

In on the Ground Level

Talented student helps paint the big picture in comparative genomics.

Few students get to witness the birth of a scientific discipline, but in Krishna M. Roskin's case, he even participated in the delivery. As a graduate student at the University of California, Santa Cruz (UCSC), Roskin helped to advance the nascent science of comparative genomics as he worked with a team of more than 20 research groups in six countries on the rat genome sequencing project. The team's results were published in the April 1, 2004, issue of *Nature*.

With the genetic sequence of the brown Norway rat in hand, the team was the first to compare three mammalian genomes (rat, mouse, and human), triangulating data in their search for molecular-scale similarities and differences between species as well as the mechanisms of evolutionary change over the past 12 to 24 million years.

"It was really great to move toward understanding the major evolutionary events and to begin to get a handle on that," says Roskin, who performed many of the computations for the analyses.

Roskin, who was home-schooled while his family traveled throughout Central America and Europe, never imagined that one day he would study genomes. His passions were computer science and mathematics, which became his majors in college. Roskin's life story took a propitious twist in 2001, his junior year, when David Haussler, an HHMI investigator who directs the UCSC Center for Biomolecular Science and Engineering, contacted him.

"I got this grandiose e-mail saying, 'How would you like to join the greatest scientific project of all time?'" Roskin recalls. He found the offer too intriguing to resist, so he joined the project as an undergraduate and then stayed on for graduate school.

First, Roskin helped develop the UCSC Genome Browser for visualizing the map of the human genome, which was roughly decoded in February 2001. Then he worked on the mouse genome draft sequence, published in December 2002. Roskin made the oft-quoted estimate that 5 percent of the mouse and human genomes had been conserved, or preserved, over evolutionary time. The estimate ignited interest in these genomic regions, which are believed to contain much more than just protein-encoding genes; two-thirds of the sequences are thought to encode other kinds of genes as well as elements that regulate transcription and other functions.

"There's now a big hunt for what we think are the regulatory elements, but we couldn't learn much from just two species," Roskin says.

Then came the project to sequence the rat genome. By this time, Roskin no longer had to work 18-hour days at the computer, as he was faster and more efficient in his computations. "I learned to pace myself—it was like a marathon

run," he says. Teams of researchers at several institutions regularly sent him sequence data, and his job was to put it all into a coherent picture and then make comparisons.

At times, Roskin says, he felt like a cartoon switchboard operator madly trying to cope with a deluge of calls. "Sometimes it seemed like I would never get all the numbers to fit together." But he had some expert collaboration. For example, the lead research teams held a conference call each Friday, often quizzing Roskin on his methods. "It was a very 'dynamic' process," he recalls, "but mostly fun."

The biggest challenge for Roskin was joining with Haussler to try to convince everyone that particular genomic regions called "ancestral repeats" were neutral sequences that could be used as baselines for measuring rates of evolutionary change. Once that was satisfactorily proven, the teams used ancestral repeats to determine that mice and rats have experienced more genetic change than have humans, evolving about three times more rapidly.

Surprisingly, they found that the rates of evolution—changes in single bases as well as insertions and deletions of longer sequences—also varied between regions of each genome. Chromosomes

also revealed differences in their rates of change. "It was interesting to see that each chromosome has its own personality and is unlike its neighbors," says Roskin. "We don't know why there are such big differences."

This mystery has captivated Roskin, who hopes to explore it further as he works toward his doctorate in computer science. He's now taking biology courses, including one on chromatin structure, to better prepare for computational questions he wants to answer. "I'm curious whether the places in the genome with lots of evolutionary change will correlate with a certain position in the nucleus," says Roskin. "Are they more accessible to mutagens and irradiation? Will there be a consistent pattern? Data coming out on that would be really exciting."

—KAREN F. SCHMIDT

Krishna Roskin got an e-mail: "How would you like to join the greatest scientific project of all time?"

