http://www.nytimes.com/2009/02/24/science/24chromatin.html?_r=1&ref=science

February 24, 2009
From One Genome, Many Types of Cells. But How?
By NICHOLAS WADE

Secrets of the Cell
A Cell's Many Faces
This is the first in a series of occasional articles on a frontier of biology - the workings of the cell.

One of the enduring mysteries of biology is that a variety of specialized cells collaborate in building a body, yet all have an identical genome. Somehow each of the 200 different kinds of cells in the human body — in the brain, liver, bone, heart and many other structures — must be reading off a different set of the hereditary instructions written into the DNA.

The system is something like a play in which all the actors have the same script but are assigned different parts and blocked from even seeing anyone else's lines. The fertilized egg possesses the first copy of the script; as it divides repeatedly into the 10 trillion cells of the human body, the cells assign themselves to the different roles they will play throughout an individual's lifetime.

How does this assignment process work? The answer, researchers are finding, is that a second layer of information is embedded in the special proteins that package the DNA of the genome. This second layer, known as the epigenome, controls access to the genes, allowing each cell type to activate its own special genes but blocking off most of the rest. A person has one genome but many epigenomes. And the epigenome is involved not just in defining what genes are accessible in each type of cell, but also in controlling when the accessible genes may be activated.

In the wake of the decoding of the human genome in 2003, understanding the epigenome has become a major frontier of research.

Since the settings on the epigenome control which genes are on or off, any derangement of its behavior is likely to have severe effects on the cell.

There is much evidence that changes in the epigenome contribute to cancer and other diseases. The epigenome alters with age — identical twins often look and behave a little differently as they grow older because of accumulated changes to their epigenomes. Understanding such changes could help address or retard some of the symptoms of aging. And the epigenome may hold the key to the dream of regenerative medicine, that of deriving safe and efficient replacement tissues from a patient's own cells.

Because the epigenome is the gateway to understanding so many other aspects of the cell's regulation, some researchers have criticized the "piecemeal basis" on which it is

being explored and called for a large epigenome project similar to the $3 billion program in which the human genome was decoded. At present the National Institutes of Health has a small, $190 million initiative, called the Epigenome Roadmap, with the money going to individual researchers.

As is often the case, academic researchers oppose a large, centralized project if the money seems likely to come out of their grants. But it is also true that such projects often fail unless carefully timed and thought out.

“Definitely this is a genome-sized thing, and I believe it will have benefits beyond what are foreseen at present,” says Richard A. Young, a biologist at the Whitehead Institute in Cambridge. But Steven Henikoff of the Fred Hutchinson Cancer Research Center in Seattle says the present methods for studying the epigenome are not yet ready to be scaled up. “It’s too early to mount a technology development that would be large scale,” he says.

The epigenome consists of many million chemical modifications, or marks as they are called, that are made along the length of the chromatin, the material of the chromosomes. The chromatin includes the double-stranded ribbon of DNA and the protein spools around which it is wound. Some of the marks that constitute the epigenome are made directly on the DNA, but most are attached to the short tails that stick out from the protein spools. Marks of a certain kind generally extend through a large region or domain of the DNA that covers one or more genes. They are recognized by chromatin regulator proteins that perform the tasks indicated by each kind of mark.

In some marked domains, the regulators cause the DNA to be wound up so tightly that the genes are permanently inaccessible. The center and tips of the chromosomes are sites of such repressive domains. So is one of the two X chromosomes in every woman’s cells, a step that ensures both male and female cells have the same level of activity of the X-based genes.

In other domains, the marks are more permissive, allowing the gene regulators called transcription factors to find their target sites on the DNA. The transcription factors then recruit other members of the complex transcription machinery that begins the process of copying the genes and making the proteins the cell needs. A third kind of domain must be established ahead of the transcription machinery to let it roll along the DNA and transcribe the message in the underlying gene.

Only a handful of domains are known so far, so it is something of a puzzle that more than 100 kinds of marks have been found in the epigenome, along with specialist protein machines that attach or remove each mark. Some biologists think so many marks are needed to specify a few kinds of domain because the system is full of backups.

The epigenome’s role in marking up the genome seems to have been built on top of a more ancient packaging role. The packaging would have been needed by one-celled organisms like yeast that keep their genome in a special compartment, the nucleus. For

multi-celled organisms to evolve, the chromatin's packaging system presumably adapted during the course of evolution to index the genome for the needs of different types of cell.

The DNA packaging system alone is an extraordinary technical feat. If the nucleus of a human cell were a hollow sphere the size of a tennis ball, the DNA of the genome would be a thin thread some 24 miles long. The thread must be packed into the sphere with no breakages, and in such a way that any region of it can be found immediately.

The heart of the packaging system is a set of special purpose proteins known as histones. Eight histones lock together to form a miniature spool known as a nucleosome. The DNA twists almost twice round each nucleosome, with short spaces in between. Some 30 million nucleosomes are required to package all the DNA of ordinary cells.

For years, biologists assumed that the histones in their nucleosome spools provided a passive framework for the DNA. But, over the last decade, it has become increasingly clear that this is not the case. The histone tails that jut out from the nucleosomes provide a way of marking up the genetic script. Although one kind of mark is attached directly to the bases in the DNA, more than a hundred others are fixed onto specific sites on the histones' tails. When the DNA has to replicate, for cell division, the direct marks pass only to the two parent strands and all the nucleosomes are disassembled, yet the cell has ingenious methods for reconstituting the same marks on the two daughter genomes. The marks are called epigenetic, and the whole system the epigenome, because they are inherited across cell division despite not being encoded in the DNA.

How is the structure of the epigenome determined? The basic blueprint for the epigenomes needed by each cell type seems to be inherent in the genome, but the epigenome is then altered by other signals that reach the cell. The epigenome is thus the site where the genome meets the environment.

The organization of the epigenomes seems to be computed from information inherent in the genome. "Most of the epigenetic landscape is determined by the DNA sequence," says Bradley Bernstein, a chromatin expert at Massachusetts General Hospital. The human genome contains many regulatory genes whose protein products, known as transcription factors, control the activity of other genes. It also has a subset of master regulatory genes that control the lower-level regulators. The master transcription factors act on each other's genes in a way that sets up a circuitry. The output of this circuitry shapes the initial cascade of epigenomes that are spun off from the fertilized egg.

The other shapers of the epigenome are the chromatin regulators, protein machines that read the marks on the histone tails. Some extend marks of a given kind throughout a domain. Some bundle the nucleosomes together so as to silence their genes. Others loosen the DNA from the nucleosome spools so as to ease the path of the transcription machinery along a gene.

Biologists had long assumed that once the chromatin regulators had shaped an epigenome, their work could not be undone because a cell's fate is essentially irreversible. But a remarkable experiment by the Japanese biologist Shinya Yamanaka in June 2007 underlined the surprising power of the master transcription factors.

By inserting just four of the master regulator genes into skin cells, he showed the transcription factors made by the genes could reprogram the skin cell's epigenome back into that of the embryonic cell from which it had been derived. The skin cell then behaved just like an embryonic cell, not a skin cell. Until then, biologists had no idea that the epigenome with its millions of marks could be recast so simply or that transcription factors could apparently call the shots so decisively.

But subsequent research has shown the chromatin regulators are not pushovers. Only one in a million of the skin cells treated with the four transcription factors reverts fully to the embryonic state. Most get stuck in transitional states, as if the chromatin regulators are resisting a possibly cancerous change in the cell's status. "The take-home story is that yes, the transcription factors are really critical players in determining cellular state, but epigenetics is important, too," Dr. Bernstein said.

The ideal of regenerative medicine is to convert a patient's normal body cells first back into the embryonic state, and then into the specific cells lost to disease. But to prepare such cells safely and effectively, researchers will probably need to learn how to control and manipulate the chromatin of the epigenome as well as the transcription factors that shape cell identity.

The treatment of many diseases may also lie in drugs that manipulate the epigenome. Rett syndrome, a form of autism that affects girls, is caused by a mutation in the gene for an enzyme that recognizes the chromatin marks placed directly on the DNA. At least in mice, the neurons resume normal function when the mutation is corrected. In several forms of cancer, tumor-suppressor genes turn out to have been inactivated not by mutation, the usual known cause, but by the incorrect placement of marks that invite chromatin regulators to silence the genes.

Drugs developed by Peter A. Jones of the University of Southern California reverse the chromatin silencing of these antitumor genes. Two have recently been approved by the Food and Drug Administration for a blood malignancy, myelodysplastic syndrome.

Besides governing access to the genome, the epigenome also receives a host of signals from the environment. A family of enzymes called sirtuins monitors the nutritional state of the cell, and one of them removes a specific mark from the chromatin, providing a possible route for the genome to respond to famine conditions. Accumulating errors in the epigenome's regulation could allow the wrong genes to be expressed, a possible cause of aging.

A principal new technique for studying the marks on an epigenome is to break the chromosomes into fragments, which are then treated with antibodies that bind to a

specific mark. The DNA fragments so designated are decoded and matched to sites on the human genome sequence. This provides a genome-wide map of how a particular mark is distributed in a particular epigenomic state. The CHiP-seq maps, as they are called, have been very useful but are far from capturing the full detail of the epigenome, a dynamic structure that can change in minutes.

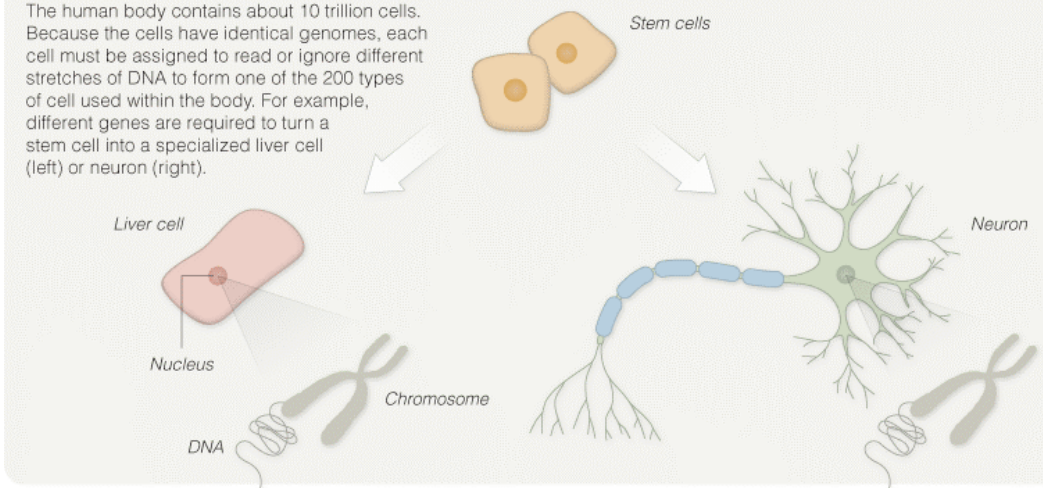
Individual researchers have made considerable progress but may not be able to assemble the comprehensive set of epigenomic marks and states that would be most useful to those developing new approaches to disease and aging. “I think the effort needs to be organized,” Dr. Young said. “It would benefit from being larger than it is.”

The Epigenome: Guiding Cells to Their Specialized Roles

Researchers are finding that a complex layer of proteins and markers called the epigenome controls access to genetic information, allowing each cell to read the genes necessary for cell-specific functions but blocking off most of the rest of the genome.

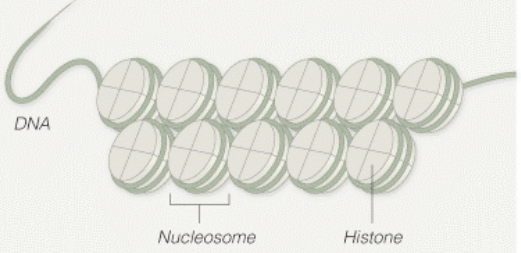
CELL DIFFERENTIATION

The human body contains about 10 trillion cells. Because the cells have identical genomes, each cell must be assigned to read or ignore different stretches of DNA to form one of the 200 types of cell used within the body. For example, different genes are required to turn a stem cell into a specialized liver cell (left) or neuron (right).



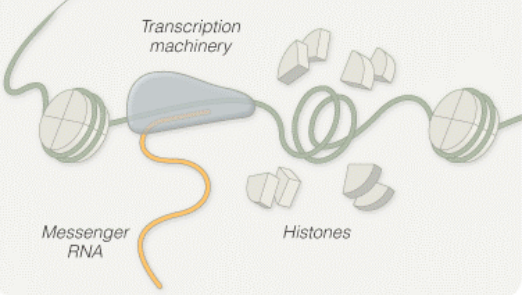
CLOSED DNA, UNREADABLE GENES

The long double strand of DNA within each chromosome is wrapped around millions of nucleosomes, small spools of eight proteins called histones. In some places the nucleosomes are tightly wound and closely packed, effectively blocking the cell from reading the genes within. Below, a gene-carrying stretch of DNA is permanently unreadable within this liver cell.



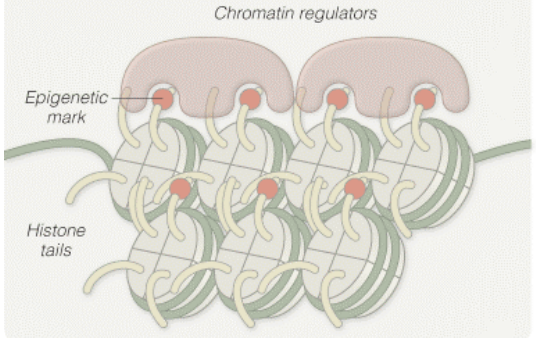
OPEN DNA, READABLE GENES

Other parts of the genome are more open and less tightly wound. Below, the same stretch of DNA within this neuron is accessible, allowing the cell's transcription machinery to read and copy the genes it contains. Other cellular machinery strips off the histones and reattaches them after the DNA has been transcribed.



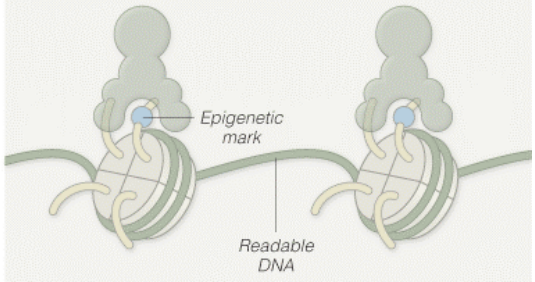
EPIGENETIC MARKERS

Access to the genes is controlled by the epigenome, a layer of millions of proteins and other marks that attach themselves to the genome. Tails extending from the histones are a common location for epigenetic marks. Below, proteins called chromatin regulators recognize a set of marks and help to lock up the nucleosomes.



RESEARCHING THE EPIGENOME

So far more than 100 kinds of epigenetic marks have been found, but researchers are still studying how the different components of the epigenome interact, replicate and adjust to the cell's environment. Below, permissive epigenetic marks attach to histone tails within this neuron, encouraging the cell's transcription machinery to latch on and begin reading and transcribing the DNA.



Sources: Cell; "Molecular Biology of the Cell," by Bruce Alberts et al.

JONATHAN CORUM/THE NEW YORK TIMES