



THE BLACK BOX OF REPROGRAMMING

BY DAVID CYRANOSKI

Scientists have been reprogramming adult cells into embryonic ones for decades — but they are only now getting to grips with the mechanics.

Eggs and sperm do it when they combine to make an embryo. John Gurdon did it in the 1960s, when he used intestinal cells from tadpoles to generate genetically identical frogs. Ian Wilmut did it too, when he used an adult mammalian cell to make Dolly the sheep in 1996. Reprogramming — reverting differentiated cells back to an embryonic state, with the extraordinary ability to create all the cells in the body — has been going on for a very long time.

Scientific interest in reprogramming rocketed after 2006, when scientists showed that

adult mouse cells could be reprogrammed by the introduction of just four genes, creating what they called induced pluripotent stem (iPS) cells¹. The method was simple enough for almost any lab to attempt, and now it accounts for more than a thousand papers per year. The hope is that pluripotent cells could be used to repair damaged or diseased tissue — something that moved closer to reality this year, when retinal cells derived from iPS cells were transplanted into a woman with eye disease, marking the first time that reprogrammed cells were

NIK SPENCER/NATURE

transplanted into humans (see *Nature* <http://doi.org/xhz>; 2014).

There is just one hitch. No one, not even the dozen or so groups of scientists who intensively study reprogramming, knows how it happens. They understand that differentiated cells go in, and pluripotent cells come out the other end, but what happens in between is one of biology's impenetrable black boxes. "We're throwing everything we've got at it," says molecular biologist Knut Woltjen of the Center for iPS Cell Research and Application at Kyoto University in Japan. "It's still a really confusing process. It's very complicated, what we're doing."

One of the problems, stem-cell biologists say, is that their starting population contains a mix of cells, each in a slightly different molecular state. And the process for making iPS cells is currently inefficient and variable: only a tiny fraction end up fully reprogrammed and even these may differ from one another in subtle but important ways. What is more, the path to reprogramming may vary depending on the conditions under which cells are being grown, and from one lab to the next. This makes it difficult to compare experimental results, and it raises safety concerns should a mix of poorly characterized cells be used in the clinic.

But new techniques are starting to clarify the picture. By carrying out meticulous analyses of single cells and amassing reams of detailed molecular data, biologists are identifying a number of essential events that take place en route to a reprogrammed state. This week, the biggest such project — an international collaboration audaciously called Project Grandiose — unveiled its results²⁻⁶. The scientists involved used a battery of tests to take fine-scale snapshots of every stage of reprogramming — and in the process, revealed an alternative state of pluripotency. "It was the first high-resolution analysis of change in cell state over time," says Andras Nagy, a stem-cell biologist at Mount Sinai Hospital in Toronto, Canada, who led the project. "I'm not shy about saying grandiose."

But there is more to do if scientists want to control the process well enough to generate therapeutic cells with ease. "Yes, we can make iPS cells and yes we can differentiate them, but I think we feel that we do not control them enough" says Jacob Hanna, a stem-cell biologist at the Weizmann Institute of Science in Rehovot, Israel. "Controlling cell behaviour at will is very cool. And the way to do it is to understand their molecular biology with great detail."

NUCLEAR TRANSFER

When Gurdon and Wilmut reprogrammed frog and sheep cells, respectively, they did it by transferring a differentiated nucleus into an egg stripped of its own DNA. Scientists knew that something in the egg was able to reprogram the nucleus, such that the genes associated with being a skin cell, for example, were switched off and those associated with pluripotency were switched on and triggered a cascade of

downstream events. In the following decade, researchers found various new ways to reprogram — adding nuclei to fertilized eggs and to embryonic stem cells — but these methods did little to clarify what it was in the cells that did the reprogramming and how the process worked.

That changed when Shinya Yamanaka and Kazutoshi Takahashi at Kyoto University made iPS cells¹. They showed that just four proteins that are usually expressed in early embryos or in embryonic stem cells could reprogram an adult cell — and, crucially, they also provided a tool that researchers could use to study reprogramming in a culture dish, something they have been doing ever since. Stem-cell biologists now

that only genes relevant for a fibroblast are expressed. It wouldn't do for a skin cell to suddenly behave like a dividing stem cell, because that can be the route to diseases such as cancer.

Scientists now have a good grip on what happens during the first 48 hours as the four Yamanaka factors, with brute force, kick cells out of this state. In embryonic stem cells, these proteins activate genes in a 'pluripotency network' that keeps cells proliferating indefinitely. But the factors act differently when shoved into a differentiated cell such as a fibroblast. When cell biologist Ken Zaret at the University of Pennsylvania in Philadelphia mapped the location of these factors during the first two days of

“THE ONE THING THAT WE KNOW IS THAT IT’S NOT MAGIC, THERE IS A MECHANISM.”

know that after introducing these proteins — sometimes known as the Yamanaka factors — there is a flurry of intense and mostly predictable gene expression. But then, after a few days, the cells enter a mysterious state in which they are dividing but stalled, failing to reprogram further. After a week or so, a slim few — only one in a thousand — become true pluripotent cells⁷.

This process is unpredictable, in the sense that it is impossible to know at the beginning which cells will reprogram, and it takes them a long time. But it is predictable in some ways. "Researchers doing it in Germany, Japan and the US will all get the iPS cells about the same time and at about the same rate," says Alexander Meissner at Harvard University in Cambridge, Massachusetts. "The one thing we know is that it's not magic, there is a mechanism. That's good news — we should be able to find it." And yet, Meissner says, it is "almost disappointing" how little progress there is from year to year.

From the cell's point of view, it is an immense task to overcome a fully differentiated state, which is like being in biological lock-down. Take fibroblasts, for example, the connective-tissue cells that scientists often extract from skin and try to reprogram. In the long process by which they gained their identity, these cells' DNA has been stamped with 'epigenetic'

markers, chemical modifications such as the addition of methyl groups or changes to the histone proteins that package up DNA. These ensure

reprogramming in human fibroblasts, he found that they were "physically blocked" from reaching their usual target genes by the conformation of the chromosomes⁸.

Instead, the proteins head for accessible areas of the chromosomes. Sometimes, they activate genes that force the cell to commit suicide; in others, they bind to distant control regions called enhancers that encourage the activation of genes known to be involved in the reprogramming process. Rudolf Jaenisch, a stem-cell scientist at the Massachusetts Institute of Technology in Cambridge, has labelled this widespread binding of the Yamanaka factors as "promiscuous".

Other studies have illuminated the sweeping changes that take place on chromosomes during this early phase. In a study published in 2011, Meissner's group showed that a type of histone modification that boosts gene expression, called H3K4me2, changes at more than 1,000 positions in the genome of these cells: it was added at many sites on pluripotency genes, and dropped from sites where genes specific for fibroblasts reside¹⁰. At the same time, the cells look and behave differently: they compact and move around less.

"Our early thought was that the factors create complete chaos," says Meissner. "But this first step is predictable and consistent across all cell types." Now he can almost foretell for a given cell type "which sites might become open to active transcription, which might be modified, and which will stay silent", he says. "That part you can predict. But that doesn't answer the question of what happens next?"

➔ **NATURE.COM**
Listen to a
podcast about
reprogramming at
go.nature.com/xydbge

The week-long lag that follows flummoxes scientists. The cells soldier on, and some express new genes, but not in a predictable or comprehensible way. Even the H3K4me2 modifications mapped by Meissner do not seem to boost gene expression until much later in the process. “Most cells reach a partially reprogrammed state. Some get beyond that, and we’re not sure why,” says Meissner. “That is the black box.” If a cell starts to pump out Sox-2 protein, however, that is a really good sign that it is progressing. “Once Sox-2 comes on, everything falls in line,” says Jaenisch, who studied the activity of nearly 50 genes in individual cells as they went through reprogramming¹¹. Within a few days, the production of this and other transcription factors necessary for pluripotency all ramp up.

But why does all this take so long, and why is it so rare? “We don’t understand why it can’t be faster,” says Woltjen. He suggests that a cell might need to go through several divisions, each taking at least half a day, to reshape its epigenetic state. “Perhaps that’s one limiting factor,” he says.

Yamanaka offers several possible explanations for the low conversion rate. One is that

idea that variability in the reprogramming process is producing fundamentally different cells. The project, launched in 2010 by some 30 senior scientists at 8 research institutes, was motivated by Nagy’s desire to open up the black box. “I wanted to find out what was in it,” he says. After triggering reprogramming with the Yamanaka factors, the team collected 100 million cells per day for a month, and then regularly analysed their production of protein and RNA, their changing methylation state and more. The methylation analyses alone produced so much data that collaborators resorted to sharing it on terabyte hard drives that they FedEx-ed around the world. The size of the undertaking also inspired the project’s title, Nagy says. “The name just came out of my head when I was considering how much data was being collected,” he says.

A CLASS OF ITS OWN

The headline finding is the new category of pluripotent cell, called F-class cells after the fuzzy appearance of the cell colonies. These cells were produced with a small tweak to the iPS-cell recipe: instead of stopping expression

true. According to this ‘stochastic’ model, as the reprogramming factors trigger cascades of molecules, some cells will drift into a reprogrammed state and some will not, and which way they go cannot be predicted.

But some studies, including one by Hanna¹², show that the reprogramming method can be tweaked to make the process more efficient — suggesting that the randomness can be controlled or even eliminated. These studies imply that reprogramming can be switched from a stochastic process to a deterministic one, in which one step inevitably follows the next to a new cell state.

Many scientists now say that reprogramming involves both deterministic phases — at the start and end — and a stochastic phase, which is the mysterious week in the middle. Hanna plays down the debate altogether, seeing little contradiction between the two sides. “I do not believe there is a stochastic versus deterministic camp.” He compares reprogramming to flipping a coin: each flip will have a random outcome, but after 100 flips, close to 50% of them will have come up heads. Similarly, whether a given cell flips into a reprogrammed state might be random. But over time, a reprogramming method should produce a certain percentage — maybe 10% — of pluripotent cells every time. Further experiments might resolve the debate, says Zaret, by pinpointing the events that snap the cells out of their week-long lethargy.

For Zaret, the reprogramming debate offers a window on a bigger concept: how order in biology arises from randomness. “Cellular systems are built upon intrinsic noise and stochastic events that somehow elicit cell fates that are locked down and do not switch back and forth,” he says. This question is at the basis of cell type control, he says, and draws him to the research.

For others, like Yamanaka, the incentive to open the black box is a practical one. More-efficient reprogramming makes for better experiments and a more reliable source of cells that can eventually be used in human medicine. “The motivation of my research is to treat patients,” he says. “Anything that helps push iPS cells into the clinic excites me.” ■

David Cyranoski reports for Nature from Shanghai.

1. Takahashi, K. & Yamanaka, S. *Cell* **126**, 663–676 (2006).
2. Hussein, S. M. *et al.* *Nature* **516**, 198–206 (2014).
3. Tonge, P. D. *et al.* *Nature* **516**, 192–197 (2014).
4. Clancy, J. L. *et al.* *Nature Commun.* <http://dx.doi.org/10.1038/ncomms6522> (2014).
5. Benevento, M. *et al.* *Nature Commun.* <http://dx.doi.org/10.1038/ncomms6613> (2014).
6. Lee, D.-S. *et al.* *Nature Commun.* <http://dx.doi.org/10.1038/ncomms6619> (2014).
7. Okita, K., Ichisaka, T. & Yamanaka, S. *Nature* **448**, 313–317 (2007).
8. Soufi, A., Donahue, G. & Zaret, K. S. *Cell* **151**, 994–1004 (2012).
9. Buganim, Y., Faddah, D. A. & Jaenisch, R. *Nature Rev. Genet.* **14**, 427–439 (2013).
10. Koche, R. P. *et al.* *Cell Stem Cell* **8**, 96–105 (2011).
11. Buganim, Y. *et al.* *Cell* **150**, 1209–1222 (2012).
12. Di Stefano, B. *et al.* *Nature* **506**, 235–239 (2014).

“I’M NOT SHY ABOUT SAYING GRANDIOSE.”

the starting cell population is a rainbow of cell types. The chunk of tissue used to derive fibroblasts, for example, probably contained a mix of subtly different cell types; even those that are fibroblasts will differ slightly in the blend of proteins and other molecules they contain. Furthermore, cells growing in culture are constantly shuttling back and forth between different states. This means that the introduced reprogramming factors will affect each cell differently. “What works for one subset of the population will not work for others,” Yamanaka says. Minor differences in cell culture and the relationship with neighbouring cells also make it difficult to control all the variables and command the cells like an obedient army, he adds. “A perfect implementation is impossible.”

Researchers are now trying to classify some of the cell types that come out of the black box, and are tinkering with reprogramming techniques to see if they can pin down how and where they diverge. Woltjen, for example, has shown that the ratio of the different reprogramming factors affects the type of cells produced. One set of conditions has a high success rate, but the resulting cells end up in a partially reprogrammed, unstable state; another has a low efficiency but produces mainly high-quality iPS cells.

Project Grandiose has also supported the

of the reprogramming factors after a few days, and the researchers continued to supply them. “That leads to a bifurcation,” says Nagy.

F-class cells are different from iPS cells because they fail one of the most stringent tests of pluripotency: when injected into mouse embryos they cannot contribute to tissues in the resulting chimaeric mice. For this reason, some critics say that F-class cells could be what other scientists have been calling ‘partially reprogrammed’ cells. But Nagy says that cells do not have to contribute to chimaeras to be considered pluripotent, and points to the cells’ other characteristics of pluripotency: for example, they form what is known as a teratoma, which contains a range of differentiated cell types.

Nagy says that others have overlooked the F-class state because they were only looking for cells that were similar to embryonic stem cells, whereas his team was “unbiased by expectation of what pluripotency should look like”. He thinks that there are more states of pluripotency to be found, and his group will be looking for them in its hard drives. “It’s a conceptually important thing, it opens up a big door,” he says.

All these studies are adding fuel to a central debate in the reprogramming community: does the process have an inherently random and unpredictable element to it? Until recently, there was a general consensus that this was