New discovery may unlock human genome

December 29 2004

Goodman laboratory devises technique to explain patterns of gene regulation

An Oregon Health & Science University-led development of a technique for identifying control elements that drive the expression of genes in brain cells could unleash the disease-fighting potential of the much-hailed human genome.
Scientists at the OHSU Vollum Institute, which headed the multidisciplinary study appearing in the Dec. 29 edition of the journal *Cell*, are calling the approach a significant advance in understanding the genome.

The Vollum's director, Richard Goodman, M.D., Ph.D., professor of cell and developmental biology, and biochemistry and molecular biology, OHSU School of Medicine, said the technique could give a critical boost to the new era of genomic discovery set forth when the Human Genome Project was completed early last year.

"The question was how to understand the enormous amount of genomic information that has been generated," Goodman said. "Our approach will help unlock the regulatory control of the genome."

The approach could heighten understanding of the pathways behind genetic aberrations that cause diabetes, Parkinson's disease, heart disease, cancer and other diseases, he said.

The Vollum team's technique, developed in collaboration with scientists

at Brookhaven National Laboratory in Upton, N.Y., and State University of New York, Stony Brook, resulted from an effort by Soren Impey, Ph.D., in Goodman's laboratory to characterize a family of genes regulated by the "cAMP response element binding" protein, or CREB. This well-characterized molecule is among a group of proteins called transcription factors that interact with regulatory elements in DNA that are responsible for increasing or decreasing the level of gene expression in cells.

The technique involves linking DNA from a cell with the transcription factor protein, then isolating the complex through a process called immunoprecipitation. Strips of 21-nucleotide-long DNA are then released from the immunoprecipitated DNA to create "genomic signature tags," which are then identified in the international genome database. The method uncovered about 6,300 regulatory regions that mapped to distinct sites on the genome.

"A subset of these regions highlight novel genes," said Impey, assistant professor of neurology, OHSU School of Medicine, and the study's lead author.

Goodman calls the process "the most comprehensive analysis to date in a metazoan system – that is, a multicellular system – of where transcription factors bind to their genomic targets." It gives scientists a system for mining the entire genome for all the regulatory sites involving a given transcription factor protein.

"You can start to put together a transcriptional map of pathways that are involved in cellular function," he said. "In the past, it's only been possible to look at a very small part of the genome, but now we can look at the whole thing. It's a big step forward."

David Ginty, Ph.D., professor of neuroscience at The Johns Hopkins

University School of Medicine in Baltimore, studies molecular control of growth and survival of neurons in the developing vertebrate nervous system as a Howard Hughes Medical Institute investigator. He said the challenge to exploiting the human genome has been to uncover the relationships between identified genes and to understand how complex patterns of gene expression take place.

But the Goodman lab's discovery, Ginty said, will help scientists understand how transcription factors coordinate complex genetic patterns and, therefore, how different cells are made and how they function.

"The study establishes a beautifully simple approach to identifying mechanisms of complex genomic control," he said. "The method should prove useful for establishing how sets of genes are turned on or off in any given cell type, and how cellular and functional diversity is achieved."

Exploration of the humane genome has been frenzied since the International Human Genome Sequencing Consortium, led in the United States by the National Human Genome Research Institute and the Department of Energy, and The Institute for Genomic Research (TIGR), a private genome sequencing company, announced the completion of the Human Genome Project more than two years ahead of schedule in April 2003. Between 20,000 and 25,000 genes coding for proteins that perform most life functions were found. But there was a problem.

"That's not very many genes," Goodman said. "And so, in a sense, declaring the genome solved was somewhat arbitrary because it's solved when you really understand it. If you look at the genome, or the database that the genome provided, what you have is a bunch of letters and it has to be decoded to understand what those letters mean."

Goodman compared the genome to a phone book in which the names were interspersed "with a lot of nonsense letters," and the names themselves were broken into pieces. "And rather than having 26 letters, there are only four, and they're all mixed up," he said. "It's hard to know where the genes start and stop."

Said Impey: "Although the Human Genome Project identified about 25,000 protein-coding genes, the instruction set that regulates these genes is, for the most part, unknown. This is important because what makes a cancerous cell different from a noncancerous cell is the set of genes that are turned on or off. These instructions or regulatory regions are believed to be far more numerous than genes, but it was not clear how to identify them."

"We developed a novel technique that is able to isolate a comprehensive set of regulatory regions and map them to the entire genome. Our work will help unravel the genomic instruction set that governs how genes are regulated in a given cell type. If one views the genome as a multidimensional puzzle, our method helps make the puzzle a little simpler."

The discovery already has demonstrated its implications for human diseases. Goodman's lab is working with the OHSU Cancer Institute on a cancer-causing oncogene that arises when a rearrangement of chromosomes generates an abnormal transcription factor. "If you had a technique that would allow you to take that factor and identify what its targets are, you would understand why that oncogene causes cancer," Goodman said.

Another project is examining a transcription factor involved in the differentiation of dopamine-producing cells. By identifying the targets of the transcription factor, stem cells differentiated as dopamine cells could be developed to treat Parkinson's disease.

And a project with Markus Grompe, M.D., of the OHSU Oregon Stem Cell Center is studying the transcription factor involved in pancreatic beta cell differentiation.

"This is a factor that drives the expression of insulin, but also other differentiated properties of a beta cell, so if we can identify all those targets, we'll understand something about the nature of the development of a beta cell," Goodman said.

Ginty, of The Johns Hopkins University, said the future could even hold answers to such questions as how a neuron in the brain stores information that forms the basis of memory.

"The future of genome exploration will bring an understanding of how the genome is controlled to yield different cell types of the body and their various functions," he said.

In addition to Goodman and Impey, study collaborators included: Hyunjoo Cha-Molstad, Jami Dwyer and Gregory Yochum, Vollum Institute; Sean McCorkle and John Dunn, Brookhaven National Laboratory; Jeremy Boss, Emery School of Medicine; Shannon McWeeney, OHSU; and Gail Mandel, State University of New York, Stony Brook.

Source: Oregon Health & Science University