MedicalXpress

Epigenetic marker 5hmC opens door to studying its role in developmental disorders and disease

February 4 2013

Nearly every cell in the human body carries a copy of the full human genome. So how is it that the cells that detect light in the human eye are so different from those of, say, the beating heart or the spleen?

The answer, of course, is that each type of cell selectively expresses only a unique suite of genes, actively silencing those that are irrelevant to its function. Scientists have long known that one way in which such gene-silencing occurs is by the chemical modification of cytosine—one of the four bases of DNA that write the genetic code—to create an "epigenetic" marker known as 5-methylcytosine (5mC). Appropriate placement of this marker is essential to many normal biological processes, not least embryonic development. Conversely, its erroneous distribution contributes to the evolution of a broad range of cancers.

But 5mC is not the only epigenetic marker on the genomic block. About three years ago at Rockefeller University, Skirmantas Kriaucionis, currently a Ludwig researcher based at Oxford University, and Nathaniel Heintz, of Rockefeller University, discovered that a second modification of cytosine that converts it into 5-hydroxymethylcytosine (5hmC) seems to play a similarly vital role in the selective expression of the genome. Since then, researchers have scrambled to figure out what precisely that role might be. In a recent issue of the journal *Cell*, Heintz, Kriaucionis and colleagues report that the 5hmC marker has an effect on gene expression opposite to that of 5mC, and identify how its signal is

detected and broadly interpreted in the healthy brain cells of mice. Since changes in the distribution of 5hmC are known to take place in a broad range of tumor cells, these findings could prove to be of great value to cancer research.

To begin, the team mapped where exactly 5hmC is found across the genomes of three types of healthy mouse neural cells. They discovered that it is largely associated with DNA that is loosely looped about its protein scaffolding in the nucleus. The 5mC signal, meanwhile, is predominantly located on more tightly packed, less accessible stretches of DNA. It is on the loosely packed DNA that most gene expression takes place.

In line with that finding, they discovered that the 5hmC marker was scattered over regions of the genome where genes are being expressed at high levels. And in concurrence with other studies, they found that the 5mC signal was mainly located on silent islands of the genome.

When the researchers peered more closely at the DNA sequence, they further discovered that 5hmC is largely planted within gene bodies—the parts of genes that encode proteins—not in intervening sequences that mark where the coding regions begin (promoters) and determine how avidly the genes are expressed (enhancers). This surprised them, as it is in such sequences that one would expect an on switch for a gene to be located. "It's easy to see how a modification of enhancer or promoter regions might affect gene expression," says Kriaucionis, PhD, and an assistant member of the Ludwig Institute at Oxford University.

"Indeed, we know 5hmC is found in those sequences in embryonic stem cells, which give rise to the whole body. But it is much less clear how the placement of this marker on the gene body—which contains the instructions for making a protein, not the sequences that determine when or how much of it is made—could have such a pronounced effect on

gene activation."

Kriaucionis and colleagues in Nathaniel Heintz' team also discovered that in places where 5hmC is common, 5mC is found at low levels, and vice versa. Further, the patterns and ratios of the two markers and the genes highly expressed in each of the cells assessed varied dramatically. This implies that preexisting molecular signals that are unique to each kind of cell are vital to determining where exactly the 5hmC marker is placed across the genome. Changes in such signals are likely to play a role in the generation of disease, including cancers.

Next, the researchers sought to determine how the 5hmc signal is read. "We must understand what 5hmC is doing in normal cells," says Kriaucionis. "Understanding that will help us trace the process by which genes are incorrectly expressed in disease." The researchers discovered that a molecule known as methyl-CpG-binding protein (MeCP2), which binds 5mC, also binds 5hmC in the cells they studied. Mutations in this protein are known to contribute to Rett Syndrome, a developmental disorder that varies in severity depending on how precisely MeCP2 has been altered. Most importantly one of those mutant MeCP2 proteins was capable of binding 5mC, but not 5hmC. That mutant is known to cause relatively less severe cognitive and speech deficits in Rett patients.

Finally, investigators show in their paper that MeCP2 binding to 5hmC drives gene expression by making DNA more accessible to the molecular machinery that decodes genetic information—while its association with 5mC has the opposite effect. Though the current study focused on healthy mouse neural tissue, the unraveling of 5hmC's tangled role in genome expression is likely to be of value to cancer research. Other researchers have shown that 5hmC is highly depleted in several cancers, including the blood cancer acute myeloid leukemia and the blood disorder myelodysplastic syndrome, which can progress to cancer. This depletion is accompanied by the disruption of a tumor suppressor gene

named TET2, which encodes a protein that creates the 5hmC signal on DNA.

Kriaucionis's lab is now assessing what role 5hmC plays in the development of different types of blood cells, with the aim of deciphering how its loss contributes to the generation of blood cancers.

Provided by Ludwig Institute for Cancer Research