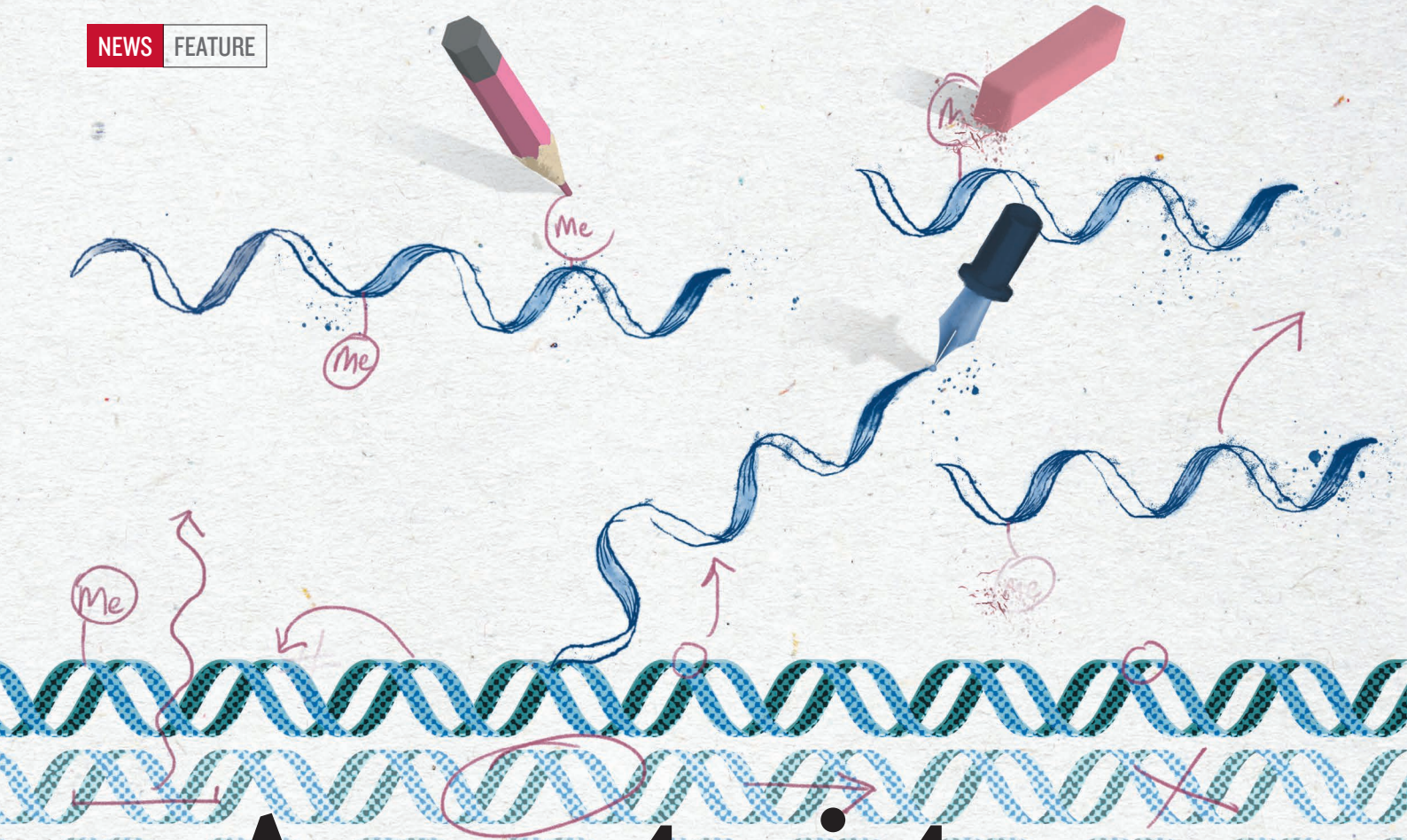


A new twist on epigenetics

How rediscovered chemical tags on DNA and RNA are shaking up the study of gene expression.

BY CASSANDRA WILLYARD

Some big ideas seem to appear out of nowhere, but in 2008 Chuan He deliberately went looking for one. The US National Institutes of Health had just launched grants to support high-risk, high-impact projects, and He, a chemist at the University of Chicago in Illinois, wanted to apply. But he needed a good pitch.

He had been studying a family of proteins that repair damaged DNA, and he began to suspect that these enzymes might also act on RNA. By a stroke of luck, he ran into molecular biologist Tao Pan, who had been investigating specific chemical marks, called methyl groups, that are present on RNAs. The pair worked in the same building at the University of Chicago, and began meeting regularly. From those conversations, their big idea took shape.

At the time, biologists were getting excited about the epigenome — the broad array of chemical marks that decorate DNA and its protein scaffold. These marks act like a chemical notation, telling the cell which genes to express and which to keep silent. As such, the epigenome helps to explain how cells with identical DNA can develop into the multitude of specialized types that make up different tissues. The marks help cells in the heart, for example, maintain their identity and not turn into neurons or fat cells. Misplaced epigenetic marks are often found in cancerous cells.

When He and Pan began working together, most epigenetic research focused on the tags associated with DNA and the histone proteins that it wraps around. But more than 100 different types of chemical mark had been identified on RNA, and nobody knew what they did. Some of the enzymes He was studying could strip off methyl groups, and He and Pan wondered whether one of them might work on RNA. If the marks could be reversed, they

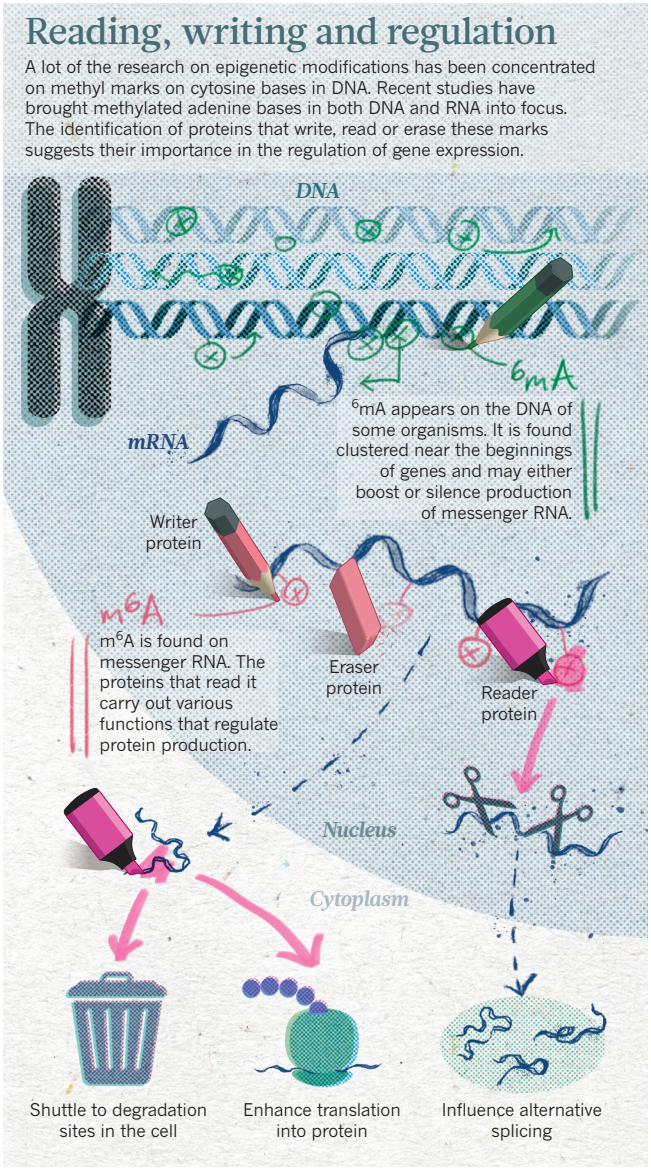
ILLUSTRATION BY NIK SPENCER/*NATURE*

might constitute an entirely new way of controlling gene expression. In 2009, they got funding to hunt for reversible marks on RNA and the proteins that erase them.

Nine years later, such research has given birth to an 'ome of its own, the epitranscriptome. He and others have shown that a methyl group attached to adenine, one of the four bases in RNA, has crucial roles in cell differentiation, and may contribute to cancer, obesity and more[1,2]. In 2015, He's lab and two other teams uncovered the same chemical mark on adenine bases in DNA (methyl marks had previously been found only on cytosine), suggesting that the epigenome may be even richer than previously imagined[3]. Research has taken off. "I think we're approaching a golden age of epigenomics and epitranscriptomics," says Christopher Mason, a geneticist at Weill Cornell Medical College in New York City. "We can actually start to see all these modifications that we knew have been there for decades."

MARKING THE MESSENGER

The governing rule of molecular biology — the central dogma — holds that information flows from DNA to messenger RNA to protein. Many scientists therefore viewed mRNA as little more than a courier, carrying the genetic information encoded in a cell's nucleus to the protein factories in the cytoplasm. That's one reason why few researchers paid much attention to the modifications made to mRNA.



They weren't a secret, though. The mark that pushed He to the forefront of epitranscriptomics was first discovered on mRNA in 1974 (ref. 4). Fritz Rottman, an organic chemist at Michigan State University in East Lansing, was trying to understand the role of RNA in regulating gene expression when he stumbled across a methyl group on adenine. The modified base is called N^6-methyladenosine, a mouthful that's commonly shortened to m^{6}A.

Rottman and his colleagues wrote that RNA methylation could be a way to select certain transcripts for translation into protein. "But that was all speculation," says Karen Friderici, an author of the 1974 paper and a geneticist at Michigan State University. The team didn't have a good way to investigate the mark's true function. "It was the beginning of molecular biology. We didn't have many of the tools that are available now," she says.

More than three decades later, He and Pan found the tools still lacking. "It's very difficult to actually study these modifications," Pan says. It requires powerful mass spectrometry and high-throughput sequencing techniques.

Two members of He's lab at the time, Ye Fu and Guifang Jia, pushed forward anyway, focusing on a protein called FTO, part of the family of methyl-stripping enzymes that He's group had been studying. Fu and Jia thought that it might remove methyl groups from RNA, but they struggled to identify its target. Fu and his colleagues began to synthesize snippets of RNA that contained different modifications, to determine whether FTO could remove them. It was slow going. Over the course of three years, the team faced a string of failures, "I almost thought I would never find the function," Fu says.

Finally, in 2010 the team decided to test FTO's activity on m^{6}A — the methylated adenine. The mark disappeared. The team had shown for the first time that RNA methylation was reversible[5], just like the marks found on DNA and histones. To He, it seemed like proof of an RNA-based system of gene regulation.

EVIDENCE MOUNTS

He's group wasn't the only one thinking about m^{6}A. In 2012, two teams of researchers independently published the first maps of where m^{6}A appears[6,7]. The studies revealed more than 12,000 methylated sites on mRNAs originating from about 7,000 genes. "After years in the dark, we were instantly facing a wide vista," wrote Dan Dominissini, an author of one of the studies, in an essay in *Science*[8].

The maps showed that the distribution of m^{6}A is not random. Its location suggested that the mark might have a role in alternative splicing of RNA transcripts, a mechanism that allows cells to produce multiple versions of a protein from a single gene.

Over the past few years, researchers have identified some of the machinery involved in regulating these marks. Each requires a writer to place it, an eraser to remove it and a reader to interpret it (see "Reading, writing and regulation'). As the identities of these proteins emerged, scientists have come to understand that m^{6}A affects not only RNA splicing, but also translation and RNA stability.

One m^{6}A reader, for example, makes mRNA degrade faster by shuttling it to decay sites in the cell. Another m^{6}A reader promotes protein production by shepherding methylated RNA to the ribosome.

Whether m^{6}A directs a cell to produce a protein or destroy a transcript depends on the location of the mark and on the reader that binds to it. But understanding how this selection works has been a major challenge, says Gideon Rechavi, a geneticist at Tel Aviv University in Israel who was involved in the mapping of m^{6}A.

What is clear is that m^{6}A has fundamental roles in cell differentiation. Cells that lack the mark get stuck in a stem- or progenitor-like state. That can be lethal: when He and his colleagues disabled the m^{6}A writer in mice, many embryos died *in utero*.

He has a possible explanation for the role of m^{6}A. Each time a cell changes from one state to another — such as during differentiation — the mRNAs in it must change too. This change in mRNA content, which He calls a transcriptome switch, requires precision and careful timing. He thinks that the methyl marks might be a way for cells to synchronize the activity of thousands of transcripts.

Self-described 'RNA geek' Wendy Gilbert, a

NIK SPENCER/*NATURE*

biologist at the Massachusetts Institute of Technology in Cambridge, says that He's explanation is plausible. "One of the things that I really like about Chuan's presentations over the last couple of years is his effort to try to speak to what is the most important aspect of the mark," she says. But she points out that there are other ways to coordinate the expression of large groups of genes, such as microRNAs, small bits of RNA that do not code for proteins and that help to silence genes. "I don't know that m^{6}A is the only way that you could do that," she says.

THE A'S HAVE IT

Although scientists have long known that RNA carries a host of modifications that decorate all four of its bases, mammalian DNA seemed to have only a few marks, all on cytosine. The most common modification in mammals, 5-methylcytosine or ^{5}mC, is so important that it's often referred to as the 'fifth base', after A, C, T and G. But He wondered whether there might be other marks hiding in the genome. Bacteria carry the DNA equivalent of m^{6}A — called N^6-methyladenine or ^{6}mA. "They use the methylation to distinguish between their own DNA or foreign DNA," says Eric Greer, a biochemist at the Boston Children's Hospital in Massachusetts. But researchers struggled to confirm its presence in more complex organisms.

In 2013, He's postdoc Fu had found an intriguing paper from the 1970s, which showed that algal DNA contains methylated adenine[9]. "Nobody ever knew the function, and nobody ever followed up," Fu says.

Fu and another postdoc, Guan-Zheng Luo, decided to take the investigation further and map the distribution of ^{6}mA in the DNA of the alga *Chlamydomonas*. They found it in more than 14,000 genes. And the distribution wasn't random: ^{6}mA clustered around the places where transcription begins. "We saw some periodic pattern of the peaks. It's like one peak after another," Fu says. It might be promoting gene activation, they reasoned.

Nearly 2,000 kilometres away in Boston, Greer and his colleagues had found ^{6}mA in the genome of a worm, *Caenorhabditis elegans*. Greer, a postdoc at the time, had been studying epigenetic inheritance using a *C. elegans* mutant that becomes less fertile with each successive generation. He wanted to understand how this infertility is transmitted from one generation to the next. *Caenorhabditis elegans* had long been thought to lack methyl marks, but Greer decided to double-check using antibodies that can bind specific methylated bases. He and his colleagues didn't find any ^{5}mC, but they did detect ^{6}mA. What's more, the levels seemed to be higher in the less-fertile generations, "raising the possibility that it could indeed be a carrier of this non-genetic information", he says. The result came as a surprise. Researchers had looked for ^{6}mA in multicellular organisms before, but they weren't able to find it because it is present at such low levels.

Greer's lab head, Yang Shi, knew that He had uncovered ^{6}mA in algae, and asked him for help. When He heard what Shi had found, he was excited. "We decided we're going to do this together," He says. A couple of months later, He met a researcher in China who had found ^{6}mA in the fruit fly *Drosophila*. "I almost fell to the floor," He says. In April 2015, the three papers came out simultaneously in *Cell*[10–12].

Andrew Xiao, who studies epigenetics at Yale University in New Haven, Connecticut, read the articles with interest. Xiao and his colleagues had identified ^{6}mA in mammalian cells, but they hadn't published their results. "Literally we thought nobody will take interest in this field," Xiao says. The *Cell* articles proved him wrong. "We realized we should hurry up."

> "It could indeed be a carrier of this non-genetic information."

Within the year, Xiao's group and another led by John Gurdon of the Gurdon Institute in Cambridge, UK, showed that ^{6}mA can be found at exceedingly low levels in a variety of vertebrate species including mouse and human[13,14]. When Xiao's group looked at the distribution of the mark in mouse embryonic stem cells, the team found the strongest peaks on the X chromosome. Here, the mark seemed to be involved in silencing gene expression. The researchers also identified an enzyme that seems to be a ^{6}mA eraser[14].

Xiao is still unravelling the function of ^{6}mA. He says that it seems to be crucial at certain developmental stages, acting like a molecular switch — barely present one moment, then there's a surge, and then it disappears.

"His paper was absolutely a bombshell," says Samie Jaffrey, a researcher at Weill Cornell Medical College. "It really showed functional roles for ^{6}mA." Both He and Shi say they have also found ^{6}mA in mammalian cells, but haven't yet published their results.

However, the significance of ^{6}mA isn't yet clear, Shi says. He points out that even with the latest technology, the modification is only borderline detectable and its precise location cannot be mapped. And the pattern of ^{6}mA will probably vary from tissue to tissue.

There are still big questions to untangle. Mamta Tahiliani, a geneticist at New York University School of Medicine in New York City calls the ^{6}mA work "incredibly exciting", but points out that researchers haven't yet shown that the mark passes from one generation of cells to its progeny, a hallmark of epigenetic modifications.

MINING FOR MORE MARKS

As some researchers dive deep to try to understand the function of m^{6}A and ^{6}mA, others are looking for new modifications. Last year, He, Rechavi and their colleagues reported[15] the discovery of another methyl mark on adenine in RNA called N^1-methyladenosine (m^{1}A). This mark also seems to promote translation, although the underlying mechanism is different from that of ^{6}mA. He says it might also have a role in synchronizing transcripts for the transcriptome switch.

Then, in January, Jaffrey and his colleagues reported on yet another kind of modification that occurs near the caps of mRNAs. The researchers found that mRNAs with this mark — called m^{6}Am — are more stable because their caps are harder to remove[16]. "It's exciting to people that the landscape of potentially regulated messenger RNA modifications that might influence gene expression could be an order of magnitude more complex than we thought before," Gilbert says.

Along with these new discoveries also come scientific squabbles. Jaffrey's work[16] suggests that FTO, which He identified as an m^{6}A eraser, actually targets m^{6}Am. And in October, He's group reported[17] that the enzyme Xiao flagged as a ^{6}mA eraser on DNA actually does a better job of stripping m^{1}A off a particular type of RNA. But such ambiguities are to be expected in a field that's experiencing a scientific gold rush.

"We are only in the beginning of the story," Rechavi says. And as the techniques improve, scientists will be able to see these marks more clearly. The wealth of research possibilities makes Mason feel "euphoric", he says. "It's like the most exciting time to be working in the field." ■ SEE TECHNOLOGY FEATURE P.503.

Cassandra Willyard *is a science journalist in Madison, Wisconsin.*

1. Ben-Haim, M. S., Moshitch-Moshkovitz, S. & Rechavi, G. *Cell Res.* **25,** 3–4 (2015).
2. Lin, S., Choe, J., Du, P., Triboulet, R. & Gregory, R. I. *Mol. Cell* **62,** 335–345 (2016).
3. Luo, G.-Z., Blanco, M. A., Greer, E. L., He, C. & Shi, Y. *Nature Rev. Mol. Cell Biol.* **16,** 705–710 (2015).
4. Desrosiers, R., Friderici, K. & Rottman, F. *Proc. Natl Acad. Sci. USA* **71,** 3971–3975 (1974).
5. Jia, G. *et al. Nature Chem. Biol.* **7,** 885–887 (2011).
6. Meyer, K. D. *et al. Cell* **149,** 1635–1646 (2012).
7. Dominissini, D. *et al. Nature* **485,** 201–206 (2012).
8. Dominissini, D. *Science* **346,** 1192 (2014).
9. Hattman, S., Kenny, C., Berger, L. & Pratt, K. *J. Bacteriol.* **135,** 1156–1157 (1978).
10. Greer, E. L. *et al. Cell* **161,** 868–878 (2015).
11. Fu, Y. *et al. Cell* **161,** 879–892 (2015).
12. Zhang, G. *et al. Cell* **161,** 893–906 (2015).
13. Koziol, M. J. *et al. Nature Struct. Mol. Biol.* **23,** 24–30 (2016).
14. Wu, T. P. *et al. Nature* **532,** 329–333 (2016).
15. Dominissini, D. *et al. Nature* **530,** 441–446 (2016).
16. Mauer, J. *et al. Nature* **541,** 371–375 (2017).
17. Liu, F. *et al. Cell* **167,** 816–828.e16 (2016).

CORRECTION

The News Feature 'A new twist on epigenetics' (*Nature* **542,** 406–408; 2017) omitted a reference (M. J. Koziol *et al. Nature Struct. Mol. Biol.* **23,** 24–30; 2016) that demonstrated ^{6}mA in mammalian cells.