BIOINFORMATICS

The Human Genome in 3D, At Your Fingertips

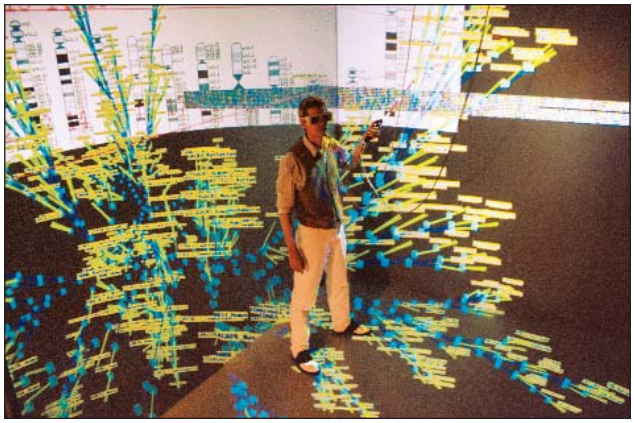
Swishing through thickets of genes with a "light saber" in a virtual-reality room could become a fun way to make sense of the flood of genome data

AMSTERDAM—In a dark chamber in the bowels of Stichting Academisch Rekencentrum Amsterdam (SARA), the Dutch national supercomputing facility, Anton Koning thrusts his hand into a galaxy. "This group here is interesting," he says, pointing to a cluster of glowing points above his head. The computer scientist thumbs a hand-held mouse, and the cluster becomes enmeshed in gleaming red lines. The galaxy is made of genes—not stars—and the red lines, Koning explains, show which are expressed in the same tissues. With another mouse click, small pennants emblazoned with names and functions pop up from each gene. "Human surfactant, human pulmonary," Koning reads. "Potential drug targets for lung disease. So let's see where they sit and compare with known disease loci." Click, click, and the 23 human chromosomes appear behind the galaxy. He pokes at the cluster with what looks like a light saber shooting from his mouse, and a ray extends from each gene to its location on a chromosome.

Saragene, the virtual-reality software driving this galactic tour of the human genome, could become a powerful new tool for scientists struggling to tap its hidden treasures. Even the most seasoned experts are daunted by the mountains of human genome data churned out by sequencing and microarray technologies. By making it possible to display vast amounts of data and interconnections all at once, virtual reality will soon offer an indispensable way to extract meaning from gargantuan data sets, predicts Peter van der Spek, director of bioinformatics drug discovery at Johnson & Johnson Pharmaceutical Research and Development in Beerse, Belgium, who initiated the Saragene project with SARA a year ago. The system could become available widely within a year. "There are some basic software engineering problems to overcome," says Antony Cox, who tackles such problems for the Wellcome Trust Sanger Institute near Cambridge, U.K., "but this is a project with legs."

Koning, bushwhacking through thickets of genes with his light saber, zeroes in on a smaller group of genes. "Now we're getting somewhere," he says, as more pennants appear. Some reveal gene-protein coupled receptors (GPCRs) that are like x's on a treasure map to drug discoverers. GPCR proteins bind signal molecules such as hormones (or imposters such as drugs) and trigger gene expression. Diseases marked by aberrant GPCR activity—schizophrenia, hypertension, and asthma, to name a few—would rate on a pharmaceutical most-wanted list. "Let's see about mouse models," says Koning, as homologous genes appear in his galaxy. What Koning has done in seconds with Saragene, van der Spek explains, could take hours in two dimensions and cannot be viewed all at once.

The virtual-reality chamber Koning is standing in is called a CAVE (for computer-assisted virtual environment), a system designed in the early 1990s and now in use at about 100 facilities, mostly universities. Koning wears a pair of glasses with shutters that open and close 120 times per second in synchrony with the lights that project the galaxy into the chamber. The slightly shifted perspectives flashing alternately at each eye fool the brain into seeing the projection in three dimensions. A sensor on the glasses relays to the computer the position of Koning's head in space, which allows him to move in and around the data intuitively.



May the force be with you. Saragene allows researchers to display vast amounts of interconnected genomic data all at once.

CREDIT: SARA

Because of supercomputing charges, wielding a light saber in a CAVE costs about $1000 per hour. The motivation for Johnson & Johnson, which pays the bills in the Saragene marriage, is to speed up the process of drug discovery. Before "wet-lab" research can begin, van der Spek's bioinformatics team hunts for promising drug targets in silico. By developing this as a public tool with the technical team at SARA, says van der Spek, Johnson & Johnson is shaping Saragene to fit its needs: "It's win-win."

Although the hardware Koning plays with is pricey, most of the data are free. Stored on site at SARA is a copy of Ensembl, a database of annotated genomes maintained in the public domain by the Sanger Institute and the European Bioinformatics Institute, based in Cambridge. The keepers of Ensembl were given a demonstration of Saragene 2 weeks ago and are now considering teaming up with Johnson & Johnson and SARA to add a virtual-reality interface to Ensembl, allowing any facility with a CAVE to jack into the Saragene world. "I found it pretty exciting," says Cox, who would lead the design modifications if Ensembl incorporated a Saragene interface, a decision that could come by next summer.

Cox is particularly attracted to the possibility of using Saragene for comparing gene regions between species, a famously difficult problem in two dimensions. In the search for drug targets, candidate genes are often found with few clues to function, but insight can be gleaned by comparing a gene's region on the human chromosome with the matching chromosome region from another species in which gene functions are better defined. Making such comparisons for more than two species when evolution has shuffled genes across chromosomes becomes very messy on a computer screen. "Stepping into the data" with Saragene, as Koning puts it, makes multispecies comparisons "much easier."

Another enthusiast is Gert Vriend, director of the Dutch Centre for Molecular and Biomolecular Informatics in Nijmegen. Vriend, who recently visited SARA, is interested in using Saragene to study evolutionary biology. "Looking at a phylogenetic tree in a CAVE is really cool. You can see much more information at once," he says. And rather than displaying one lineage at a time, Saragene creates a phylogenetic "forest." "You can put thousands of sequences in one single tree," beams Vriend.

Just as Koning seems to be closing in on a culprit in lung disease, his light saber lodges in chromosome 13. Even supercomputers freeze sometimes. After a reboot, he once again loses himself in a galaxy of genes.

–JOHN BOHANNON