

ESSAY

GENOMICS AND PROTEOMICS

Retracing embryological fate

Cloned organoids derived from adult mice enable reconstruction of embryonic lineages

By Sam Behjati^{1,2}

Mammalian development is a beautifully orchestrated process of cell division and differentiation during which the various cell lineages arise that form an organism. The precise nature, origin, and fate of these lineages remain a mystery in humans and in other mammals. In addition to illuminating fundamental developmental biology, mapping human cell lineages may offer insights into a range of physiological and pathological processes, such as stem cell development, congenital diseases, and childhood cancer (1).

Tracing a cell back to its origins is a challenge. In model systems, it is possible to label early embryonic cells prospectively and, subsequently, to trace the fate of each cell during development. However, in humans, this is not possible and necessitates an alternative approach to decoding cellular development retrospectively. This can be achieved by investigating the cellular DNA mutations acquired during embryonic development.

Infidelities during DNA replication in dividing cells can manifest as mutations. As an organism develops and cell lineages arise, each cell acquires a specific combination of mutations. Thus, cells within the same lineage share a set of mutations that are absent from other cells. These mutations, present in some, but not all, cells of an organism, are referred to as mosaic mutations. They encode the embryological history of an organism [see the figure (Top)].

When I joined Mike Stratton's group at the Wellcome Trust Sanger Institute as a doctoral student in 2011, others had attempted to reconstruct embryological lineages utilizing mosaic mutations. These studies tracked the fate of short insertion and deletion mutations (indels), by genotyping polymerase

chain reaction (PCR)-amplified DNA (2, 3). However, indels frequently present at repetitive regions of the genome and are consequently unstable and polymorphic, and these hamper the precise reconstruction of embryological lineages. We therefore adopted a different approach, focusing on single base substitutions, which are highly stable mutations, and harnessing the power of next-generation DNA sequencing.

Methods to amplify DNA from single cells replicate the genome incompletely and introduce artifactual DNA variants, which preclude an accurate portrayal of a genome.

By contrast, the physiological replication machinery utilized in cloning has very high fidelity. Thus, instead of sequencing individual cells, we studied clonal lines derived from single cells, which provided sufficient DNA without the need for biochemical amplification.

A major challenge we faced was the need to clone single cells reliably across different tissues from the same organism. We therefore sought help from Hans Clevers' group at the Hubrecht Institute, who pioneered the generation of three-dimensional clonal lines known as organoids. This cloning method enables efficient generation of cells from clonal structures, such as colonic crypts. In a major effort, Clevers' team generated 25 clonal organoids derived from multiple tissues of two mice for us. With their help, we were able to study organoids derived from the stomach, small bowel, colon, and prostate for evidence of mosaic mutations acquired during embryogenesis.

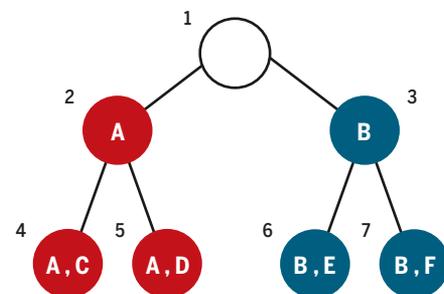
We sequenced the whole genomes of each organoid, as well as tail tissue of both mice, in hopes of finding mosaic variants shared between different tissues. From our experience sequencing cancer genomes, we were aware that artifactual mutations were a major potential source of false positives. Thus, after systematically harvesting a catalog of putative mosaic mutations, we validated each candidate mutation by capillary sequencing to confirm their presence in mutated samples and their absence from

unaffected samples. Using this approach, we were able to precisely define 35 mosaic mutations, 20 in mouse 1 and 15 in mouse 2.

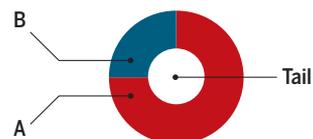
Using phylogenetic reconstruction methods, we then derived the embryological lineage of each cell [see the figure (Top)], bearing in mind a number of potential pitfalls, as discussed in our report (4).

Reconstructing the lineages of the 25 organoids revealed two key findings. First, the mutation rate of the developing mouse embryo was ~1.5 mutations per cell division. With this rate, we were able to define 17 of the 23 theoretically possible cell divisions. Thus, the intrinsic mutation rate of mice provided sufficient resolution for capturing most cell divisions. From what is known about mutational process in normal human

Simplified schema of the experimental approach



Lineage tree. Each circle represents a cell. Each letter represents a specific mutation, color coded by lineage. Numbers refer to specific cells. In this hypothetical experiment, cells 4, 5, 6, and 7 would be sequenced. Mosaic mutations are called from these sequences. Embryonic precursor cells (here, cells 2 and 3) can be inferred from these mutations. For example, as cells 4 and 5 both share mutation "A," it follows that they are derived from the same lineage and that a precursor that harbored only mutation "A" must have existed.



Contribution of cells to adult mouse. The circle represents the tail. Differentially colored segments represent the mutant read frequency of mutations "A" (red) and "B" (blue). To determine how much each mutation and/or precursor cell contributed to the adult mouse, we measured the mutant read frequency in tail tissue, as a proxy of the mouse. In this example, mutations "A" and "B," which define the daughter cells of the fertilized egg (cell 1, above), were measured. Together, the cells account for 100% of the tail, which makes it likely that they represent the first cell division. Note that the contribution of the precursor cells to the tail is unequal.



¹Cancer Genome Project, Wellcome Trust Sanger Institute, Hinxton CB10 1SA, UK. ²Department of Pediatrics, University of Cambridge, Cambridge CB2 1TN, UK. Email: sb31@sanger.ac.uk



CATEGORY WINNER: GENOMICS AND PROTEOMICS

Sam Behjati

Originally from Germany, Sam Behjati read medicine at Oxford University and pursued postgraduate clinical training in London, UK. Funded by the Wellcome Trust, he joined the Cancer Genome Project of the Sanger Institute (UK) for doctoral research. Using a “mutational postcode” that cells acquire as they divide, Behjati has shown that it is feasible to reconstruct the early developmental process in the adult mouse. Building on ideas developed during his Ph.D. preparation, he now aims to define the embryonic origin and fate of childhood cancer cells.

cells, embryonic mutation rates are similar in humans. Thus, it would likely be feasible to apply our approach to precisely reconstructing embryology in humans.

A second key finding was that embryonic cells contributed unequally to the adult mouse. To assess how much of each embryonic mutation was present in the whole mouse, we measured the relative proportion of mutant sequencing reads in the tail, as a proxy. For example, one of the immediate daughter cells of the fertilized eggs made up 75% of the tail of mouse 1 [see the figure (Bot-

tom)], demonstrating a skewed contribution of embryonic cell lineages to the adult mouse. Using a completely different prospective approach, the same observation was made by Magdalena Zernicka-Goetz’s group in 2005 (5). Labeling the first cells of the mouse embryo and tracing their development, they found that one derivative lineage contributed mainly to the embryo, whereas the others formed extra-embryonic structures. The results of these experiments suggest that complex cell fate decisions are taking place in the early embryo.

Our study represents the first demonstration that an exact map of the embryo can be derived retrospectively from mosaic mutations at the resolution of individual precursors and cell divisions. The challenges of scaling up this approach may be overcome by advances in the fidelity of single-cell DNA sequencing. Our experiment is a prelude to a large-scale, systematic effort to apply genomic science to define mammalian embryology. Deriving a precise developmental map of humans in health and in disease has now become a tractable endeavor. ■

REFERENCES AND NOTES

1. E. Shapiro, T. Biezuner, S. Linnarsson, *Nat. Rev. Genet.* **14**, 618 (2013).
2. C. A. Carlson *et al.*, *Nat. Methods* **9**, 78 (2012).
3. A. Wasserstrom *et al.*, *PLOS ONE* **3**, e1939 (2008).
4. S. Behjati *et al.*, *Nature* **513**, 422 (2014).
5. B. Plusa *et al.*, *Nature* **434**, 391 (2005).

ACKNOWLEDGMENTS

I wish to thank my supervisor, M. Stratton, for guiding me through the labyrinth of the genome.

10.1126/science.aal2189

EXTENDED PDF FORMAT
SPONSORED BY



Retracing embryological fate

Sam Behjati (December 1, 2016)

Science **354** (6316), 1109. [doi: 10.1126/science.aal2189]

Editor's Summary

This copy is for your personal, non-commercial use only.

- Article Tools** Visit the online version of this article to access the personalization and article tools:
<http://science.sciencemag.org/content/354/6316/1109.2>
- Permissions** Obtain information about reproducing this article:
<http://www.sciencemag.org/about/permissions.dtl>

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2016 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.