groups, ranging from −0.3 °C to +0.6 °C.

Digging further into these differences, Chan *et al.* realized that measurements from Japanese ships in the North Pacific suddenly became about 0.35 °C cooler after 1930 when compared with measurements from other countries. This change was caused by the Japanese switching from recording temperatures in whole-degrees Fahrenheit to taking readings in degrees Celsius and then dropping any numbers after the decimal point. The authors identified a similarly large change in the North Atlantic that is associated with German readings, but the cause of this change is less clear.

Chan and colleagues' results suggest that scientists have been overestimating warming in the North Atlantic and substantially underestimating warming in the North Pacific during the early twentieth century because of not fully accounting for biases in bucket measurements (Fig. 1). These findings bring the difference in estimated warming between the two regions in line with projections from climate models. However, there are still large differences between models and observations in the overall rate of global ocean warming during this period.

The authors' approach of comparing groups of proximate-ship measurements is conceptually similar to that used in identifying problems in the land temperature record, whereby each weather station is compared with its neighbours to find and remove localized biases[6]. The method offers an innovative solution to the lack of good ship metadata during the early twentieth century and provides a major advance in our understanding of historical ocean measurements.

This study, and recent major updates to the SST record at the UK Met Office's Hadley Centre[7], provide a useful reminder that large systematic biases might remain in our observational temperature records. Improved quantification of these biases is still a key technical challenge for researchers, and will help to address questions about the performance of climate-model simulations of the past and the role of intrinsic climate variability in historical temperature change. ■

Zeke Hausfather *is in the Energy and Resources Group, University of California Berkeley, Berkeley, California 94720, USA.*
e-mail: hausfath@berkeley.edu

1. Haustein, K. & Otto, F. E. L. *J. Clim.* https://doi.org/10.1175/JCLI-D-18-0555.1 (2019).
2. Chan, D., Kent, E. C., Berry, D. I. & Huybers, P. *Nature* **571**, 393–397 (2019).
3. Kent, E. C. *Bull. Am. Meteorol. Soc.* **98**, 1601–1616 (2017).
4. Lenssen, N. J. L. *et al. J. Geophys. Res. Atmos.* https://doi.org/10.1029/2018JD029522 (2019).
5. Kennedy, J. J., Rayner, N. A., Smith, R. O., Parker, D. E. & Saunby, M. *J. Geophys. Res. Atmos.* **116**, D14103 (2011).
6. Menne, M. J. & Williams, C. N. Jr. *J. Clim.* **22**, 1700–1717 (2009).
7. Kennedy, J. J., Rayner, N. A., Atkinson, C. P. & Killick, R. E. *J. Geophys. Res. Atmos.* https://doi.org/10.1029/2018JD029867 (2019).

GENETICS

How mutations express themselves

A method for detecting mutations and measuring gene-expression levels in the same cell has enabled an investigation into the effects of mutations in a specific gene on the emergence of a form of blood cancer. SEE ARTICLE P.355

SIDDHARTH RAJU & CHUN JIMMIE YE

The cells that circulate in the bloodstream perform various functions and, in adults, are derived from progenitor cells in the bone marrow. Mutations in the DNA sequences of progenitor cells can lead to changes in blood-cell development, sometimes resulting in cancer. Owing to technical constraints, elucidating the effects of progenitor mutations on blood-cell development has been challenging. On page 355, Nam *et al.*[1] report a method for detecting mutations and measuring gene expression in individual blood progenitor cells, and use it to analyse a mixture of progenitors with or without mutations in a cancer-linked gene. They show that progenitors that have the same mutation can give rise to cells with different gene-expression profiles.

Haematopoiesis — the process through which mature blood cells are formed from progenitors — is tightly regulated. The 'decision' that progenitor cells make as to which cell type to become is generally determined by the signals that they receive from their immediate surroundings. However, mutations that sometimes arise in these progenitor cells can result in the signals being blocked, overamplified or simply ignored, resulting in the enrichment or depletion of specific cell types and, in some cases, production of cancerous clones. Understanding how mutations in progenitor cells lead to changes in the production of different cell types is a key question.

Investigating how mutations in a progenitor cell affect its gene expression, and thus its identity and function, has been highly challenging, largely because mutant cells can be rare and often do not express molecular markers that can be used to separate them physically from non-mutant cells. Strategies to simultaneously detect genetic differences and measure gene expression in single cells have been used to assign cells from a mixture of immune blood cells to their human donor of origin[2], and to study changes in populations of host and

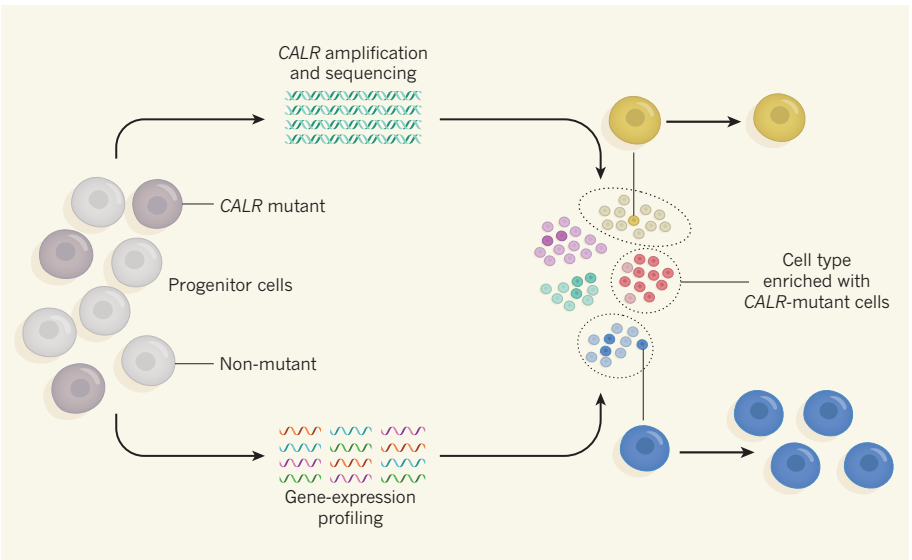

Figure 1 | An analysis of mutation status and gene expression in single cells. Nam *et al.*[1] sampled progenitor cells that give rise to blood cells from individuals who have a type of blood cancer that is caused by progenitor cells with mutations in the *CALR* gene. To distinguish mutant from non-mutant cells, the authors amplified and sequenced the *CALR* gene of individual cells. The authors also measured the levels of gene expression in each cell. They identified different cell types on the basis of a statistical analysis of the cells' gene-expression profiles (dotted circles represent statistical, rather than physical, cell groupings), and examined which of the cells in these different types had *CALR* mutations. Certain cell types were enriched in *CALR*-mutant cells, and *CALR* mutations had different effects (for example, on proliferation) in cells of different types.

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.