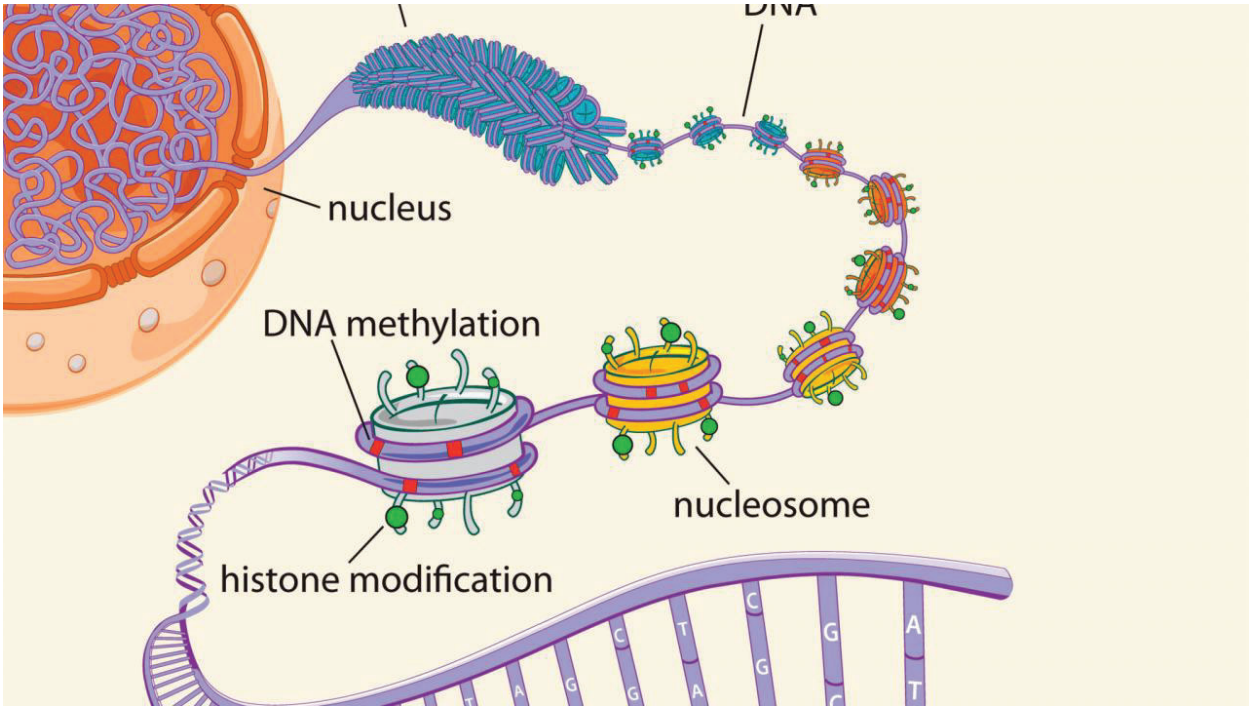

The genome comes tightly packed into chromosomes, and epigenetics helps unwind DNA so its genes and regulatory regions can become active.

Adapted from Illustration by Sigrid Knemeyer, previously adapted by Lauren Soloman

Massive project maps DNA tags that define each cell's identity

By Elizabeth Pennisi Feb. 18, 2015 , 7:00 AM

Our DNA may encode all the instructions needed to build the human body and keep it running, but each of our cells must follow just a subset of those instructions in order for the body to work properly. Thanks to the $300 million, National Institutes of Health (NIH) Roadmap Epigenomics Project, researchers have **now identified most of the chemical tags on DNA and its associated proteins that influence gene function and help define more than 100 different kinds of human cells**. The knowledge of these so-called epigenetic modifications has already led to new insights into Alzheimer's disease, cancer, and development.

Hundreds of publications have already used the epigenetic data the project has gathered over the past 8 years, and today, more than 20 reports appear in *Nature* and its associated journals. The flood of studies marks the official end of the project, but

researchers predict they will be using its information on how genes are regulated for years to come.

"The reality is that it's very challenging to understand the genetic basis of disease," says Janine LaSalle, an epigeneticist at the University of California (UC), Davis, who was not involved in the project. Because of its complexity, "it's really important to have the Roadmap."

When the human genome was deciphered more than a decade ago, researchers already knew that there was a layer of control to gene function beyond the sequence of the DNA: an epigenome. "Epi" means "on, upon, at, by, or near" and includes molecules that latch on to the DNA or stick to the proteins that surround DNA and make up the chromosome. These chemical tags include so-called methyl or acetyl groups; they don't change the DNA sequence itself, but they can make the DNA nearby fold or unfold, blocking access to a gene so it can't turn on, or exposing a gene so it can be activated. In doing so, the epigenome helps determine the genes a cell can use, turning it, for example, into a liver cell or neuron.

"By studying the epigenomes of cells, we come to learn about how they became what they are," says Christopher Glass, a genomics scientist at UC San Diego who was not involved in the work. And unlike a person’s genome sequence, the epigenome can readily change in response to diet, disease, and environment factors, allowing cells—and therefore the body—to adjust to new conditions.

But until recently, epigenomes were hard to decipher. An earlier NIH-funded project called ENCODE helped develop several efficient techniques for determining epigenomes by marking the places along each chromosome where methyl groups and other epigenetic modifications attached. But that project focused on so-called cell lines, immortalized versions of cells that can grow indefinitely in lab dishes. Those cells don't necessarily accurately reflect what's happening in normal cells, Glass says, so the Roadmap project instead examined samples taken from the body directly, such as cells from the heart, liver, kidney, muscle, intestines, skin, fat, and brain, as well as fetal tissue. In addition, some of the project's investigators tested several kinds of stem cells and even sampled stem cells as they differentiated into nerve cells or other tissues.

When the teams were done mapping and comparing the epigenomes of these 111 tissues and cells, they were able to define active and inactive regions of each cell type's genome. They were also able to pinpoint active regulatory DNA called enhancers, which are under the sway of epigenetic tags and control where and when different genes are expressed. In any given cell type, **about 5% of the epigenome regulates gene expression**, report Manolis Kellis, a computational scientist at the Broad Institute in Cambridge, Massachusetts, and his colleagues online today in *Nature*. Further analyses indicated that enhancers specific to particular cell types "appear to be where the major action is for disease," LaSalle says. These enhancers are often located where large-scale population studies of genetic variation—so-called genome-wide association studies—pinpoint an increased risk of disease.

Kellis and his colleagues also found a surprise in enhancers involved with Alzheimer's disease. They looked at the epigenome and patterns of gene expression in the brain cells of a mouse as they underwent degeneration akin to Alzheimer's disease. They then compared what they found with the epigenome and gene activity of the brains of Alzheimer's patients who recently died. The analysis showed that enhancers and genes involved in nervous system activity tended to become less active in both the diseased mouse and the human, but that enhancers that turn on immune response genes revved up.

"That is not what we would have predicted in Alzheimer's and may be a new wrinkle that might be helpful" for understanding or treating the disease, says Henk Stunnenberg, an epigenomicist at Radboud University in Nijmegen, the Netherlands, who was not involved with the work. Adds Glass: "The brain studies suggest that this could be a very valuable approach for approaching many complex diseases going forward."

Bing Ren, a molecular geneticist at the San Diego branch of the Ludwig Institute for Cancer Research, and colleagues studied another important aspect of the epigenome. To fit into a nucleus, the DNA is tightly folded but unfurls at some spots to allow specific genes to be activated. His group has mapped these folding patterns in developing stem cells and found them to be changing quite a bit in some places. **The work indicates that enhancers play an important role in shaping the 3D structure of the DNA**, the team reports.

In many ways, though officially over, the Roadmap Epigenomics Project is just a beginning. There are hundreds of other cell types waiting to be mapped, as well as to be studied under different conditions and at different time points, and an international consortium has set it sights on generating 1000 epigenomes in the next few years. And there's much left to be done with the information already gathered, Glass says. Although "this effort generated a huge amount of data, they've left most of the work of testing the hypotheses [generated] for people who will read these papers."

Posted in:

- Biology
- Health

doi:10.1126/science.aaa7875

Elizabeth Pennisi

Liz is a senior correspondent covering many aspects of biology for *Science*.

- **Email Elizabeth**
- **Twitter**