50 Years Ago

This year is the bicentenary of the granting of patents for two inventions which played a crucial part in making Britain the most important nineteenth century industrial power. In 1769, James Watt patented his separate condenser, which proved to be the greatest single improvement ever made in steam engines, and Richard Arkwright patented his spinning machine, which, strictly speaking, was … a successful exploitation of a much earlier machine which never quite worked. To mark the occasion, the Science Museum in London has arranged a characteristically subdued exhibition of the two original patents … a little biographical material … and eight or nine cases containing recent and contemporary models and drawings of Watt's work and Arkwright's original spinning machines.

From *Nature* 19 July 1969

100 Years Ago

With the view of honouring some of those who helped to win the war … the North-East Coast Institution of Engineers and Shipbuilders held a Victory meeting … Lady Parsons read a paper on women's work in engineering and shipbuilding during the war. … There is no doubt that many women developed great mechanical skill and a real love of their work. The engineering industry is again barred to women by an agreement made between the Treasury and the trade unions … The meeting agreed with Lady Parson's condemnation of the Labour party, which, while demanding full political equality for women and their right to sit in the House of Lords and to practise at the Bar and as solicitors, will not grant to women equality of industrial opportunity.

From *Nature* 17 July 1919

donor cells in individuals with a type of blood cancer who received stem-cell transplants[3]. However, combined approaches have not been extensively used to examine the effects of mutations in cancer-associated genes on blood-cell development.

Nam *et al.* designed a method called 'genotyping of transcriptomes' (GoT) by combining an existing platform for profiling gene expression[3] with a technique for amplifying a specific genetic sequence to detect mutations in it (Fig. 1). They used this method to analyse thousands of progenitor cells sampled from the bone marrow of five individuals with a form of blood cancer that is caused by mutations in the *CALR* gene, and that is characterized by overproduction of platelet cells. GoT enabled the authors to ascertain which of the sampled cells carried a *CALR* mutation and which did not.

The authors used a statistical analysis to 'group' the sampled progenitor cells into different types on the basis of their gene-expression profiles (Fig. 1). All of the identified types contained both cells with and without the *CALR* mutation. However, *CALR*-mutant cells were more likely to follow certain differentiation pathways and therefore to become certain types of blood cell. Furthermore, Nam and colleagues found that the effects of the mutation, when present in the progenitor cells, were noticeable only at later stages of cellular differentiation; the progeny of *CALR*-mutant cells were more abundant than the progeny of their non-mutant counterparts and had a distinct gene-expression profile. Such observations would not have been possible using standard techniques, which demonstrates the value of this method.

Although GoT has its limitations, they can probably be addressed by adapting it to new single-cell workflows. First, GoT currently requires that the identity of the mutated gene, or a small set of potentially mutated genes, is known in advance. As an example, the authors used a multiplexed version of their analysis that can simultaneously target multiple prespecified parts of the genetic sequence to probe three genes. If no specific mutations, genes or regions of the genome have been prespecified for analysis (for example, on the basis of an association with disease progression), multiplexed analyses can, in theory, be used to cover larger panels of genes; however, this might not be cost-effective.

Second, GoT is less effective at detecting mutations that occur near the middle of a gene than those that occur near the ends. One solution to this problem would be to use a lower-throughput platform that allows the analysis of full-length RNA transcripts in single cells[4,5]; in theory, this approach could detect mutations anywhere in the RNA-encoding parts of genes. Nam *et al.* present an alternative approach by showing that a technique called nanopore sequencing, in which full-length transcripts are sequenced by passing them through a tiny pore, is compatible with their high-throughput platform.

Third, GoT cannot detect mutations in genetic sequences that are not transcribed but that may affect gene expression. Investigation of such sequences might be possible by combining GoT with a technique that measures how accessible certain DNA sequences in a cell are to enzymes[6].

A recent paper[7] used a different high-throughput approach to implement a similar targeted-amplification strategy to study a blood cancer that is thought to be partly caused by disruption of haematopoiesis by progenitor-cell mutations. The authors of that paper also identified a set of genes that were co-expressed only in malignant progenitors (that is, progenitor cells with a cancer-associated mutation), and described a machine-learning approach that used gene-expression data to distinguish malignant cells from non-malignant ones, even without using prespecified gene-sequence information. It would be interesting to see whether the same machine-learning approach could use Nam and colleagues' gene-expression data to distinguish the malignant cells from non-malignant cells. Obtaining gene-sequence information from single cells remains more challenging than assessing gene expression; therefore, a method for predicting malignancy solely on the basis of single-cell gene expression would have vast clinical implications.

"Understanding how mutations in progenitor cells lead to changes in the production of different cell types is a key question."

In theory, GoT and similar approaches could be used to study any cancer. They have the potential to precisely determine the effects of mutations in known genes on downstream cell-development states and to establish whether certain mutations are sufficient to induce cancer. These insights, in turn, could shed light on the mechanisms that underlie the evolution of clonal lineages of cells in cancer. ■

Siddharth Raju *and* **Chun Jimmie Ye** *are in the Department of Medicine, Institute for Human Genetics, and at the Bakar Institute for Computational Health Sciences, University of California, San Francisco, San Francisco, California 94143, USA.*
e-mail: jimmie.ye@ucsf.edu

1. Nam, A. S. *et al. Nature* **571**, 355–360 (2019).
2. Kang, H. M. *et al. Nature Biotechnol.* **36**, 89–94 (2018).
3. Zheng, G. X. Y. *et al. Nature Commun.* **8**, 14049 (2017).
4. Gupta, I. *et al. Nature Biotechnol.* **36**, 1197–1202 (2018).
5. Macaulay, I. C. *et al. Nature Methods* **12**, 519–522 (2015).
6. Buenrostro, J. D. *et al. Nature* **523**, 486–490 (2015).
7. van Galen, P. *et al. Cell* **176**, 1265-1281 (2019).

This article was published online on 3 July 2019.