

MARCH 18, 2010

Transcription Factors May Dictate Differences Between Individuals

Researchers are only beginning to understand how individual variation in gene regulation can have a lasting impact on one's health and susceptibility to certain diseases. Now, an ambitious survey of the human genome has identified differences in the binding of master regulators called transcription factors to DNA that affect how genes are expressed in different people.

The study, which is published in the March 18, 2010, issue of *Science*, looked at two common transcription factors. HHMI medical research fellow Maya Kasowski and her colleagues in the laboratory of molecular biologist Michael Snyder at Yale University conducted the work with Jan Korbel at the European Molecular Biology Laboratory. Snyder has since joined the faculty at Stanford University.

Transcription factors account for as much as 10 percent of the coding genome in humans and other organisms. When activated, transcription factors switch on or off hundreds or thousands of genes, a cascade that programs cells to grow or divide. "The activity of transcription factors determines what a cell is doing at any given moment," says Kasowski, who was a medical student at Yale when she received her HHMI medical research fellowship. She has since decided to pursue an M.D./Ph.D. degree.

Despite their large numbers and critical role, many aspects of transcription factor biology remain poorly understood. Until now, no one had looked at whether there was any variability in the targets of transcription factors from one person to the next. The current study found a "number of differences between individuals" in the binding sites of two transcription factors, Snyder says.

"We know there are differences in gene expression between people. Understanding the differences in how genes are regulated could help us understand human diversity."

- Maya Kasowski

Transcription factors bind to the human genome within areas of the genome still viewed as a black box—vast stretches of DNA sequence between known genes. Gradually, biologists have found that much of this DNA performs a vital function—helping turn genes on and off in specific situations. Some of the regulatory regions, known as binding regions, serve as handholds for transcription factors.

“We know there are differences in gene expression between people.” Kasowski says. “Understanding the differences in how genes are regulated could help us understand human diversity. But identifying the regulatory DNA that controls expression is much more difficult than looking for differences in the regions of the genome that code for genes.”

For the current study, Kasowski, Snyder, and their colleagues examined two important transcription factors: RNA polymerase II and NFkappaB. RNA polymerase II, which is active in all cells, transcribes DNA into RNA. NFkappaB is activated by stress, plays a key role in immune responses to infections, and has been implicated in several diseases, including cancer.

The team mapped every binding region for these two factors inside the genomes of 10 individuals. To do so, they deployed a new technology that uses chemicals to freeze transcription factors as they bind to the genome. The scientists then sequenced the segment of DNA to which the transcription factor bound. After the team combined the data from all 10 individuals, they found around 19,000 binding regions for RNA polymerase II and another 15,500 binding regions for NFkappaB.

They discovered that the number of transcription factors binding at the different sites often varied near different genes, which in many cases influenced how much of the gene was expressed. Hence, variation in transcription factor binding can help explain why one person may make more of a certain gene product than another, Snyder says. Among any two individuals, the team found that 25 percent of the RNA polymerase II binding regions varied in time or frequency, as did 7.5 percent of the NFkappaB binding regions.

Closer examination of these variable binding regions showed that single-letter differences in the genome—called SNPs—accounted for some of the difference in transcription factor binding. That is, in some of individuals, a single letter change at a certain binding region influenced the likelihood that the transcription factor would bind there. “We found that differences in DNA sequence contributed to differences in transcription factor binding,” Kasowski says. “The more SNPs we found in a particular binding region, the more variation in binding we saw.”

Other, larger differences in the genome, called structural variation, also accounted for a number of the differences in transcription factor binding.

Structural variation happens when large segments of the genome are deleted, duplicated, or inverted. It varies widely among humans, and the role of such variability in human biology is not well understood.

But the new study shows that SNPs and structural variation can either increase or decrease transcription factor binding, and, hence, the amount of a protein that gets made from a particular gene. “We found that about one third of the differences in binding was caused by SNPs and structural variation,” Snyder says. “This is the first time anyone has shown that SNPs and structural variation affect large number of regulatory elements that control gene expression. Normally, people look at differences in the gene themselves rather than in the regulatory regions, because they are difficult to identify.”

The study also reports differences in binding of RNA polymerase II and NFkappaB near genes implicated in many major diseases, including type 1 diabetes, lupus, chronic lymphatic leukemia, schizophrenia, asthma, Crohn’s disease, and rheumatoid arthritis. “Variation in the regulation of genes might eventually help account for some of the varying susceptibility to diseases we see in the population,” Kasowski says.

In addition to looking at humans, Kasowski, Snyder and their colleagues looked at transcription factor binding for a single chimpanzee. The study shows that 32 percent of RNA polymerase II binding regions differed between the humans in the study and the chimp. Snyder says that he included the chimp out of curiosity to see how transcription factor binding might account for differences between ourselves and our closest genetic cousin. But the 32 percent difference between chimps and humans was not that much larger than the 25 percent difference in RNA polymerase II binding found among two individuals.

Still, Snyder says that the study opens a new genomic frontier for biologists. “Only about two percent of our DNA codes for genes,” he says. “Studying the rest of the genome, including gene regulation and transcription factors, is the next wave in understanding human variation.”