

BIOLOGICAL TECHNIQUES

The dynamics of each gene in every cell

A method has been developed to infer whether the expression of each gene in a single cell is increasing or decreasing, and at what rate, using RNA-sequencing data. This tool has many potential applications.

ALLON M. KLEIN

To understand and control complex systems, we must be able to measure the dynamics of their constituent parts. In biology, gene expression is probably the ultimate example of a complex system, with more than 20,000 genes orchestrating the functions of human tissues. However, we lack the tools to measure how genes vary in expression in individual cells over time. In a paper in *Nature*, La Manno *et al.*¹ describe a powerful method that enables the level and rate of change of expression to be estimated simultaneously for each gene in a single cell. The approach has considerable implications for studying cellular dynamics, particularly in disease progression and in complex processes such as embryonic development.

Biologists face an operational problem when trying to understand the dynamic changes in gene expression that occur as cells age, differentiate or become diseased. On the one hand, techniques that enable researchers to broadly measure the expression of all genes in a given cell involve destroying the cell of interest. This prohibits analysis over time and so provides only a snapshot of gene expression. On the other hand, techniques that enable the long-term measurement of gene expression in living cells can be used to track only a limited number of genes².

My group and many others previously attempted to infer the expression dynamics of all genes in a cell from destructive measurements of single cells, by organizing snapshot measurements into continuous ‘trajectories’ that approximate expression dynamics. But, because various gene-expression dynamics could give rise to the same snapshots, even the most sophisticated algorithms can produce incorrect results³.

La Manno *et al.* partially overcame this operational problem by realizing that existing snapshots of gene expression can, in fact, provide bona fide dynamic information. The authors analysed data that were generated by single-cell RNA sequencing. This approach is used typically to measure the abundance

of messenger RNA transcripts for each gene in every cell of a sample. But the researchers showed that these RNA sequences also provided information on whether the expression of each gene was increasing or decreasing at the moment when the measurement was taken.

La Manno *et al.* exploited the fact that freshly transcribed mRNA contains segments that are later cut out (spliced) during the formation of mature mRNA. For a gene that is stably expressed, a small fraction of its

mRNA will always be found in the immature, unspliced form, because older transcripts are replaced continuously with new ones. When a gene has just been activated, for a short time there will be a much higher proportion of immature transcripts. Conversely, when the expression of a gene is repressed, the proportion of short-lived, unspliced transcripts will drop before the longer-lived mature mRNA transcripts decay. Therefore, for each gene in a cell, the ratio of unspliced mRNA to spliced mRNA can be used to directly infer instantaneous expression dynamics — that is, the ‘RNA velocity’ of each gene, which can then be used to deduce the cellular changes that are taking place in a tissue (Fig. 1).

This approach has already been used on bulk RNA-sequencing data sets^{4,5}. La Manno *et al.* realized that the method can be applied to single-cell data, for which it is considerably more useful. These data provide a much higher-resolution picture of dynamic processes — particularly in complex tissues, which contain many cell types with varied gene-expression patterns that are amalgamated in bulk

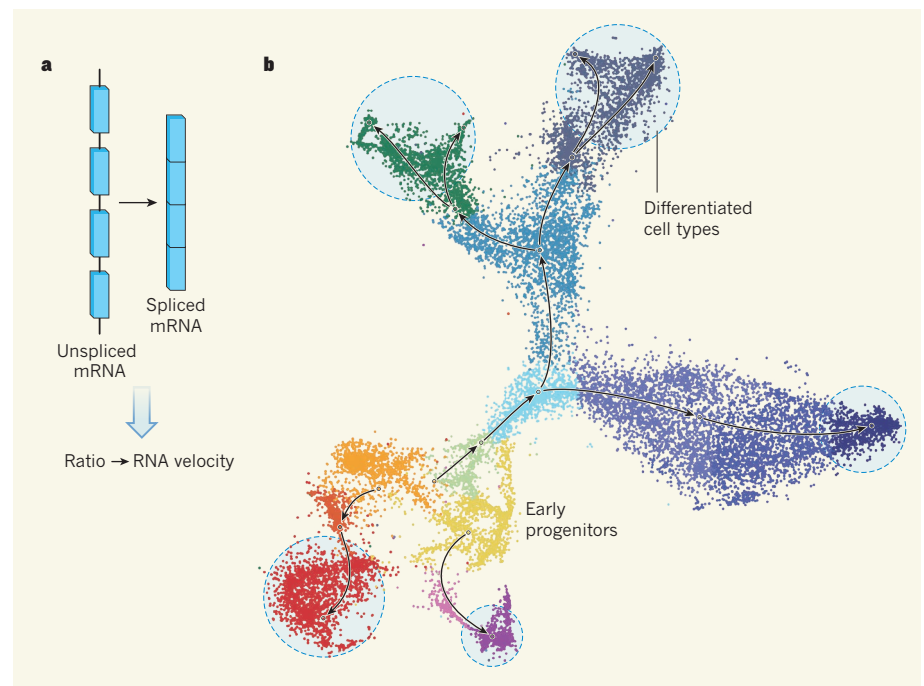


Figure 1 | Measuring dynamic changes in gene expression across complex tissues. **a**, As messenger RNA matures, sections of the immature transcript are removed — a process called splicing. When the expression of a gene increases, a transient increase in the proportion of immature, unspliced transcripts compared with that of mature, spliced transcripts is observed in the cell. By contrast, a higher proportion of spliced transcripts is seen for a short time when expression of the gene decreases (not shown). La Manno *et al.*¹ measured the ratio of unspliced to spliced transcripts for each gene in a single cell to calculate a quantity called the RNA velocity, which reveals how gene expression is changing. **b**, By measuring RNA velocity in thousands of cells in a tissue (here, in neurons in the developing mouse brain), the authors could generate maps that show not only how closely related cells are to one another (with closeness indicated by similar colours), but also which cells they will become similar to in the future (indicated by arrows), according to the gene-expression changes they are undergoing. RNA velocity successfully tracks early progenitors (orange and yellow) that eventually give rise to a range of differentiated cell types (blue dashed circles). (b adapted from Fig. 3c of ref. 1.)

analyses. The authors found that existing algorithms for the analysis of data from single-cell RNA sequencing routinely discard information about immature, unspliced mRNA. By completely reworking their computational pipelines to salvage these data, they could recover information about both spliced and unspliced forms of each transcript, and therefore predict RNA velocity.

As is often the case, much effort and technical ingenuity were required for La Manno *et al.* to translate their initial idea into a robust set of working algorithms. Among the challenges that they had to overcome was the fact that measuring gene expression in single cells can be noisy. This is because most of the mRNA molecules in each cell are lost in the attempt to sequence them, leaving researchers with only a patchy picture of gene expression. Another challenge was determining how to infer the baseline ratio of spliced to unspliced transcripts for each gene when it is undergoing stable transcription. The authors needed to apply cutting-edge approaches in statistics and machine learning to solve these problems.

La Manno and co-workers beautifully demonstrate the usefulness of their approach using both published and newly collected data sets. For instance, they showed that RNA velocity could accurately detect the increases

and decreases in gene expression that cells in the embryo are known to undergo as they differentiate from a cell type called a neural crest cell into chromaffin cells of the adrenal glands. The authors also used RNA velocity to investigate gene-expression dynamics in the developing hippocampus of the mouse brain, during intestinal stem-cell differentiation and more. This array of examples suggests that the method will have wide value. Among the group's most important achievements was the analysis of human embryonic tissue, in which other forms of dynamic measurement would be very difficult, or even impossible, to carry out because of the technical and ethical issues that are associated with studying living human embryos.

Developing an analysis of RNA velocity for single cells is a major breakthrough. But, of course, it has limitations. By its nature, RNA velocity cannot actually track a given cell over time, it is limited to the study of mRNA, and it does not provide information about the spatial organization of cells. These limitations could be restrictive when exploring phenomena in stem-cell biology, embryonic development or the onset of disease, which are likely to depend on the lineage and arrangement of cells, and which can be driven by mechanisms other than transcription, including protein

phosphorylation. The method gives only a probabilistic description of cell dynamics, which is pieced together from instantaneous velocities. As a result of these limitations, there is little doubt that the spatio-temporal expression dynamics of genes will continue to be studied using complementary methods such as live imaging.

Nonetheless, the ability to infer true, instantaneous RNA velocities in single cells is a leap forward for studies of gene-expression dynamics on the whole-genome scale. Indeed, the authors' approach has already been applied by other researchers⁶. In the immediate future, I can foresee RNA velocity easily becoming an essential tool for single-cell analysts. ■

Allon M. Klein is in the Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA.
e-mail: allon_klein@hms.harvard.edu

1. La Manno, G. *et al.* *Nature* <https://doi.org/10.1038/s41586-018-0414-6> (2018).
2. Specht, E. A., Braselmann, E. & Palmer, A. E. *Annu. Rev. Physiol.* **79**, 93–117 (2017).
3. Weinreb, C., Wolock, S., Tusi, B. K., Socolovsky, M. & Klein, A. M. *Proc. Natl Acad. Sci. USA* **115**, E2467–E2476 (2018).
4. Zeisel, A. *et al.* *Mol. Syst. Biol.* **7**, 529 (2011).
5. Gray, J. M. *et al.* *PLoS ONE* **9**, e89673 (2014).
6. Plass, M. *et al.* *Science* **360**, eaag1723 (2018).