have fewer olfactory-receptor genes. Of particular interest for pharmacological studies are the detoxifying P450 genes, of which mice and rats have more than humans. So, it may be more difficult than thought to use the toxicity of drugs in rats as a guide to their toxicity in humans — because rats may be better at removing toxins from their system. Still, having the genome sequence makes it possible to assess whether the P450 genes acting on a specific drug are present in both rats and humans. Moreover, this gene number is more similar between rats and humans than between mice and humans, which could help explain rats' superiority over mice in pharmacology.

In the initial comparison between the mouse and human genomes<sup>2</sup>, the high degree of sequence conservation found hinted that some 5% of each genome contains functional elements — genes, regulatory elements and so on. Including the rat genome has significantly increased the resolution of these analyses<sup>1</sup>. The precision of gene prediction (annotation) will continue to increase with each vertebrate genome that is sequenced, and with the availability of other resources, such as the Mammalian Gene Collection<sup>11</sup>,

that catalogue physical evidence of gene transcription. This is encouraging, because we are far from knowing every gene in the human genome, and even further from knowing how each gene is regulated. Using comparative sequence analysis, researchers can draw on the experiments performed by evolution, where functional elements are

conserved and neutral sequences — those sequences that have no particular function — change over time.

To accomplish this in the next few years, a working group<sup>12</sup> has been charged with developing a plan to select the appropriate species and sequencing strategy by which to identify all the functional elements in the human genome. Several other mammalian genomes (those of the chimpanzee, macaque, dog, cow and opossum; Fig. 2) are currently being sequenced, and when completed will further enhance the three mammalian genomes that we already have. The high quality of the rat genome sequence<sup>1</sup>, and the speed with which it was produced, leads us to expect great things in the near future.

Kerstin Lindblad-Toh is at the Broad Institute, 320 Charles Street, Cambridge, Massachusetts 02141, USA.

e-mail: kersli@broad.mit.edu

1. Rat Genome Sequencing Project Consortium *Nature* **428**, 493–521 (2004).
2. Mouse Genome Sequencing Consortium *Nature* **420**, 520–562 (2002).
3. International Human Genome Sequencing Consortium *Nature* **409**, 860–921 (2001).
4. Venter, J. C. et al. *Science* **291**, 1304–1351 (2001).
5. Abbott, A. *Nature* **428**, 464–466 (2004).
6. <http://rgd.mcw.edu>
7. Kuramoto, T., Kuwamura, M. & Serikawa, T. *Brain Res. Mol. Brain Res.* **122**, 103–108 (2004).
8. The International HapMap Consortium *Nature* **426**, 789–796 (2003).
9. Wade, C. M. et al. *Nature* **420**, 574–578 (2002).
10. Weinreich, D. M. *J. Mol. Evol.* **52**, 40–50 (2001).
11. <http://mgc.nci.nih.gov>
12. <http://www.genome.gov/1000218>

## Earth science

# Inside history in depth

David Stevenson

The history of how Earth's interior evolved, and how it accounts for many aspects of our planet's behaviour, remains largely unwritten. Taking water into account could well help to explain a great deal more.

The basic divisions of Earth's internal structure (crust, mantle and core) have been known for a long time. But the evolutionary path that gave us this structure, and that provides the dynamics of plate tectonics, volcanism and magnetic-field generation, remains poorly understood. Why do we have plate tectonics? What is the nature and extent of melting deep within Earth? How does the core manage to keep generating such a richly complex magnetic field?

These were among the questions debated on the first day of a workshop\* organized under the auspices of the CSEDI (Cooperative Study of the Earth's Deep Interior), a National Science Foundation programme that is now eight years old. From the talks and ensuing discussions among deep-Earth devotees, it is evident that we need a better

knowledge of the processes that govern deep-Earth history, and the material parameters that control those processes, before any kind of 'standard model' can be constructed — much though such a model may be desired by some participants.

A simple view of Earth's evolution invokes the first law of thermodynamics, estimates of radiogenic heat production — that produced by radioactivity — and some straightforward scaling arguments that emerge from our theoretical and experimental understanding of thermal convection. In this view, Earth started hot and cooled through geological time at a rate that closely mimics the decreasing rate at which the radiogenic heat production declines. This decreasing rate arises because the radiogenic heat sources have a variety of half-lives, and the more long-lived sources become increasingly important with time. Earth is

\*CSEDI Science Plan Workshop. La Jolla, California, USA, 22–23 February 2004.