

Behind the Scenes of Gene Expression

Researchers studying epigenetics are turning up the many ways that proteins and RNA can alter gene activity

Some of the weirdest genetic phenomena have very little to do with the genes themselves. True, as the units of DNA that define the proteins needed for life, genes have played biology's center stage for decades. But whereas the genes always seem to get star billing, work over the past few years suggests that they are little more than puppets. An assortment of proteins and, sometimes,

that doesn't involve changes to the DNA code and that can persist through one or more generations.

Over the past 5 years, researchers have shown that gene activity is influenced by the proteins that package the DNA into chromatin, the protein-DNA complex that helps the genome fit nicely into the nucleus; by enzymes that modify both those proteins and the DNA itself; and even by RNAs (see sidebar on p. 1066). These proteins and RNAs control patterns of gene expression that are passed on to successive generations. "The unit of inheritance, i.e., a gene, [now] extends beyond the sequence to epigenetic modifications of that sequence," explains Emma Whitelaw, a biochemist at the University of Sydney, Australia.

Indeed, the chromatin-modifying enzymes are now considered the "master puppeteers" of gene expression.

During embryonic development, they orchestrate the many changes through which a single fertilized egg cell turns into a complex organism. Such epigenetic phenomena may in fact underlie the problems encountered in mammalian cloning (see Rideout *et al.* Review, p. 1093). And throughout life, epigenetic changes enable cells to respond to environmental signals conveyed by hormones, growth factors, and other regulatory molecules without having to alter the DNA itself.

"[Epigenetic effects] give you a mechanism by which the environment can very stably change things," says Rudolph Jaenisch, a developmental biologist at the Whitehead Institute for Biomedical Research in Cambridge, Mas-

sachusetts. Researchers are hoping to harness these effects to design drugs that correct cancer and other diseases brought on by gene misregulation.

Shaky beginnings

Although cell biologists are just now probing epigenetic mechanisms, the late Conrad Waddington coined the term almost 50 years ago to help explain his ideas about development. At the time, many biologists thought that "you change the genome all the time as cells differentiate," explains Reik. Liver cells, for instance, became liver cells by losing unnecessary genes, such as those involved in making kidney or muscle cells.

In contrast, Waddington's epigenetic hypothesis proposed that the complement of genes remains constant, but they are switched on and off differently to make the various cells in the body. In other words, patterns of gene expression, not genes themselves, define each cell type. A series of experiments in the 1950s helped convince Waddington's colleagues that genes weren't lost during development.

But figuring out what might turn genes on and off was tough. One of the best clues came from the realization that the addition of methyl groups to DNA plays some role in silencing genes—and that somehow the methylation pattern carries over from one generation to the next. In the 1970s, for example, cancer biologists observed that the DNA in cancer cells tends to be more heavily methylated than DNA in healthy cells. They suspected that methylation might contribute to cancer development by altering gene expression, but proof was elusive.

"For 20 years, we correlated and correlated; it was boring because there were no mechanisms," recalls Jaenisch. No one could find the enzymes that added methyl groups



Too big. Apparently as a result of abnormal imprinting, the cloned lamb at left is bigger than the normal lamb at right. Cloned animals often have other health problems as well.

RNAs, pull the strings, telling the genes when and where to turn on or off.

The findings are helping researchers understand long-standing puzzles. Why, for example, are some genes from one parent "silenced" in the embryo, so that certain traits are determined only by the other parent's genes? Or how are some tumor suppressor genes inactivated—without any mutations—increasing the propensity for cancer?

For most of the past 40 years, researchers studying such phenomena had failed to explain why gene expression had gone awry. As a result of this lack of progress, "[the work] was looked upon as not quite serious science," says Wolf Reik, a developmental geneticist at the Babraham Institute in Cambridge, United Kingdom. "It had a little bit of the smell of something odd."

But in reality, those oddball phenomena were early clues suggesting that gene expression is not determined solely by the DNA code itself. Instead, as cell and molecular biologists now know, that activity also depends on a host of so-called epigenetic phenomena—defined as any gene-regulating activity



Epigenetically deprived. A mouse lacking an imprinted gene called *Mest* (right) fails to retrieve and care for her pups the way a normal female (left) does.

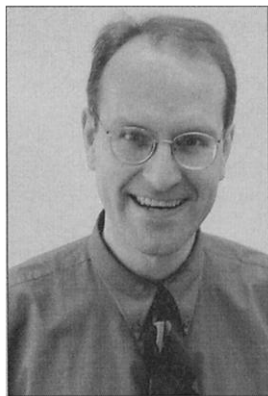
CREDITS: (LEFT TO RIGHT) SIMON WALKER, A. SURANI/WELLCOME CRIC INSTITUTE/U OF CAMBRIDGE

Champion of Chromatin: Alan Wolffe (1959–2001)

Even a good idea needs a spokesperson, and the idea that chromatin plays a dynamic role in regulating gene activity has had one of the best. Traditionally viewed as little more than DNA and its protein packaging, chromatin has recently gained new respect for its regulatory role (see main text). And although dozens of biologists have worked hard over the past decade to make that link, none have been as eloquent as Alan Wolffe in making the scientific community sit up and take notice. "He moved the field forward in a way nobody else was capable of," says Jeffrey Hayes, a biochemist at the University of Rochester Medical Center in New York.

Sadly, Wolffe died in a car accident in late May while attending a scientific meeting in Brazil. He was 41 and, at the time, senior vice president of the biotech firm Sangamo BioSciences Inc. in Richmond, California. Part of his legacy is the emerging field of epigenetics, in which researchers are learning about the many factors that influence gene expression.

Wolffe grew up in a small village in Staffordshire, England, where he worked in his father's store. He studied at Oxford as an undergraduate. Then as a graduate student with Jamshed Tata at the Medical Research Council in Mill Hill, U.K., he worked on hormonal influences on gene expression. Afterward, Wolffe was so highly recommended that Donald Brown, a developmental biologist at the Carnegie Institution



Silenced scientist. Before his untimely death, Alan Wolffe helped foster an appreciation of epigenetics.

in Baltimore, accepted the young biologist sight unseen as a postdoctoral fellow.

When Wolffe showed up at Brown's lab in 1984, he didn't waste any time: Within the hour he started his first experiments. "He was one of the most intense scientists I've ever met," Brown recalls. "We would talk over an experiment, and it would be done the next day." Yet at the same time, Wolffe liked to have fun, whether relaxing over a beer with his colleagues after a long day at talks or engaging a friend in a game of tennis.

At the Carnegie Institution, Wolffe began to look closely at what controls the expression of a small gene called 5S. During that work, he became interested in how chromatin affects the binding of transcription factors, proteins that control gene activity, to the DNA. That interest expanded after Wolffe joined the National Institutes of Health (NIH) in 1987, where he stayed for the next 13 years.

When Wolffe joined NIH, chromatin had slipped out of mainstream biology. But he changed that, first by organizing joint seminars that helped bring together the NIH chromatin researchers, and later by helping to build an active, international chromatin community and writing a book devoted to the topic. At one time, he jetted around the world giving some 31 talks in just 24 days. "His meteoric rise in the field was partly because he pushed things along with his enthusiasm," explains David Clark, a biochemist at the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). "He told people why they should care."

During that time Wolffe and his postdoctoral fellows—there would typically be two dozen in the lab at any one time—explained a wide range of chromatin-related phenomena. For example, they helped demonstrate that histones, the proteins in chromatin, are modified in different ways to modulate gene expression. Wolffe also helped elucidate the connection between chromatin remodeling and DNA methylation, a chemical modification often used to inactivate, or "imprint," genes. "His ability to bring it all together was amazing," says NIDDK's Gary Felsenfeld.

—E.P.

to DNA, nor could they even show that methylation is important. Eventually attention shifted to the easier-to-track-down genetic mutations that set off a tumor's uncontrolled growth.

In the 1990s several key experiments helped to bring epigenetics out of biology's backwater. In 1993, Jaenisch's team knocked out the gene for the enzyme that methylates DNA in mice and found that the resulting animals failed to develop properly, proving that methylation is important. Then at Johns Hopkins University School of Medicine in Baltimore, Maryland, Stephen Baylin and his colleagues found increased methylation of the tumor suppressor gene *p16* in a variety of human tumors. Treating these cells with a demethylating agent restored *p16* gene activity, suggesting that methylation could deactivate a perfectly good tumor suppressor gene. Others found extensive methylation of promoters—regulatory sequences that foster

gene activity—near tumor suppressor genes. This chemical change effectively shut down those genes, illustrating yet another route to cancer beyond genetic mutations.

Epigenetics got another boost in the early 1990s when biochemists began isolating the enzymes that add methyl groups to DNA on its cytosine bases. Among oth-

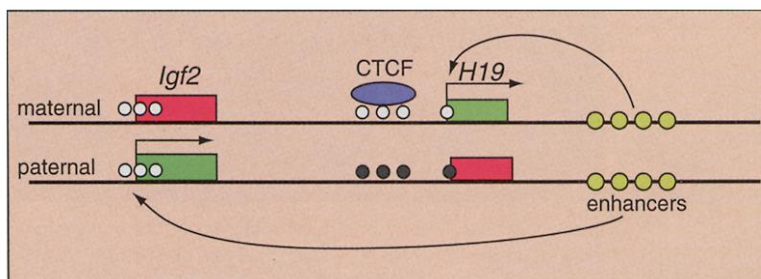
not just for cancer but for biology, Baylin recalls. Most recently, researchers have begun to learn what determines which part of the genome becomes methylated.

Making an imprint

This work, too, is the outgrowth of one of those odd genetic phenomena observed over the ages. According to historians, mule breeders first noticed 3000 years ago that a mare crossed with a donkey yields a mule, whereas a stallion crossed with a donkey produces a hinny, which has shorter ears, a thicker mane and tail, and stronger legs than the mule.

This made researchers aware that there could be parent-specific effects in the offspring. Other observa-

tions through the centuries suggested that the genes passed on by each parent had somehow been permanently marked—or imprinted, as it eventually came to be known—so that the expression patterns of the maternal



Setting up boundaries. When the CTCF protein binds to DNA, it blocks regulatory DNA downstream from interacting with the *Igf2* gene and only the *H19* gene is expressed. If methyl groups (black) prevent CTCF binding, *Igf2* is active, but *H19* is silenced.

er things, researchers found that methylation appears to be widespread, occurring in plants, animals, and fungi (see Martienssen and Colot Viewpoint, p. 1070). These reports "began to define a new paradigm"

RNA Rules—At Least Sometimes

A great deal of recent evidence has shown that proteins, especially enzymes that modify either DNA itself or the histone proteins that package it into chromatin, can alter gene activity, sometimes permanently (see main text). But proteins aren't alone in producing such epigenetic effects. A variety of RNAs can interfere with gene expression at multiple points along the road from DNA to protein.

A decade ago, plant biologists recognized a phenomenon called posttranscriptional gene silencing in which RNA causes structurally similar mRNAs to be degraded before their messages can be translated into proteins. That's apparently the side effect of a plant defense aimed at getting rid of pathogenic viruses, many of whose life cycles involve a double-stranded RNA. But in 1998, researchers found a similar phenomenon in nematodes, and it has since turned up in a wide range of other organisms, including mammals (*Science*, 25 May, p. 1469). This work is now providing researchers with easy ways to study genes without destroying them (see Matzke *et al.* Viewpoint, p. 1080).

RNAs can also act directly on chromatin, binding to specific regions

to shut down gene expression. Sometimes an RNA can even shut down an entire chromosome. One of the problems that sexual reproduction poses for the organism is the inheritance of two copies of the same sex chromosome. For example, female mammals have two X chromosomes, and if both were active, their cells would be making twice as much of the X-encoded proteins as males' cells do. Newly formed female embryos solve this "dosage compensation" problem with the aid of an RNA called XIST, translated from an X chromosome gene (see Park and Kuroda Viewpoint, p. 1083). By binding to one copy of the X chromosome, XIST somehow sets in motion a series of modifications of its chromatin that shuts the chromosome down—permanently. Male fruit flies also use an RNA to solve the dosage compensation problem; in their case it turns up the gene activity of the males' single X chromosome to match that of females' two.

And these may not be the only RNAs that influence gene function, because there are hints from other work of more to come. "The question now is how widespread are regulatory noncoding RNAs in the genome," says Denise Barlow of the Institute of Molecular Biology in Salzburg, Austria. She and others are eager to find out. —E.P.

and paternal genes differ in their progeny. These so-called imprints have since been found in angiosperms, mammals, and some protozoa. Not until 1991, however, did researchers begin isolating a variety of genes whose expression depended on their parents of origin. That year, researchers identified two genes, *Igf2r* and *H19*, that are active only when inherited from the mother; a third, called *Igf2*, is turned on only when inherited from the father.

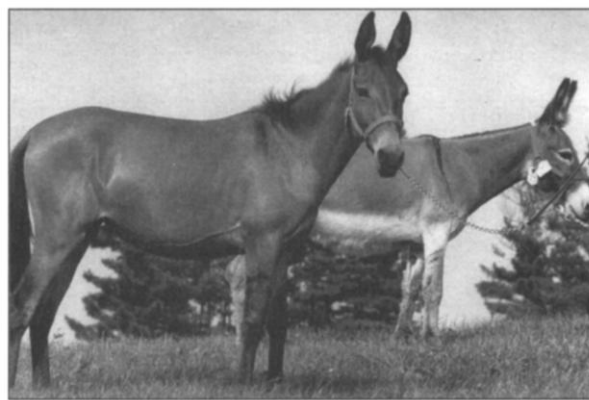
The discovery of these three genes has prompted an all-out search for other imprinted genes, as well as for the epigenetic tags, such as unusual methylation, that identify one copy as paternal and the other as maternal. Other researchers are investigating the molecular means by which silencing is initiated and enforced (see Ferguson-Smith and Surani Review, p. 1086).

So far, more than 40 imprinted genes have been found; about half are expressed when they come from the father and half when they come from the mother. Among these are a number of disease genes, including the *necdin* and *UBE3A* genes on chromosome 15 that are involved in Prader-Willi and Angelman syndromes, and possibly *p73*, a tumor suppressor gene involved in the brain cancer neuroblastoma. Seven, including *Peg3* and *Igf2*, affect embryonic growth or are expressed in the placenta.

In each case, either the maternal or paternal gene itself, or DNA located close by, is somehow earmarked for methylation. Sometimes methylation results in the gene's silencing, but other times, such as is the case for *Igf2r* and *Igf2*, methylation turns it on. Researchers don't know what accounts for that difference, and exactly how the earmarking occurs is also still a mystery. "The big questions are how is [the mark] set and what are the mechanisms to set it," says Vin-

cenzo Pirrotta, a molecular geneticist at the University of Geneva, Switzerland.

Researchers think that there are some clues lurking in the way that imprinted genes are arranged in the genome. "In the last 2 years, particularly with comparative sequencing approaches of the mouse and human, a lot of organizational features have been discovered," says Reik. In particular, researchers often find that imprinted genes are clustered. For example, the *H19* and *Igf2* genes and six other imprinted genes are located near one another on human chromosome 11 (11p15.5). Last fall, Randy Jirtle and his colleagues at Duke University Medical Center in Durham,



Parents matter. Hybrid of a horse and donkey, the hinny (foreground) differs from the mule because of parent-of-origin effects.

North Carolina, found that the imprinted genes *DDK1* and *GTL2* are neighbors on human chromosome 14q32, arranged much the same way that Princeton's Shirley Tilghman had found them in the mouse. The organization of the DNA around both these gene clusters is similar, suggesting that the surrounding DNA somehow specifies the imprinting arrangement.

On both chromosomes, genes next to one another are imprinted so as to be reciprocally expressed—that is, one is turned off when the other is turned on, depending on whether the chromosome comes from the mother or the father. And in both cases one gene in the pair on each chromosome codes not for a protein but for an RNA that never gets translated into a protein. Indeed, an estimated one-quarter of the imprinted genes produce these noncoding RNAs. Finally, the researchers have found that on both chromosomes, the pairs of genes within the clusters are separated by a stretch of DNA that includes so-called CpG islands, regions of DNA where the bases cytosine and guanine alternate with one another (see diagram, p. 1065).

That stretch of DNA contains a binding site for a protein called CTCF, which forms a chromosomal "boundary." When CTCF is attached, it isolates DNA upstream of the binding site from DNA downstream. Last year, several research groups showed a connection between methylation of some of the CpG islands, CTCF binding, and the activity of the *H19* and *Igf2* genes.

The *Igf2* gene is located before the *H19* gene on chromosome 11; farther along the chromosome, after both genes, are regulatory regions called enhancers. Transcription can occur only if the enhancers interact with promoters located near each gene. Last year Gary Felsenfeld and A. C. Bell of the National Institute of Diabetes and Digestive and Kidney Diseases found that CTCF binding blocks the enhancers' access to the *Igf2* promoter, thereby

CREDIT: PERMISSION OF CORNELL UNIVERSITY PHOTOGRAPH, CHARLES HARRINGTON, PHOTOGRAPHER

silencing that gene. However, the enhancers can still interact with the *H19* promoter, which coincides with the CpG island and CTCF binding site. Thus *H19* is active. But when the CpG island at the CTCF binding site is methylated, the enhancers cannot interact with the *H19* promoter and instead cause the *Igf2* genes to turn on.

Because the imprint region Jirtle studies on chromosome 14 is very similar to that on chromosome 11, he thinks it works the same way. He cautions, however, that “no one has gone in and knocked out these [chromosome 14] sites to see if they are functional.” He and others are looking at other imprinting clusters and their surrounding DNA to characterize more boundaries and potential epigenetic elements. In the imprinting field, says Marisa Bartolomei, a molecular geneticist at the University of Pennsylvania in Philadelphia, “the boundaries are by far the most exciting thing that’s happened in the last year.”

Histone code

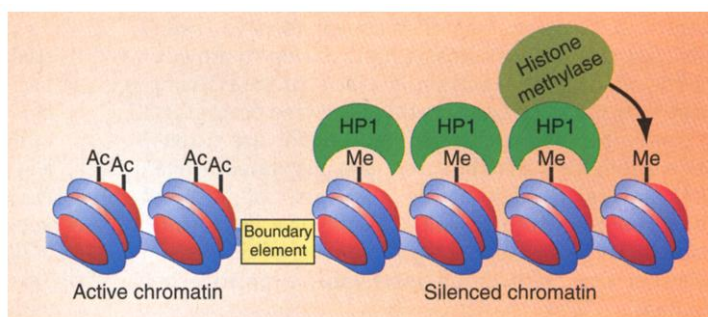
Boundaries may explain how one gene can be active while the next one down the line is not. But they don’t explain how the decision to raise or lower the boundary is made, that is, what determines what DNA gets methylated, and what maintains that decision through successive generations. Although methylation is key in some cases of imprinting, as the work on the *H19* and *Igf2* genes shows, it can’t be the whole story.

Increasing evidence suggests that the chromatin also plays a pivotal role. In the chromatin, DNA wraps around groups of eight histone proteins; each histone octet and its DNA make up a beadlike structure called a nucleosome. Until recently, “chromatin was thought of as just a way to package the DNA to keep it quiet,” Pirrotta says.

But in 1993 the late Alan Wolffe (see sidebar on p. 1065) and his colleagues showed that the addition of acetyl groups to chromatin’s histones alters other proteins’ access to the DNA, possibly influencing gene expression. Over the next 5 years, researchers isolated enzymes that either add acetyls to histone (acetylases) or remove them (deacetylases). They also found some intriguing interactions between these enzymes and other regulatory proteins. For example, these enzymes form complexes with transcription factors that turn genes on and

off. Because addition of acetyl groups to histones apparently opens up the chromatin, this acetylation may foster gene transcription by providing access for the transcription factors and other components of the gene-copying machinery.

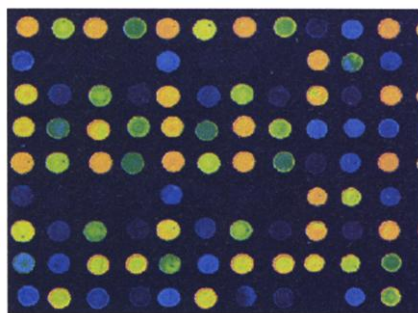
In 1998, Adrian Bird and his colleagues at the University of Edinburgh, U.K., showed that enzymes that pull acetyl groups off histones can work in conjunction with enzymes that add methyl groups to DNA. Methylation typically silences genes, but if the re-



Chromatin chemistry. Chemical modifications—acetylation (Ac) or methylation (Me)—of histone proteins determine whether genes on the surrounding DNA are active. HP1 is a transcription-inhibiting protein.

searchers inhibited the deacetylase, the genes remained active.

Chromatin histones are chemically modified by other mechanisms as well, and these modifications may also affect gene activity, either temporarily or in a more long-term, “epigenetic” way. Last year Thomas Jenuwein of the Research Institute of Molecular Pathology in Vienna, Austria, identified an enzyme that adds methyl groups to the same part of the histones that is acetylated. As he and C. David Allis of the University of Virginia, Charlottesville, are finding, histones with methyl groups are not acetylated and vice versa.



Methyl detector. DNA spotted on a glass slide binds differentially to matching DNA depending on its methylation state. Yellow indicates binding of highly methylated DNA while blue represents unmethylated samples.

binding of another protein called HP1 that prevents transcription.

All this work shows that chromatin’s proteins are much more than static scaffolding. Instead, they form an interface be-

tween DNA and the rest of the organism. Chemical modifications alter the chromatin structure, sometimes clearing the way for transcription and other times blocking it. The exact nature of these modifications remains mysterious. But Allis thinks methylation, acetylation, phosphorylation, and other changes likely occur in combinations that define a “histone code” that fine-tunes gene expression. “The different modifications mean different things, because they recruit different kinds of pro-

teins and prevent other kinds of modifications,” Allis explains. (See Jenuwein and Allis Review, p. 1074.)

Even as Allis and cell biologists work to decipher this code, enterprising biomedical entrepreneurs are trying to put the new insights about epigenetics into practice. Their goal is to understand how to diagnose diseases based on their epigenetics and then

to develop drugs that can alter the epigenetics and treat disease. A biotechnology start-up named Epigenomics, based in Germany and Seattle, Washington, is automating the determination of DNA methylation patterns. Aberrant patterns will likely be indicative of disease, and thus the approach could prove useful for diagnosing various cancers or other diseases. In collaboration with a Europe-based consortium, the company is also beginning to build the “Human Epigenome,” which will attempt to show the methylation profile of the entire human genome.

Companies, such as Sangamo BioSciences in Richmond, California, also hope to develop drugs for altering epigenetic marks. Cancer is one important target. Already, clinical trials are under way to treat leukemia with agents that remove methyl groups from genes; the goal is to restore the function of tumor suppressor genes. And at Memorial Sloan-Kettering Cancer Center in New York City, oncologist Paul Marks and his colleagues have just started two clinical trials of a drug that interferes with the removal of acetyl groups from histones. The drug has helped rid mice of a variety of tumors.

By changing the epigenetic status, “you can increase the expression of a gene [that’s needed], or you can increase the expression of a gene that will counter [a bad gene’s activity],” says Marks. “I think it’s going to be an extremely important area in the next few years.”

—ELIZABETH PENNISI