

THE TRICKIEST FAMILY TREE IN BIOLOGY

Scientists are striving for a deeper view of development, from embryo to adult, cell-by-cell.

BY EWEN CALLAWAY

For 18 months in the early 1980s, John Sulston spent his days watching worms grow. Working in twin 4-hour shifts each day, Sulston would train a light microscope on a single *Caenorhabditis elegans* embryo and sketch what he saw at 5-minute intervals, as a fertilized egg morphed into two cells, then four, eight and so on. He worked alone and in silence in a tiny room at the Medical Research Council Laboratory of Molecular Biology in Cambridge, UK, solving a Rubik's cube between turns at the microscope. "I did find myself little distractions," the retired Nobel prize-winning biologist once recalled.

His hundreds of drawings revealed the rigid choreography of early worm development, encompassing the births of precisely 671 cells, and the deaths of 111 (or 113, depending on the worm's sex). Every cell could be traced to its immediate forebear and then to the one before that in a series of invariant steps. From these maps and others, Sulston and his collaborators were able to draw up the first, and so far the only, complete 'cell-lineage tree' of a multicellular organism¹.

Although the desire to record an organism's development in such exquisite detail preceded Sulston by at least a century, the ability to do so in more-complex animals has been limited. No one could ever track the fates of billions of cells in a mouse or a human with just a microscope and a Rubik's cube to pass the time. But there are other ways. Revolutions in biologists' ability to edit genomes and sequence them at the level of a single cell have sparked a renaissance in cell-lineage tracing.

The effort is attracting not just developmental biologists, but also geneticists and technology developers, who are convinced that understanding a cell's history — where it came from and even what has happened to it — is one of biology's next great frontiers. The results so far serve up some tantalizing clues to how humans are put together. Individual cells from an organ such as the brain could be related more closely to cells in other organs than to their surrounding tissue, for example. And unlike the undeviating developmental dance of *C. elegans*, more-complex organisms invoke quite a bit of improvisation and chance, which will undoubtedly complicate efforts to unpick the choreography.

But even incomplete cellular ancestries could be informative. Sulston's maps paved the way for discoveries surrounding programmed cell

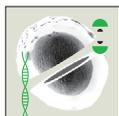
death and small, regulatory RNA molecules. New maps could elucidate the role of stem cells in tissue regeneration or help combat cancer — a disease of unharnessed lineage expansion. "There's a real feeling of a new era," says Alexander Schier, a developmental biologist at Harvard University in Cambridge, Massachusetts, who is using genome editing to trace the cell-lineage history of zebrafish and other animals.

RECONSTRUCTING HISTORY

A cell's history is written in its genome: every mutation acquired that gets passed on to daughter cells serves as a record. In 2005, the computer scientist Ehud Shapiro at the Weizmann Institute of Science in Rehovot, Israel, calculated that researchers could use the natural mutations in individual human cells to piece together how they are related². He conceived of a corollary (in concept at least) to the *C. elegans* cell map, which he called the Human Cell Lineage Project. But the field, he says, wasn't ready. "When we offered this vision, neither the field nor the name of single-cell genomics existed."

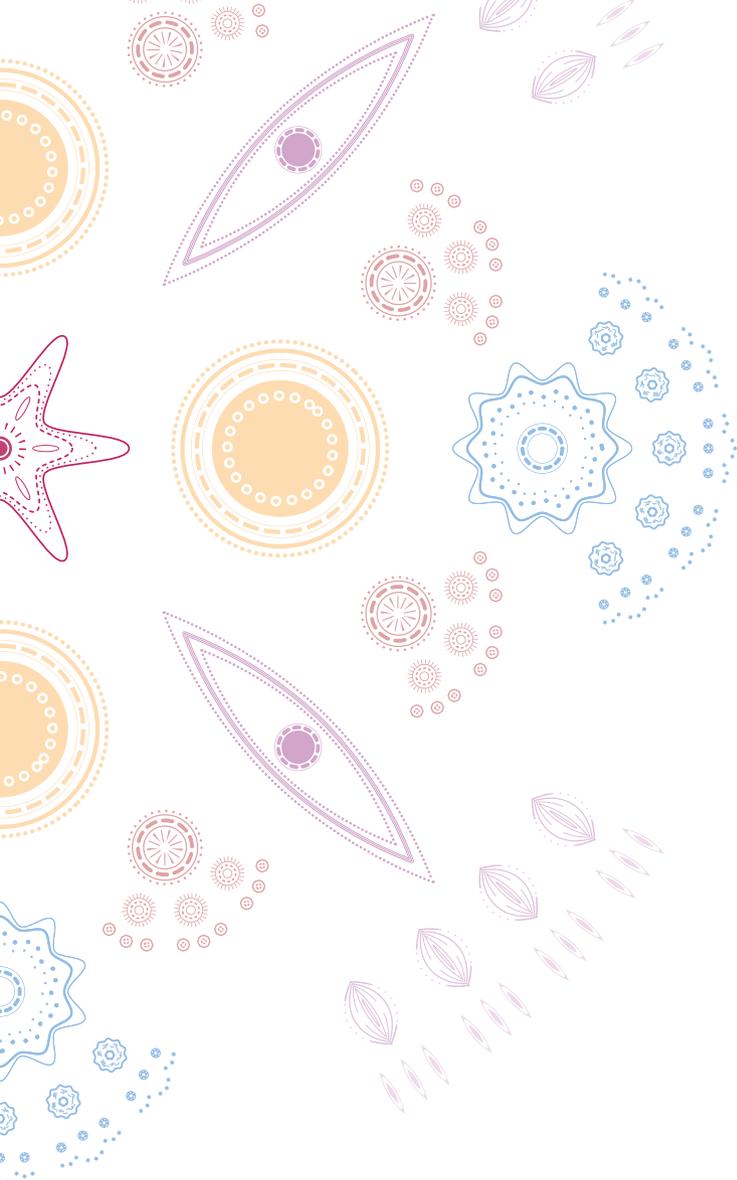
Fast forward a decade, and researchers have developed a suite of powerful tools to probe the biology of lone cells, from their RNA molecules and proteins to their individual and unique genomes. Now, he envisions a way of capturing the developmental course of a human, frame by frame, from fertilized egg to adult. "You want the whole movie with 3D frames from beginning to end," he says. To make such a film, it's not even necessary to look at the entire genome. Shapiro's team is focusing on repetitive stretches of DNA peppered across the genome called microsatellites. These sequences tend to mutate more frequently than other bits of the genome, and his team is

ILLUSTRATION BY JASIEK KRZYSZTOFIAK/NATURE



SINGLE-CELL BIOLOGY

A Nature special issue
go.nature.com/singlecell



working on sequencing tens of thousands of them across the genomes of hundreds of individual human cells to determine how they relate.

Christopher Walsh, a neuroscientist and developmental biologist at Boston Children's Hospital and Harvard Medical School, doubts that researchers will ever reconstruct a complete human cell-lineage map to match that of *C. elegans*, but even a less than complete tree will pay dividends, he says. "I've been studying cell lineage in the cortex for 25 years, and the idea of studying it directly in the human brain was an inconceivable dream. Now it's a reality."

In experiments described in 2015, Walsh's team sequenced the complete genomes of 36 cortical neurons from 3 healthy people who had died and donated their brains to research³. Reconstructing the relationship between the brain cells in an individual revealed that closely related cells can be spread across the cortex, whereas local areas can contain multiple distinct lineages. Successive generations of cells seem to venture far from their ancestral homes. One cortical neuron, for instance, was more closely related to a heart cell from the same person than to three-quarters of the surrounding neurons. "We were not expecting to find that," Walsh says.

Walsh's team is trying to understand how mosaicism in the brain — in which some cells harbour different gene variants — affects health. They have identified, for example, forms of epilepsy that occur even when just a small percentage of cells in a tiny brain region carry a disease-causing mutation. And they have found that individual neurons from healthy individuals can bear mutations that would cause seizures and schizophrenia if present more widely. It seems from this work that it matters which cells end up with a mutation. "The lineage basically determines what diseases are possible," Walsh says.

Other scientists are uncovering records of life's earliest events in the genomes of adult cells. In experiments published this year⁴, Michael Stratton, a geneticist at the Wellcome Trust Sanger Institute in Hinxton, UK, and his team sequenced white blood cells from 241 women with breast cancer and looked for mutations found in only a subset of their blood cells. The study revealed mutations that occurred very early in development, perhaps as far back as the two-cell embryo. And they noted that the descendants of these cells do not contribute equally to the blood system of adults. This could be because one cell multiplies more efficiently than the other; or it could, as Stratton suspects, be that by chance one ends up contributing more to a developing fetus than to a placenta or other supporting tissues.

Future studies, Stratton says, will look for bottlenecks in development that limit the contribution of some cell lineages. "We're beginning to see the rules of development in normal human beings," he says.

FROM BLOBS TO BARCODES

Jay Shendure, a geneticist at the University of Washington in Seattle, still remembers the day he became fascinated with cellular histories. As a 14-year-old with an interest in biology and computers, he wrote a program that modelled a mass of multiplying cells to impress his uncle, a reconstructive surgeon visiting from India. "He said, 'This is amazing. One day you'll do the same thing, and instead of a blob it will be a whole baby,'" Shendure recalls.

Nearly a decade later, Shendure was a first-year graduate student working for the Harvard geneticist George Church. Church presented a list of ideas ("all of which, at the time, seemed totally absurd," Shendure says); one of them was to reconstruct the lineages of many cells at once, in a single experiment. Shendure toiled for six months trying to use DNA-flipping enzymes called recombinases to create a readable record in the genomes of bacteria as they divide. Rather than relying on naturally acquired mutations in the genome, the system would essentially create variants to keep track of.

Shendure eventually switched projects, but he revived the idea a few years ago when graduate students Aaron McKenna and Greg Findlay joined his laboratory in Seattle. They realized that the popular genome-editing tool CRISPR-Cas9 would be ideal for introducing traceable mutations to whatever part of the genome they wanted (see "The lines of succession"). Teaming up with Schier's lab, they unleashed CRISPR-Cas9 in two single-cell zebrafish embryos and instructed it to edit DNA 'barcode' sequences that had been engineered into their genomes. They then sequenced these barcodes in cells of an adult animal and used the mutations in them to piece together their lineage⁵.

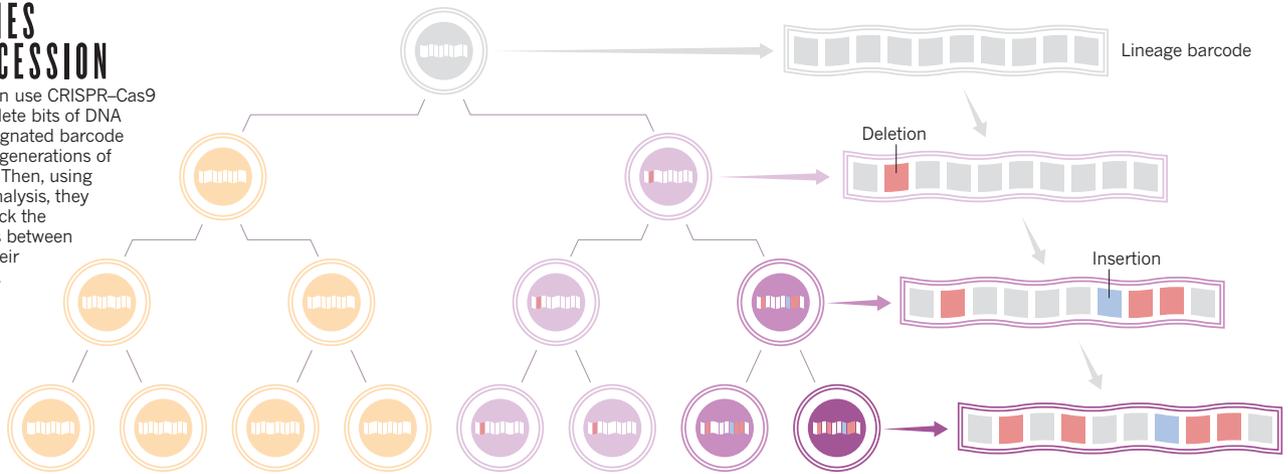
The trees they produced show that a small number of early-forming embryonic lineages give rise to the majority of cells in a given organ. More than 98% of one fish's blood cells, for instance, came from just 5 of the more than 1,000 cell lineages that the team traced. And although these five contributed to other tissues, they did so in much lower proportions. They were almost entirely absent from the muscle cells in the heart, for example, which was mostly built from its own small number of precursors. "It was profoundly surprising to me," says Shendure. His colleague Schier says he is still trying to make sense of the data.

Jan Philipp Junker, a quantitative developmental biologist at the Max Delbrück Center for Molecular Medicine in Berlin, says that the cell-lineage trees of early embryos probably vary greatly between individuals, and that the dominance of particular lineages observed by Shendure and Schier's team could be the result of chance events. The cells of an early embryo move around, and only a fraction of them contribute to the final organism, for example. It would be more revealing, he adds, to track later developmental events, such as the formation of the three germ layers that give rise to different organs, because these events are less governed by luck.

Junker and others have developed a bevy of other CRISPR-based techniques for piecing together developmental histories. He and Alexander van Oudenaarden, a systems biologist at Utrecht University in the Netherlands, applied such an approach to track the regeneration

THE LINES OF SUCCESSION

Scientists can use CRISPR–Cas9 to add or delete bits of DNA within a designated barcode over several generations of cell division. Then, using single-cell analysis, they can trace back the relationships between cells from their shared edits.



of a damaged fin in zebrafish. Regeneration, they discovered, occurred in the same kind of way as development: few of the cell lineages that gave rise to the original fin were lost when it was remade from stem cells. The finding confirmed previous studies, but the CRISPR-based methods allowed the team to trace lineages of thousands of cells in a single experiment⁶.

Church says his team has used CRISPR to study mouse development and has managed to record the embryonic cell divisions that give rise to the three major germ layers, which form all the body's organs⁷. "I don't think we're that far away from doing a complete lineage," he says.

Some researchers strive to know not just how an organism's cells relate to one another, but what happened to them along the way. Michael Elowitz and Long Cai, both at the California Institute of Technology in Pasadena, have developed a lineage tracer that creates fluorescent probes to help them observe the histories of cells as they develop⁸. Their method can track whether certain developmental genes have been turned on in the past for a given lineage. On 5 July, Elowitz, along with Shendure and Schier, were awarded a 4-year, US\$10 million grant from the Paul G. Allen Frontiers Group to combine their technologies. The trio plan to develop synthetic chromosomes that act as tape recorders for cell-lineage history and molecular events.

Such recordings might allow scientists to tinker with a cell's development in more delicate ways than current cell-reprogramming techniques allow, says Tim Lu, a synthetic biologist at the Massachusetts Institute of Technology in Cambridge who is also working on a technology to record a cell's history⁹. "You might see some version of these recorders being inserted into the cell therapies of the future," although it won't be for a while, he cautions. "I'm not going to go and inject my CRISPR recorder into a patient."

LINEAGES FOR LIFE

Cancer is where new lineage-tracing methods are likely to make waves first. "Cancer is a disease of lineage — it's a disease of stem cells," says Walsh. One question that researchers are starting to tackle is the origin of metastatic cells, which emerge from the primary tumour and invade sometimes distant organs. They tend to be the hardest tumour cells to vanquish and the ones most likely to kill patients.

A team led by cancer geneticist Nick Navin at the University of Texas MD Anderson Cancer Center in Houston published lineage maps of two colon cancers in May¹⁰. The results showed that liver-invading metastatic cells shared many DNA mutations with the primary tumours they came from, suggesting that the metastasis had emerged at a late stage and hadn't needed a bunch of new mutations to spread. Lineage mapping could also show whether tumours really develop from single cells, as geneticists have argued, or whether they originate from multiple cells, as some imaging studies have suggested. Navin suspects that similar work could be used to direct treatment. His team and others are tracing cancer-cell lineages in patients as they begin taking drugs. They

hope these studies can spot resistant lineages, allowing doctors to pick better treatments and switch medicines in time to make a difference.

At the moment, however, promise in the field far exceeds the reality. And Sulston's lineage maps of *C. elegans* still loom large over current efforts. Stephen Quake, a bioengineer at Stanford University in California, devised his own method for tracking cellular ancestry through CRISPR and decided to test it in the worm¹¹. "It's nice to have a gold standard," Quake says. He and his team sequenced the cells of a mature animal after CRISPR had mutated its genome during development. The efforts took much less time than the year and a half that Sulston spent with his microscope. But Quake says that the picture they developed was also less than complete. Yes, it captured a key transition in roundworm development — the segregation of cells bound for the intestine and those that give rise to the rest of the body — but it lacked the exquisite detail Sulston observed with his eyes. "I'll be perfectly blunt. I'm not very impressed with my results," says Quake, who hadn't even planned to publish the work until he saw the rush of other papers using similar techniques. "No one has really got it licked yet," he says.

There is an argument to be made that Sulston set the bar too high with *C. elegans*. "This whole concept of a lineage tree is very much influenced by this classic work," says Junker. And that may deserve a rethink.

In fish, mice and humans, no two individuals' cell lineage trees are likely to look exactly the same, and each probably changes throughout the individual's lifetime, as tissues repair and regenerate themselves. Junker and others hope that the new techniques will allow biologists to ask questions about the variability in lineage trees — between individuals, between their organs and as they age. As Schier puts it: "We don't know how many ways there are to make a heart."

It is that vast unknown that could make such work transformative, says Elowitz: "It would change the kinds of questions you could ask." Sulston's map led biologists into uncharted territory, says Schier, and this could do the same. "We can't tell you what exactly we're going to find, but there is a sense that we're going to find some new continents out there." ■

Ewen Callaway writes for Nature from London.

1. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. *Dev. Biol.* **100**, 64–119 (1983).
2. Frumkin, D., Wasserstrom, A., Kaplan, S., Feige, U. & Shapiro, E. *PLOS Comput. Biol.* **1**, e50 (2005).
3. Lodato, M. A. *et al. Science* **350**, 94–98 (2015).
4. Ju, Y. S. *et al. Nature* **543**, 714–718 (2017).
5. McKenna, A. *et al. Science* **353**, aaf7907 (2016).
6. Junker, J. P. *et al.* Preprint at <http://biorxiv.org/content/early/2017/01/04/056499> (2017).
7. Kalhor, R., Mali, P. & Church, G. M. *Nature Methods* **14**, 195–200 (2017).
8. Frieda, K. L. *et al. Nature* **541**, 107–111 (2017).
9. Perli, S. D., Cui, C. H. & Lu, T. K. *Science* **353**, aag0511 (2016).
10. Leung, M. L. *et al. Genome Res.* <http://dx.doi.org/10.1101/gr.209973.116> (2017).
11. Schmidt, S. T., Zimmerman, S. M., Wang, J., Kim, S. K. & Quake, S. R. *ACS Synth. Biol.* **6**, 936–942 (2017).