New maps of chemical marks on DNA pinpoint regions relevant to many developmental diseases

July 30 2020

Credit: CC0 Public Domain

In research that aims to illuminate the causes of human developmental disorders, Salk scientists have generated 168 new maps of chemical

marks on strands of DNA—called methylation—in developing mice.

The data, published July 29, 2020, in a special edition of *Nature* devoted to the ENCODE Project (a public research effort aimed at identifying all functional elements in the human and mouse genomes), can help narrow down regions of the human genome that play roles in diseases such as schizophrenia and Rett syndrome. The paper's authors are also on two additional papers in the special edition.

"This is the only available data set that looks at the methylation in a developing mouse over time, tissue by tissue," says senior author and Howard Hughes Medical Institute Investigator Joseph Ecker, a professor in Salk's Genomic Analysis Laboratory. "It's going to be a valuable resource to help in narrowing down the causal tissues of human developmental diseases."

While the sequence of DNA contained in every cell of your body is virtually identical, chemical marks on those strands of DNA give the cells their unique identities. The patterns of methylation on adult brain cells, for instance, are different than those on adult liver cells. That's in part because of short stretches in the genome called enhancers. When transcription factor proteins bind to these enhancer regions, a target gene is much more likely to be expressed. When an enhancer is methylated, however, transcription factors generally can't bind and the associated gene is less likely to be activated; these methyl marks are akin to applying the hand brake after parking a car.

Researchers know that mutations in these enhancer regions—by affecting the expression levels of a corresponding gene—can cause disease. But there are hundreds of thousands of enhancers and they can be located far from the gene they help regulate. So narrowing down which enhancer mutations may play a role in a developmental disease has been a challenge.

In the new work, Ecker and collaborators used experimental technologies and computational algorithms that they previously developed to study the DNA methylation patterns of cells in samples of a dozen types of tissues from mice over eight developmental stages.

"The breadth of samples that we applied this technology to is what's really key," says first author Yupeng He, who was previously a Salk postdoctoral research fellow and is now a senior bioinformatics scientist at Guardant Health.

They discovered more than 1.8 million regions of the mouse genome that had variations in methylation based on tissue, developmental stage or both. Early in development, those changes were mostly the loss of methylation on DNA—akin to removing the brake on gene expression and allowing developmental genes to turn on. After birth, however, most sites became highly methylated again, putting the brakes on gene expression as the mouse approaches birth.

"We think that the removal of methylation makes the whole genome more open to dynamic regulation during development," says He. "After birth, genes critical for early development need to be more stably silenced because we don't want them turned on in mature tissue, so that's when methylation comes in and helps shut down the early developmental enhancers."

In the past, many researchers have studied methylation by homing in on areas of the genome near genes called CpG islands—sections of DNA that have a lot of cytosine and guanine base pairs in them, since typical methylation occurs when a methyl is added to a cytosine that's followed by a guanine. However, in the new work, He and Ecker showed that 91.5 percent of the methylation variations they found during development far away from CpG islands.

"If you only look at those CpG island regions near genes, as many people do, you'll miss a lot of the meaningful DNA changes that could be directly related to your research questions," says He.

To show the utility of their new data set, the researchers looked at genetic variations that had been linked to 27 human diseases and disorders in previous genome-wide association studies (GWAS). They found associations between some human disease mutations and tissue-specific methylation patterns in corresponding regions of the mouse genome. For instance, mutations associated with schizophrenia were more likely to be found in suspected gene control regions in the mouse genome that undergo methylation changes in an area of the brain called the forebrain during development. Such patterns could help other researchers narrow down which mutations found in a GWAS they should focus on.

Other researchers on the study were Manoj Hariharan, Chongyuan Luo, Rosa Castanon, Joseph Nery, and Huaming Chen of Salk; David Gorkin, Ah Young Lee, Yuan Zhao, Hui Huang, Rongxin Fang, Bin Li and Bing Ren of the University of California San Diego; Diane Dickel, Axel Visel and Len Pennacchio of Lawrence Berkeley National Laboratory; and Brian Williams, Diane Trout and Henry Amrhein of the California Institute of Technology.