

TECHNIQUE

Genome editing for cell lineage tracing

A thorough understanding of organismal developmental biology and how it is disrupted in disease requires an understanding of cellular lineage histories. A new study describes how genome-editing technology can be leveraged as a powerful means to track cell lineages *in vivo*.

Various genetic techniques have been used to trace cell lineages, each with their own strengths and limitations. Retroviral delivery of DNA barcode libraries can generate a complex starting population of uniquely labelled cells. Following cell growth and differentiation, sequencing across the barcode can reveal the cellular progeny derived from each initially transduced cell, although the high genetic stability of the barcodes means that subclones arising subsequently to the initial labelling cannot readily be distinguished. Alternative methods can track lineages over time based on the accumulation of mutations. However, such systems are constrained by the underlying mutation rate of the system studied and can require extensive sequencing to detect mutations across the genome.

McKenna, Findlay, Gagnon *et al.* reasoned that genome-editing systems could be targeted to a single lineage-tracing barcode to induce serial mutations during development, thus combining the main advantages

of the barcode and somatic-mutation-based approaches. They carried out initial proof-of-principle tests of their genome editing of synthetic target arrays for lineage tracing (GESTALT) method in cultured human HEK293T cells. Two constructs were delivered to cells: one expressing Cas9 nuclease and a guide RNA (gRNA), and the other expressing a DNA barcode consisting of an array of ten different target sites for the gRNA. The target sites ranged from perfect complementarity for the gRNA to sites with numerous mismatches; the aim was that the sites in the DNA barcode would be targeted at different efficiencies by the gRNA and that multiple sequential editing events can be combinatorially detected using a single sequencing read across barcode-derived DNA or RNA.

After 7 days of exposure to the editing system, 1,650 unique barcodes were detected by DNA sequencing on the basis of Cas9-induced small insertions and deletions (indels). Further technical optimization of the system assessed the consequences of reagent delivery methods (DNA transfection versus retroviral transduction), Cas9 expression levels and the numbers and types of target sites in the barcodes. As proof of the biological applicability of the system, >90% of cells exposed to two rounds of editing could be

assigned unambiguously to lineage clades based on the combinations of edits that were shared with other cells.

The investigators applied their system *in vivo* to zebrafish single-cell embryos, injecting a ribonucleoprotein complex of Cas9 protein and ten distinct guide RNAs, each with perfect complementarity to one of the ten different target sites in the transgenic genomic barcode. There were no noticeable detrimental effects on organismal development, indicating that the system itself does not substantially skew developmental trajectories. As for cultured cells, high allelic diversity was generated: from 1,961 cells collected from a single embryo at 30 hours post-fertilization, 1,323 distinct barcodes were detected, 98% of which could be lineage related to other barcodes based on at least one shared edit. Editing activity tapered off after the first 4 hours of embryogenesis; hence, analyses of subclonal events later in development will require the editing machinery to be inducible, delivered later or to have sustained activity. Despite this transient activity, analyses of adult tissues revealed notable insights. In particular, adult tissues are typically derived from a small number of embryonic cell clones (for example, only five distinct barcodes were found across 98% of blood cells examined), but different adult tissues mostly originate from distinct embryonic progenitor cell clones.

It will be interesting to see the applicability of this lineage tracing technique (and future optimizations) in diverse developmental and pathological contexts in different organisms. Given that it requires minimal endogenous cellular machinery, it should be compatible with any organism for which the components can be effectively delivered.

Darren J. Burgess

ORIGINAL ARTICLE McKenna, A., Findlay, G. M., Gagnon, J. A. *et al.* Whole organism lineage tracing by combinatorial and cumulative genome editing. *Science* <http://dx.doi.org/10.1126/science.aaf7907> (2016)

“multiple sequential editing events can be combinatorially detected using a single sequencing read”