

## UC San Diego Bioinformatics Experts Help Reconstruct The Genomic Makeup Of Our Ancestors As International Consortium Completes Rat Genome Sequencing

March 31, 2004

Doug Ramsey

Scientists have generated and begun to analyze the rat genome, paving the way for comparisons with the two other mammalian genomes sequenced so far - human, and mouse. The primary results of the Rat Genome Sequencing Project Consortium (RGSPC) are presented in the April 1 issue of *Nature*, and an additional thirty manuscripts describing further detailed analyses are contained in the April issue of *Genome Research*.

The cover of *Genome Research* (see end of release) was produced by University of California, San Diego professors Pavel Pevzner and Glenn Tesler and their co-author on the journal paper, Guillaume Bourque of the University of Montreal. It depicts the course of evolution for the X chromosome in humans, rats and mice from a common ancestor over 80 million years ago and, for the first time, reconstructs the genomic architecture of mammalian ancestors. "It contributes to the solution of the so-called original synteny problem in biology," said Pevzner, the Ronald R. Taylor Professor of Computer Science at UCSD's Jacobs School of Engineering. "While scientists routinely find bones that lead to (often unrealistic) reconstructions of dinosaurs and other prehistoric animals on movie screens and in toy stores, this is the first rigorous reconstruction of the genomic makeup of our mammalian ancestors."

Pevzner and Tesler are among the more than 200 co-authors of the *Nature* article, and expanded on their part of the research in a *Genome Research* paper titled "Reconstructing the Genomic Architecture of Ancestral Mammals: Lessons from Human, Mouse and Rat Genomes." "Having the third genome allows us to reconstruct the putative genomic architecture of our mammalian ancestor," said Pevzner. "Our contribution has been to demonstrate how to look at the human, mouse and rat genomes - each roughly three billion letters in length - and then infer the evolutionary earthquakes that shaped their genomic architectures."

Pevzner and his colleagues contend that those earthquakes - major genomic rearrangements - tend to occur at evolutionary hot spots known as breakpoints, which are similar to fault zones where earthquakes are more likely to hit. "We found a few hundred strings of roughly a million letters long, and we specifically focused on those large blocks that are shared between the human, mouse and now rat genomes," said Tesler, an assistant professor of mathematics at UCSD. "After sequencing these three genomes, it is clear that substantial rearrangements in the human genome happen only once in a million years, while the rate of rearrangements in the rat and mouse is much faster."

Comparison of the rat genome to human and mouse allows a unique view of mammalian evolution. The rat data shows about 40 percent of the modern mammalian genome derives from the last common mammalian ancestor. These 'core' one billion bases encode nearly all the genes and their regulatory signals, accounting for the similarities among mammals. "We now have information on all three genomes and we can see how many common architectural blocks we share," said Pevzner. "It is almost like a triangle: in the case of the X chromosome, mouse and rat are about the same distance apart as rat and human, and human and mouse."

The study found the rat genome contains similar numbers of genes to the human and mouse genomes, but at 2.75 billion nucleotides is smaller than human (2.9 billion) and slightly larger than mouse (2.6 billion). Almost all human genes known to be associated with diseases have counterparts in the rat genome and appear highly conserved through mammalian evolution. A selected few families of genes have been expanded in the rat, including smell receptors and genes for dealing with toxins, and these give clues to the distinctive physiology of the species.

"The rat genome allows us to reconstruct the genome's architecture, especially for the sex X chromosome, which doesn't exchange genetic material with the other chromosome," explained Pevzner. "We can come to a very reliable evolutionary scenario and genomic architecture of the X chromosome. So essentially we are solving the original phylogeny problem. - how to reconstruct the genomic architecture of our mammalian ancestor."

The Rat Genome Sequencing Project Consortium is led by the Human Genome Sequencing Center at Baylor College of Medicine (BCM-HGSC) in Houston, in conjunction with the National Heart, Lung and Blood Institute (NHLBI) and the National Human Genome Research Institute (NHGRI). "This is an investment that is destined to yield major payoffs in the fight against human disease," said NIH Director Elias A. Zerhouni, M.D. "For nearly 200 years, the laboratory rat has played a valuable role in efforts to understand human biology and to develop new and better drugs. Now, armed with this sequencing data, a new generation of researchers will be able to greatly improve the utility of rat models and thereby improve human health."

The rat sequence is the third complete mammalian genome to be sequenced to high quality and described in a major scientific publication. Three-way comparisons with the human and mouse genomes will help to resolve details of mammalian evolution. "The sequencing of the rat genome constitutes another major milestone in our effort to expand our knowledge of the human genome," said NHGRI Director Francis S. Collins, M.D. "As we build upon the foundation laid by the Human Genome Project, it's become clear that comparing the human genome with those of other organisms is the most powerful tool available to understand the complex genomic components involved in human health and disease."

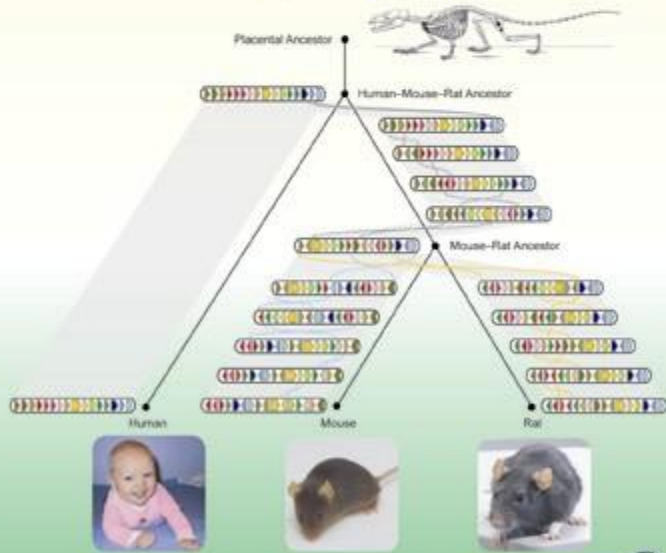
A network of centers generated data and resources for the RGSP, including the BCM-HGSC, Celera Genomics, Genome Therapeutics Corporation, British Columbia Cancer Agency Genome Sciences Centre, The Institute for Genomic Research, University of Utah, Medical College of Wisconsin, The Children's Hospital of Oakland Research Institute, and Max-Delbrück-Center for Molecular Medicine (Berlin). After assembly of the genome at the BCM-HGSC, analysis was performed by an international team, representing over 20 groups in six countries and relying largely on gene and protein predictions produced by the Ensembl project of the EMBL-EBI and Sanger Institute (UK). Funding for the RGSP was largely provided by the NHLBI and the NHGRI with additional private funding provided to the BCM-HGSC by the Kleberg Foundation.

Media Contact: Doug Ramsey, (858) 822-5825

# GENOME RESEARCH

April 2004  
Volume 14 Number 4

## Rat Genome Special Issue



Cold Spring Harbor  
Laboratory Press