Title
Single-cell analysis tools for drug discovery and development.

Journal
Nature reviews. Drug discovery, 15(3)

ISSN
1474-1776

Authors
Heath, James R
Ribas, Antoni
Mischel, Paul S

Publication Date
2016-03-01

DOI
10.1038/nrd.2015.16

Peer reviewed

Single-cell analysis tools for drug discovery and development

James R. Heath[1], Antoni Ribas[2] and Paul S. Mischel[3]

Abstract | The genetic, functional or compositional heterogeneity of healthy and diseased tissues presents major challenges in drug discovery and development. Such heterogeneity hinders the design of accurate disease models and can confound the interpretation of biomarker levels and of patient responses to specific therapies. The complex nature of virtually all tissues has motivated the development of tools for single-cell genomic, transcriptomic and multiplex proteomic analyses. Here, we review these tools and assess their advantages and limitations. Emerging applications of single cell analysis tools in drug discovery and development, particularly in the field of oncology, are discussed.

There have been significant recent advances in the development of single-cell analysis tools. For example, approximately 5 years ago, patch-clamping electrophysiology methods[1], fluorescence *in situ* hybridization[2,3], flow cytometry[4,5] and enzyme-linked immunospot[6] assays were among the few single-cell molecular analysis tools available. Most of those methods could only analyse between 1 and 3 molecules from a given cell, although multicolour flow cytometry could capture approximately 12 cell surface protein markers[7].

This landscape is rapidly changing, and several technologies that can comprehensively analyse the single cell at the molecular level have now emerged. For example, some single-cell tools and methods can assay reasonably large numbers (>40) of secreted proteins[8], equally large numbers of cell surface markers[9] and elements of phosphoprotein signalling pathways[10,11]. In addition, single cells can now be analysed for the genome at focused[12,13] or high coverage[14], the transcriptome at sparse coverage[15,16] or the entire transcriptome with moderate[17] or high[18] cell statistics.

Additional reports in which integrated measurements of genes and transcripts[19], limited numbers of proteins, transcripts[20,21] and genes[22], and panels of proteins and metabolites[23] from single cells have also appeared. Microfluidics methods permit molecular analysis to be correlated with measurements of specific cellular functions (such as motility) or enable the analysis of defined, small populations of cells (that is, two to three cells)[24–26]. Microfluidic designs can also permit cell analysis within highly controlled, custom environments[27–29], or can allow for non-destructive cell analysis so that cells identified as interesting, such as B cells producing specific antibodies, can be harvested for further use[30,31]. Two recent tissue staining methods, *in situ* RNA profiling via sequential hybridization[32–34] and proteomic analysis via ion beam profiling[35], enable the analysis of single cells within fixed, intact tissues, with a level of multiplexing that significantly exceeds traditional immunohistochemical staining methods. The level of analyte quantification varies from measurements that yield copy numbers per cell[18,32,36] to relative quantification between cells. Many of these methods generate relatively new types of data and are therefore being integrated with new computational approaches[37–41]. In fact, the development of computational tools that can analyse increasingly large single-cell datasets is lagging behind the advances in experimental methods.

Although these diverse and rapidly evolving single-cell technologies provide remarkable opportunities for drug discovery and development, they also provide a deluge of information for the non-technologist. This Review is therefore intended to serve as a guide for the non-specialist. Here, we describe the state of the art of single-cell biology tools for different analyte classes and discuss the new types of biological information that can be gleaned through the use of these tools, highlighted using three examples. To illustrate the broader application of these emerging technologies, these tools are placed within the context of two classes of cancer therapies. The first is the development and use of targeted inhibitors for treating heterogeneous tumours. The second is cancer immunotherapy, which is an area in which several single-cell analysis tools are already having important roles.

Single-cell analysis tools can be grouped according to the measured analytes, that is, genomics-, transcriptomics-, proteomics- or metabolomics-based approaches, or by a combination of these. It is anticipated that the methods described here are likely to emerge in the marketplace

[1]*California Institute of Technology Division of Chemistry and Chemical Engineering, MC 127–72, 1200 East California Boulevard, Pasadena, California 91125, USA.*
[2]*Department of Medicine, University of California, Los Angeles, 10833 Le Conte Avenue, Los Angeles, California 90095, USA.*
[3]*Ludwig Institute for Cancer Research San Diego, Department of Pathology and Moores Cancer Center, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA.*

Correspondence to J.R.H.
heath@caltech.edu

doi:10.1038/nrd.2015.16
Published online 16 Dec 2015

within a couple of years, although earlier generation variants are, in many cases, already commercially available as whole platforms, commercial services or through the purchase of essential reagents.

Single-cell analysis tools

Below, we review the state of the art in analytical tools designed for single-cell genomics, transcriptomics and proteomics, with a particular emphasis on quantitative, highly multiplex assays that can perform measurements on many single cells in a given experimental run. Many of the reviewed methods are very new.

Single-cell genomics. The rapid technological advances in DNA sequencing tools have exposed the whole genome, the exome and the transcriptome for single-cell analysis. For single-cell whole-genome sequencing[12,42,43], the genome must be amplified before sequencing. In principle, amplification can be performed with PCR-based whole-genome amplification methods[44], but such methods are prone to bias because random genes can be over- or under-amplified by the nonlinear PCR process[45]. A commonly used alternative is the multiple displacement amplification method, which is a technique that utilizes the φ29 DNA polymerase enzyme for DNA synthesis[46], and can amplify DNA isothermally at 30 °C. Multiple displacement amplification provides an improved representation of the entire genome, but the φ29 enzyme is still a nonlinear amplifier (similar to PCR), and can therefore yield bias. Such bias in turn makes it difficult to discern copy number variations and single-nucleotide variations, although Dago and co-workers have reported measurements of such quantities from single circulating tumour cells (CTCs) originating from prostate cancer[47]. A second whole-genome amplification approach, called multiple annealing and looping-based amplification cycles (MALBAC), has been recently reported[48]. MALBAC is designed so that the initial polymerase amplification steps yield an amplicon that cyclizes due to complementary sequences incorporated into the 3′ and 5′ ends, which makes the amplicon unavailable as a template. The cyclization of the amplicon keeps the initial genome amplification process linear and reduces amplification bias. As a result, copy number variations and single-nucleotide variations can be reliably quantitated at the single-cell level. As an illustrative example, MALBAC has been extended to the analysis of CTCs from patients with lung cancer[49]. For certain challenging genes, such as oncogenes with multiple variants[50] or the T cell receptor-α (TCRα) and TCRβ genes (*TCRA* and *TCRB*, respectively), nested PCR methods[51,52] coupled with Sanger sequencing are used. Recent, highly parallel, multi-step RT-PCR-based techniques, coupled with next-generation sequencing tools, now allow such sequences to be determined from many (100 or more) single cells in parallel[53].

Various target-enrichment strategies have been developed to broadly select genomic regions of interest for sequencing[54]. For example, as methods for exome sequencing have become standardized[55,56], they have been extended to single-cell analysis[57]. Exome sequencing is a relatively cost-effective procedure that yields an enriched dataset of highly penetrant variants, such as those that are relevant to genetic disorders or diseases that exhibit a genetic instability, such as many cancers. Examples of single-cell investigations include capturing the genetic heterogeneity of tumours[57,58] or comparing CTCs with the originating tumour or metastatic lesion[59]. Exome sequencing is a technique of rapidly increasing relevance to immunotherapy, as discussed further below.

Single-cell transcriptomics. Although the analysis of gene expression at the single-cell level dates back to the early 1990s[60], the field has rapidly advanced over the past 5 years, with RNA sequencing (RNA-seq) exploiting the success of next-generation sequencing tools[61]. Indeed, RNA-seq has advanced at such a rapid pace that a new report emerges almost every month describing a new set of protocols that enable an increasingly deeper and more quantitative analysis of larger numbers of single cells[18,39,62–65]. Applications of RNA-seq include the analysis of immune cells[64] and CTCs[66], and capturing the transcriptional heterogeneity of various healthy[16,67] and diseased tissues[17]. The basic biochemical method of RNA-seq is PCR, but the major technical challenges have been to engineer contamination-free methods that can account for PCR bias correction and yield absolute quantification. This goal has been best accomplished through the combined use of microfluidics platforms[63], including microdrop technologies[68,69] and molecular barcoding techniques[62,69]. The microfluidics character of these approaches implies that individual cells are isolated in volumes ranging from a few tens to a few hundreds of picolitres, and this produces several advantages. First, molecular diffusion times within such small volumes are short, which can then significantly shorten the times required for chemical reactions that are part of the processes flow. Second, the small volume raises the relative concentration of the cellular analytes being investigated and lowers the copy numbers of any molecular contaminants. Finally, small volumes limit reagent costs and allow many cells to be interrogated in parallel.

Two very recent quantitative single-cell transcriptomic methods — CytoSeq and inDrop — are conceptually similar, but distinct in practice (FIG. 1). Fan *et al.*[18] reported the CytoSeq technique, which utilizes dilute cell loading into 20 pl volume microwells. Each microwell contains a 20 μm magnetic bead that is functionalized with many oligonucleotide primers, each containing a universal PCR priming site, a combinatorial cell label (the barcode), a unique molecular index (UMI)[70] and an mRNA capture sequence. All primers on each bead contain the same cell label but incorporate a diversity of molecular indices. Many mRNA molecules from a lysed cell are captured on a single bead, and all beads are combined for amplification and sequencing. Each sequence carries the barcode (single-cell identity), the molecular index (one index per transcript) and the gene identity, thus yielding a relatively deep, bias-free and quantitative analysis of the transcriptome from many single cells in parallel.

Two droplet-based microfluidics variants of this barcoding approach for single-cell transcriptomics are the DropSeq method[69] and the (simultaneously published)

Whole-genome amplification
A method, first reported using PCR by Arnheim's group, for nonselectively amplifying all DNA sequences present in a given sample, including a single cell.

Multiple displacement amplification
A non-PCR based, room temperature DNA amplification technique reported by Lasken's group that is commonly used for whole-genome amplification.

Multiple annealing and looping-based amplification cycles
(MALBAC). A PCR-type approach reported by Xie's group in which the enzymatic amplification of cDNAs proceeds via a linear process.

Exome sequencing
Genome sequencing that is limited to only the small fraction (1 %) of the genome that is protein encoding.

RNA-sequencing
(RNA-seq). Also called whole transcriptome shotgun sequencing, RNA-seq is a method for analysing the transcriptome of a sample using next-generation sequencing tools.

Molecular barcoding
An approach through which a DNA sequence or some other molecular identifier is used as an identifier of a specific cell or a specific transcript generated by that cell.

Cytoseq
A microchip-based single-cell transcriptomics method reported by Fodor's group at Cellular Research in 2015.

inDrop
A nanodrop-based single-cell transcriptomics method reported by Klein and others in 2015.

Unique molecular index
(UMI). A molecular barcode used to identify a specific transcript from a specific cell.

DropSeq
A nanodrop-based single-cell transcriptomics method reported by Macosko and others in 2015.

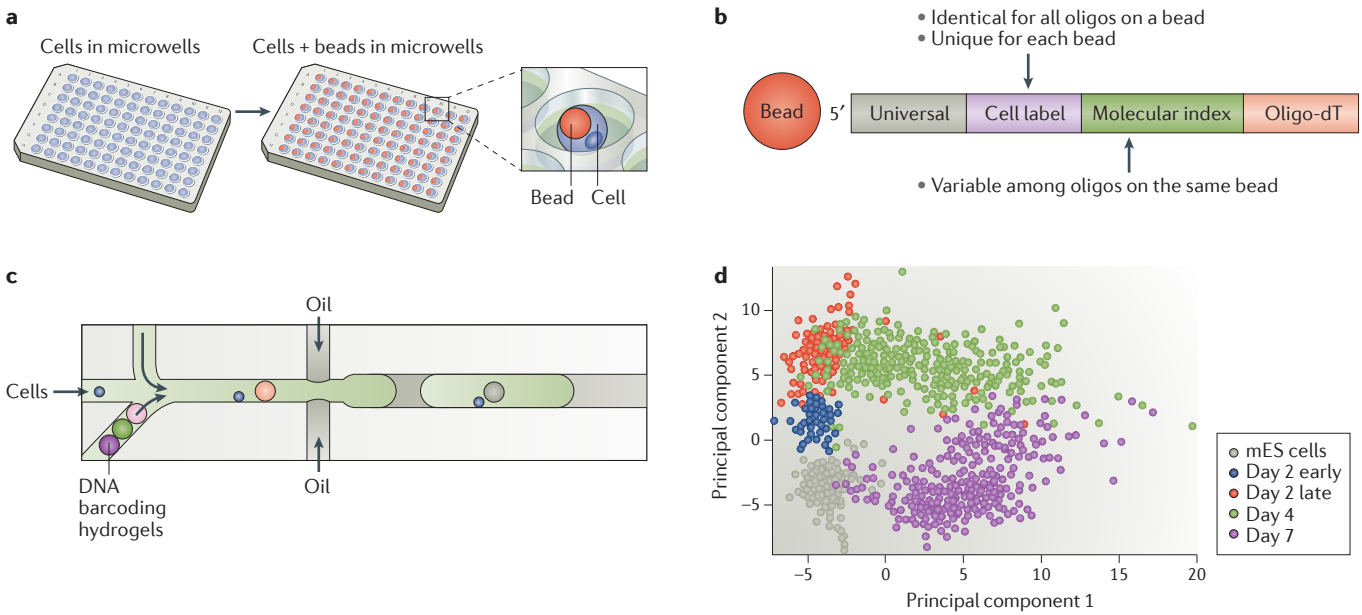


Figure 1 | **Quantitative single-cell transcriptomic methods.** Two separate but conceptually similar methods and with similar throughput capabilities are illustrated in this figure, along with representative data. **a** | The CytoSeq method is based on isolating individual cells within 30 µm diameter (20 pl volume) wells, and then placing into each well a single barcoded bead. **b** | Each barcoded bead is designed with the shown structure. Each bead contains tens to hundreds of millions of distinct oligonucleotide (oligo) primers that are each composed of a barcode that identifies the bead (and thus the single cell) plus a unique molecular index (UMI) that is associated with a particular mRNA capture sequence. After bead and cell colocalization within a well, cells are lysed and mRNAs are captured via hybridization onto specific bead-bound oligonucleotides. The beads are then all removed from the multiwell plate, and all amplification reactions are carried out in a single tube. **c** | The microdrop-based inDrop technique for single-cell transcriptomics. For this method, single cells are entrained into a single droplet along with a hydrogel microsphere. Each hydrogel microsphere contains photocleavable oligonucleotide primers that have a similar construction to the bead shown in part **b**, while the droplets contain the cell lysis buffers and reverse transcription reagents. The whole process from cell capture and lysis to signal amplification occurs separately in each droplet. **d** | A snapshot of representative data from an inDrop study of the kinetics of differentiation of mouse embryonic stem (mES) cells following leukaemia inhibitory factor (LIF) withdrawal. For this plot, datasets representing five time points are analysed using principal component analysis to reveal asynchrony in mES cell differentiation. Each dot represents a single cell. Figure parts **a** and **b** are adapted from Fan, H. C., Fu, G. K. & Fodor, S. P. Combinatorial labelling of single cells for gene expression cytometry. *Science* **347**, 1258367 (2015). Reprinted with permission from AAAS. Figure parts **c** and **d** are adapted with permission from REF. 68.

inDrop approach[68]. The basic concept of droplet microfluidics is to use microfluidic channel designs and flow control to combine oil and water so that the water divides into sub-nanolitre volume droplets separated by oil. Each of those nanodrops can be seeded with, for example, a cell, a barcoded microbead (or equivalent) or cell lysis reagents, such that each nanodrop comprises a self-contained reaction vessel. Advanced microchip designs allow virtually the entire process, from cell introduction to delivery of reagents for sequencing, to be automated on a microchip that is approximately the size of a microscope slide. The DropSeq method was utilized for the analysis of ~45,000 single mouse retinal cells, which is a testament to the scalability of droplet microfluidics, and similar to the capabilities of the CytoSeq method.

A common concern with single-cell methods is the relationship between what is measured and the copy numbers of the analyte that were actually in the cell. Even genetically identical cells, cultured side by side, will naturally exhibit significant variations in copy numbers of transcripts, proteins, metabolites and other analytes[38,71] (BOX 1). For any analyte, capture efficiency is always an issue, and can be highly challenging to quantify. For transcriptomics, this concern is complicated by the fact that the actual mRNA transcript is not measured. Instead the mRNA transcript is a cDNA complement, amplified to many copies, that provides the input into the sequencer. Different mRNAs can be differentially amplified, and noise can be amplified along with signal[45,72–74]. Of course, having a single-cell technique that captures the biological heterogeneity of the cells under study, rather than the measurement noise of the technique itself, is advantageous. Various methods have been used to increase and/or characterize the quantitative nature of single-cell transcriptomics[75].

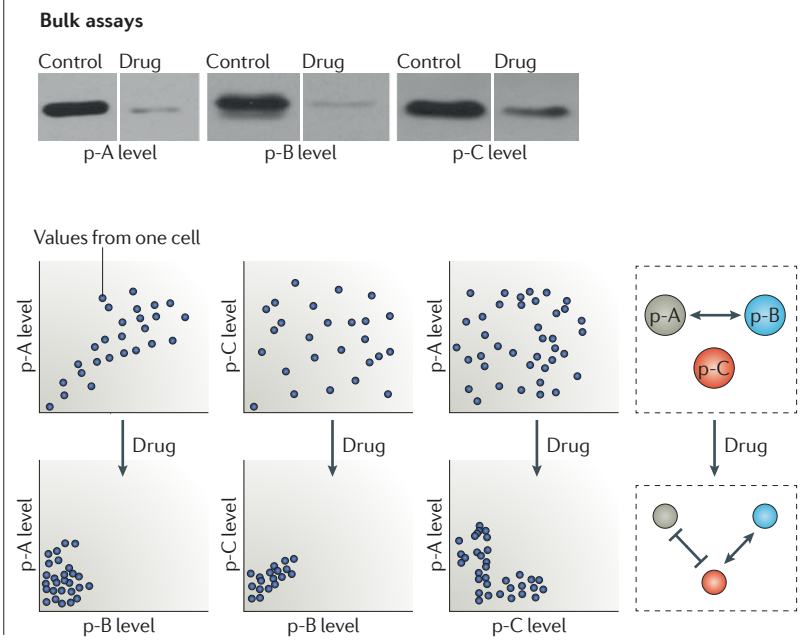
The UMI[62,70,71] mentioned in the description of CytoSeq, which is also used by inDrop and DropSeq, is a protocol designed to limit amplification bias by associating a unique molecular signature to each mRNA copy that is captured. A related method was reported by Fu and co-workers[76,77]. Thus, if ten copies of a specific transcript are captured from one cell, each will have the same

Nanodrops
Microfluidics methods in which individual assays are carried out in isolated nanolitre-size droplets of water, separated from one another by oil.

Box 1 | **The biophysical interpretation of single-cell data**

The ability to quantitate the level of analytes from single cells provides fundamentally new insight into cellular biology. For example, the abundance distribution of an analyte, as tabulated across many single cells, is also called the fluctuations of that analyte, and represents a unique single-cell measurement. A certain width of the fluctuations is fundamental and unavoidable, and is reflective of the statistics associated with the many steps through which signals are processed by gene and protein networks. For purely stochastic[96] processes, the distribution width should narrow as the square root of the average copy numbers per cell of the analyte increases[162]. However, most analytes will not behave according to this limit and, in fact, the shapes of analyte distributions can reveal new biology, such as evidence of bistable steady states[163], evidence that the cells are in a stable steady state[38] or are unstable and responding to a perturbation (for example, a drug). A relevant example of bistability might be a cell population that is composed of both a quiescent state and an active state[164,165], and thus yields differential responses to drugging.

Measurements of multiple analytes from the same single cells can be used to extract quantitative analyte–analyte correlations (and anti-correlations). Again, this is a unique single-cell measurement. Consider, for example, the levels of the three hypothetical phosphoproteins (p-A, p-B and p-C) shown in the figure. These proteins represent a small signalling network within a cell. Stimulation (or drugging) of the cell may collectively repress these phosphoprotein levels, as is reflected in the bulk immunoprecipitation assays. However, a more in-depth picture of the signalling is revealed by an analysis of a statistical number of single cells, such as is presented in the two-dimensional scatter plots. Note that in the plots for the undrugged cells, all phosphoprotein levels are high, but only p-A and p-B are strongly correlated. On drugging, all phosphoproteins are repressed, but p-A and p-B are non-correlated, p-C and p-B are strongly correlated, and p-A and p-C are anti-correlated. This inferred correlation network is shown in the figure, in which the protein levels are indicated by the sizes of the spheres and the correlations are indicated by the edges. Correlation, of course, does not mean causation, but a correlation network generated at the single-cell level can provide a rich set of testable hypotheses that may ultimately allow the chemical kinetic relationships that comprise a signalling network to be extracted. In principle, if one knows these relationships, then one can make accurate predictions regarding how a specific drug will disrupt the cellular signalling machinery. Improved measurement quantification provides significant additional value. This is because most signalling cascades actually behave as excitable devices with built-in excitability thresholds, enabling them to integrate diverse temporal and spatial inputs to produce specific signalling responses[166]. In other words, the outputs of a signalling cascade are not typically linearly dependent on the inputs, and quantitative assays permit such input–output relationships to be more accurately defined.



barcode but a different UMI; therefore, the copy numbers of a given mRNA captured is simply the number of unique UMIs for a given barcode. The use of UMIs, although a significant step towards absolute quantification, does have limitations for counting low copy number transcripts[71]. A second major issue is that of capture efficiency, which can vary from <5%[71] to ~20%[78], and can be assessed by counting mRNA copies using UMIs relative to those recorded using fluorescence *in situ* hybridization[71]. Of course, a low capture efficiency will have a correspondingly large variance across many mRNAs or between different single cells, especially for low copy number transcripts. Thus, quantitating and increasing capture efficiency is an analytical frontier of the field.

Multiplex single-cell proteomics methods. Multiplex single-cell proteomic methods (FIG. 2; TABLE 1) are classified as either microfluidics platforms or flow cytometry[7] or mass cytometry[9] (for example, CyTOF) tools. Each of these methods rely on antibodies; therefore, in contrast to mass spectrometry proteomics of bulk samples[79,80], single-cell proteomics methods cannot yet serve as discovery-level tools. Among the microfluidics platforms, the microengraving technique[31,81], single-cell barcode chips (SCBCs) and single-cell western blottings[82] (scWesterns) provide the most advanced capabilities. Several alternative approaches, typically with reduced levels of multiplexing, have been reported, including high-throughput microdroplet-based screening approaches[29,83–86], some of which have been reviewed elsewhere[87].

For analysis and cell sorting based on cell surface markers, flow cytometry-based fluorescent-activated cell sorting (FACS) is the mature single-cell proteomics method[4], and interfaces with almost all other single-cell methods described in this Review. FACS is routinely used to analyse and sort viable cells based on six or more cell surface markers, and is therefore tremendously useful for purifying cellular phenotypes for subsequent analysis.

The analysis of cellular function at the molecular level, or the specific influence of drugs on that function, typically requires the analysis of functional analytes, such as phosphorylated kinases or secreted cytokines, apoptotic or proliferation markers, and/or metabolites. In general, these different classes of molecules can require different assay methods. For analysing functional cytoplasmic proteins, CyTOF is the most mature tool[11], although SCBCs have emerged with similar and complementary capabilities[10]. ScWesterns represent the youngest technology, even though the origins of this technique can be traced back to single-cell gel electrophoresis assays, known as comet assays[88,89]. ScWesterns are also perhaps the method that is most closely aligned to standard biology practice. Each of these tools has advantages and limitations. For analysing secreted proteins, SCBCs have the unique capability of capturing large panels (>40) of proteins secreted from viable cells[8]. Microengraving tools capture only a few secreted proteins, but permit kinetic studies of protein secretion from individual cells[90]. For both SCBCs and microengraving, cells that exhibit unique or desirable protein signatures may be further analysed[30]. CyTOF can capture large panels of 'secrete-able' proteins, but protein

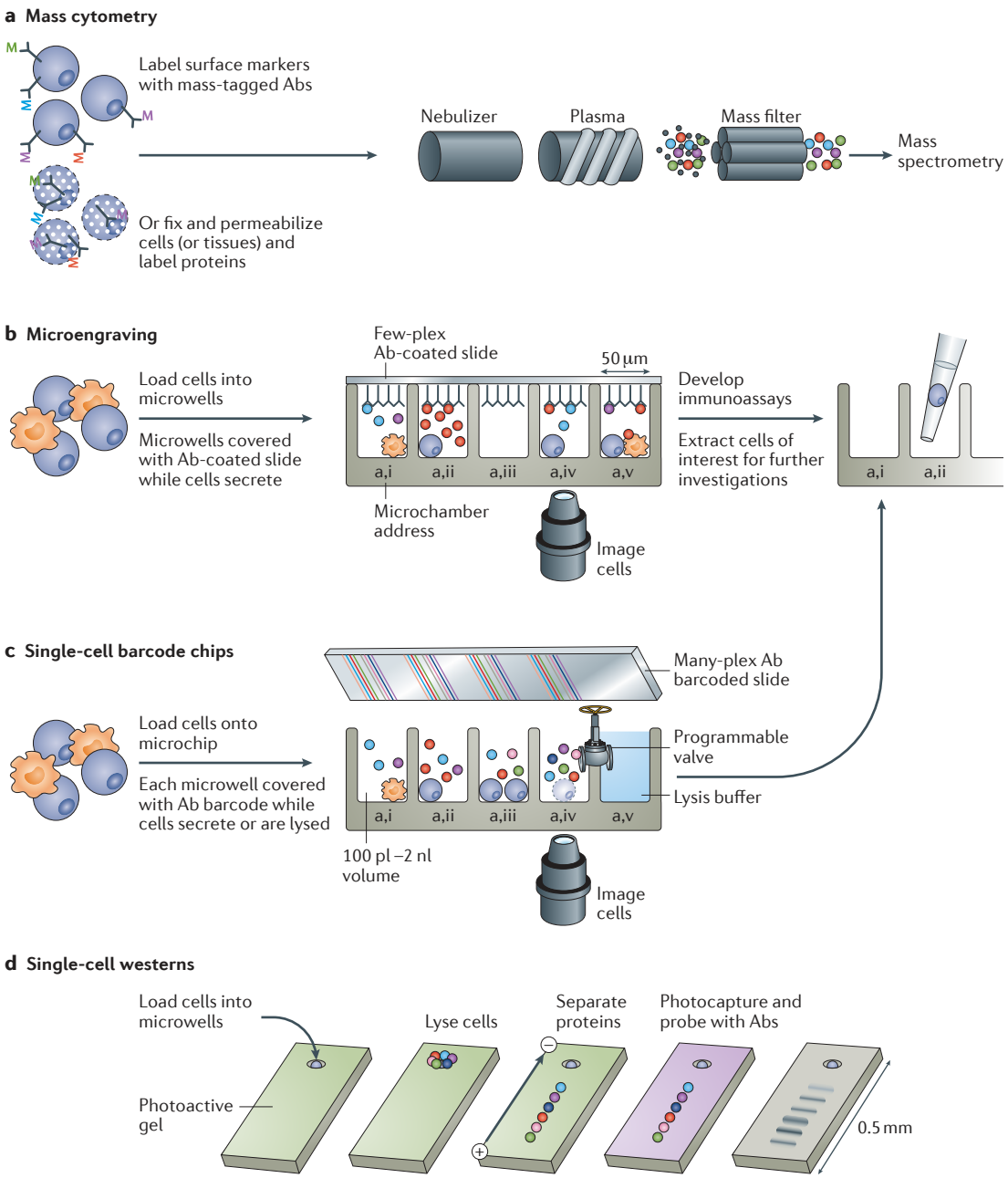


Figure 2 | **Emerging single-cell proteomics methods. a** | Mass cytometry uses antibodies (Abs), encoded with transition metal-containing mass tags, to label proteins of interest. Cells are fixed and permeabilized to permit antibody staining of cytoplasmic proteins. Single cells are entrained into vapour and atomized. A mass filter separates the transition metal atoms, which are then mass analysed. The abundance and identities of the transition metal atoms are traced back to the antibody staining reagents. **b** | The microengraving technique utilizes a microchip with many thousands of microwells, into which none to a few cells of interest are loaded. An antibody-coated coverslide (few-plex Ab-coated slide) is placed over the microchip to capture a few specific secreted proteins. Microchip addresses are correlated with regions on the coverslide and with microscopy images to associate a given cell with a given secretion profile. Captured proteins are detected using fluorescent secondary antibodies, with different proteins identified using different fluorophores. The coverslide can be replaced during the time course of an experiment to capture single-cell secretion kinetics. Cells of interest may be removed for further analysis. **c** | Single-cell barcode chips contain up to a few thousand microchambers, into which none to a few cells are loaded. An antibody-barcoded glass slide (many-plex Ab barcoded slide) is patterned so that each microchamber contains a complete, miniaturized antibody array onto which many cytoplasmic or membrane proteins are captured following their secretion or release upon cell lysis.. Protein assays are developed using fluorescently labelled secondary antibodies, with different proteins identified according to the spatial location of the immunoassay within the barcode. If cells are not lysed (only secreted proteins detected), then the cells remain viable and may be further investigated. **d** | Single-cell westerns are miniaturized variants of traditional western blotting methods, with ~1,000 single cells analysed per microchip.

Mass cytometry
A single cell proteomics method based on traditional flow cytometry methods but uses mass labels and mass spectrometry for protein analysis.

Microengraving
A microfluidics single-cell proteomics method.

Single-cell barcode chips
(SCBCs). A single-cell proteomics method

Single-cell western blottings
(scWesterns). A microchip-based method for carrying out western blotting assays on single cells.

Table 1 | **Characteristics and capabilities of single-cell proteomics methods**

Method	Protein detection method	Comments	Refs
Fluorescence activated cell sorting (FACS)	Staining with fluorophore-labelled antibodies	• Standard for cell sorting based on membrane protein cell surface markers • High-throughput tool with excellent statistics • Mature technique • Multiplexing is colourimetric • Typically requires large sample sizes • Sorted and analysed cells are viable for subsequent analysis • Commercial product (many vendors)	4,7
Mass cytometry (CyTOF)	Staining fixed cells with mass-tag labelled antibodies	• Good for cytoplasmic proteins • Excellent statistics • Demonstrated as a drug screening tool • >30 proteins assayed per cell • Multiplexing is via mass spectrometry • Applicable to fixed-tissue analysis • Commercial product (Fluidigm)	9,11,35, 41,104
Single-cell barcode chips (SCBCs)	Spatially encoded antibody array for fluorescent immunoassays of secreted proteins or analytes released from lysed cells	• Permits absolute quantitation • Suitable for small (100–1,000 cells) biospecimen sizes • Demonstration of >40 proteins assayed per cell • Secreted proteins detected from viable cells • Some designs integrate cell lysis to permit cytoplasmic protein assays and integrated protein and metabolite assays • Analysis of cell–cell interactions • Cost-effective • Multiplexing is via spatially encoded arrays • Commercial service (Isoplexis)	8,10,23, 89,91
Microengraving	Fluorescent immunoassays of secreted proteins	• Small numbers of secreted proteins • >10,000 single cells assayed in parallel • Cost-effective • Permits kinetic studies of protein secretion • Recovery of analysed cells for further analysis • Analysis of cell–cell interactions • Suitable for small (100–1,000 cells) biospecimen sizes • Multiplexing is colourimetric	31,76, 85,99
Single-cell western blotting (scWestern)	Miniaturized, automated western blotting on a microchip	• Suitable for small (100–1,000 cells) biospecimen size • 1,000 cells assayed per microchip • Multiplexing to ~12 proteins demonstrated • Permits cytoplasmic proteins from lysed cells • Reasonably fast (4 hours) • Provides protein ladder reference • Relative quantitation	77

secretion must be blocked and the cells fixed before analysis. Consequently, the detected proteins are not actually secreted and the cells cannot be further analysed[9]. For the microfluidics tools, the cells can be imaged *in situ* and factors such as cell motility or morphology can be correlated with the secretion of specific proteins[91]. The microfluidics tools also permit assays on discrete numbers of cells[24,40,92]. As CyTOF utilizes antibody staining of fixed cells, staining can be performed within fixed tissues, thus permitting CyTOF to be used as a powerful variant of immunohistochemical staining[35]. Each of these tools requires significant user skill, although that requirement is likely to diminish as the platforms mature.

Similar to transcriptomic methods, quantitative assessment of single-cell protein levels is an increasingly important issue. All single-cell proteomics methods utilize antibodies as the dominant detection technology. A recent publication provided a protocol for establishing a clear, quantitative metric for antibody performance[93] and raised serious questions regarding whether a given antibody even detects its intended target. Of the 1,124 antibodies tested, only 452 recognized their target in HEK293 cell lysates[88]. Given that large caveat, the use of antibodies for staining (as with flow cytometry or CyTOF methods) is markedly different from their use in western blotting or fluorescent sandwich immunoassays (SCBCs and microengraving), with each affording different quality checks. scWesterns, similar to standard immunoprecipitation western blotting methods, provide two separate measurements of each protein: the mass ladder (albeit of lower resolution than is typical for bulk western blotting assays) plus a primary detection antibody[94]. However, absolute quantification and absolute assessments of experimental uncertainty can be challenging. For multiplex fluorescent sandwich immunoassays, each individual protein assay provides two separate measurements per cell (as two antibodies per protein are used). Each individual assay can also be compared against every other assay in the panel for crossreactivity[95], and each assay can be calibrated against

solutions spiked with recombinant standards[10,95], thus providing assay readouts in terms of copy numbers per cell. However, on a cautionary note, recombinant standards may not be commercially available or may be modified from the corresponding protein produced within the cells. SCBC platforms have an additional quality check in that individual protein levels can be assayed multiple times from the same single cell[10], thus providing a metric for experimental accuracy. Single-cell methods that rely on antibody staining of cells are the most challenging to quantitate, although experiments on FACS-sorted cells can provide validation that the antibodies used for staining surface markers are effective, thus providing a level of quantification regarding the cell fractions that are positive or negative for specific markers.

Applications of single-cell analysis

With the development of tools that can analyse larger numbers of single cells with an increasing depth of analysis, a central emergent theme is that cellular biology is highly heterogeneous at virtually all molecular levels beyond the genome. Some of this heterogeneity is intrinsic to the nature of single cells (BOX 1), although some of it is reflective of genetic or epigenetic influences[96,97]. In many cases, it is becoming apparent that such heterogeneity is not arbitrary, and may be mined to yield new biological information. A second emergent theme is that a few cells can bias a population average[98,99].

Single-cell genomic or transcriptomic analysis can permit lineage tracing of rare cell types (see below), which can provide insight into the origin (for example, primary tumour or metastatic site) of CTCs, or into the use of CTCs as a liquid biopsy that reflects the originating lesion[49,59,100,101]. A second application, pioneered by Quake and colleagues, has been to provide a deep, molecular view of healthy[16] or diseased[102] tissue development via lineage tracing at the transcript and protein level. Related work has focused on identifying how multiple genetic defects associated with a single gene, but non-uniformly distributed throughout the tumour, influence tumour development and drug response[50].

Advances in single-cell proteomics have largely exploited the ability to interrogate combinations of secreted (or secrete-able) cytokines, chemokines and cytotoxic granules from highly defined cells of the haematopoietic lineage[103,104]. These advances have permitted comparisons of the importance of immune cell function versus immune cell abundance[7,9,90,95] (see below), and also revealed deeper insight into the haematopoietic lineage. Such studies are being applied in cancer immunotherapies[25,53,99,105], as discussed later. Single-cell proteomics has also provided detailed characterizations of the structure of phosphoprotein signalling pathways[10,11] (see below and BOX 1).

An emerging frontier is the use of microfluidics platforms that permit highly customized assays designed to correlate weak perturbations to single cells with changes in the transcriptome or proteome[38]. Two examples are studies that correlated cell motility with proteomic[87] or transcriptome analysis[28]. A third example is an analysis of how specific cancer cells respond to targeted inhibitors as the physical environment is altered from normoxia to hypoxia[27]. Other examples include studies of cellular responses to engineered molecular stimulations (that is, periodic versus continuous)[106,107], or studies designed to interrogate how one cell is influenced by another[108], including how that influence depends on cell–cell separation distance[24,40]. Such studies are enabled by the standardization of the relevant single-cell assay biochemistries, and are limited only by the imagination of the researchers. Together, such research represents tremendously powerful approaches for decoding how genetic and epigenetic influences (such as drugs) are processed by living organisms.

Lineage tracing of cellular phenotypes. Single-cell proteomics and transcriptomics can be used to understand the origins of cellular heterogeneity, as demonstrated by Dalerba and co-workers in colon cancer[102] (FIG. 3). It was found that the transcriptional diversity of a human tumour could be largely explained by *in vivo* multilineage differentiation[102]. These findings are consistent with additional models[109] and mechanistic[110] investigations that demonstrate the ability of cancer cell differentiation (and de-differentiation) to maintain a phenotypic equilibrium within certain tumours. The study by Dalerba and colleagues[102] was limited by the numbers of transcripts per cell, and the numbers of single cells, that could be analysed a few years ago. The recent advent of high-throughput, single-cell global transcriptome analysis and exome sequencing should allow for such lineage tracing studies to dive significantly deeper into a host of developmental biology problems with relevance to both healthy and diseased states.

Understanding cellular functionality. Immune cells of the myeloid lineage are often considered the first responders of host defence against bacterial infection, whereas haematopoietic stem and progenitor cells (HSPCs) are thought to respond in a delayed manner to ensure the sufficient production of myeloid cells, which are consumed during an infection[111–113]. This response of HSPCs was considered to be a passive response to the depletion of downstream immune cells. However, recent evidence suggests that HSPCs may participate directly by sensing systemically elevated cytokines as well as bacterial and viral components through cytokine receptors and Toll-like receptors, respectively[114,115].

Single-cell functional proteomics (12-plex SCBC assays), combined with flow cytometry cell sorting and genetically engineered mouse models, indicated that short-term haematopoietic stem cells (HSCs) and multipotent progenitor cells also have the capacity to respond to bacterial components via the Toll-like receptor–nuclear factor-κB axis[103]. In fact, HSPCs were significantly more potent cytokine producers in terms of speed, breadth and especially quantity than the conventional cytokine producers of the immune system, such as myeloid cells and lymphocytes[95]. Clustering of data from HSC SCBC assays revealed four functional subsets of LKS HSCs (defined as Lineage$^-$Sca1$^+$cKit$^+$); these cells secreted a set of lymphoid cytokines, a set

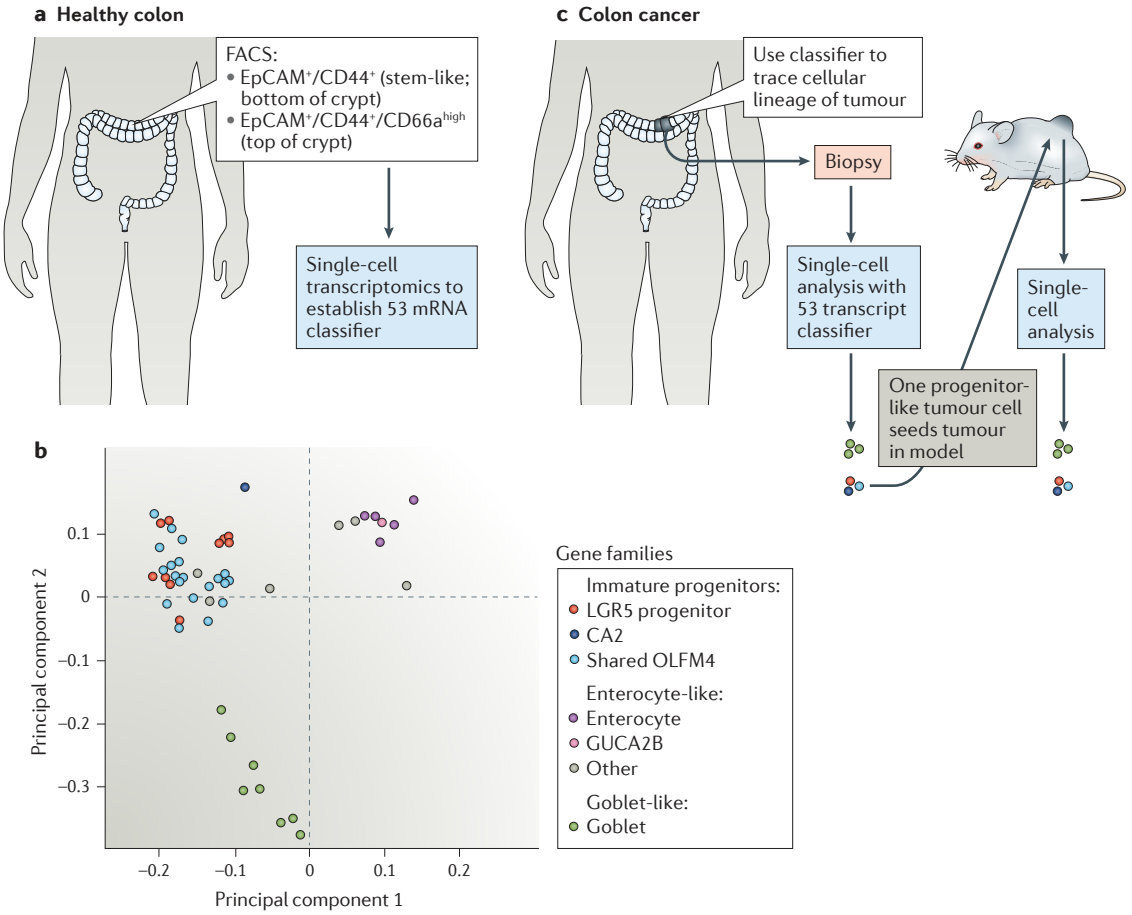


Figure 3 | **Single-cell analysis traces the lineage of a colon cancer.** The work flow proceeds from the left. **a** | A biopsy of a healthy colon is analysed using fluorescent-activated cell sorting (FACS) to separate cells extracted from the crypt-like structures of the colon epithelium. The bottom regions of the crypts are enriched in stem cell-like populations, with those cells identified as epithelial cell adhesion molecule (EpCAM)$^+$/CD44$^+$. More differentiated enterocyte and goblet cells are found near the top of the crypts and are defined as EpCAM$^+$/CD44$^-$/CD66a^{high}. **b** | Single cell, multiplex transcriptomics is used to develop a 53 gene expression classifier. Principal component analysis of the single-cell data resolves the major cellular subpopulations. The genes that define these subpopulations are plotted with respect to how they are represented within the two dominant principal components. The plot reveals how the classifier resolves immature progenitors (top left of graph), enterocyte-like cells (top right), and goblet-like cells (bottom left). Classifiers of these populations, also identified from hierarchical clustering of the single cell transcriptome data, provide the colour coding for each mRNA on the plot. **c** | Once established, the classifier can be used to analyse cells collected from a patient colon cancer tumour. In this case, the tumour cells (shaded in grey) are largely goblet-like and immature progenitors. A single immature progenitor tumour cell is sorted from the patient tumour using FACS, and implanted into a mouse model to grow a monoclonal tumour. Analysis of that tumour reveals a cellular composition reminiscent of the original patient tumour, implying that the tumour cellular heterogeneity can originate from expansion and lineage differentiation of a single progenitor-like cell. CA2, carbonic anhydrase 2; GUCA2B, guanylate cyclase activator 2B; LGR5, leucine-rich repeat containing G protein-coupled receptor; OLFM4, olfactomedin 4. Figure part **b** is adapted with permission from REF. 102.

of myeloid-associated cytokines, produced all proteins or were completely silent[95]. The overall findings indicated that short-term HSCs and multipotent progenitor cells, although rare cells, can aggressively translate danger signals arising from an infection into the vigorous production of cytokine signals that allow them to directly self-regulate stress-induced haematopoiesis. These findings have multiple implications, with one possibility being related to patients who have undergone lymphodepletion regimens as part of a therapeutic procedure. As the single-cell functional proteomics assays are non-destructive to the cells, a logical next step in this type of work would be to analyse those functional subsets at the transcriptome level to identify whether there are specific cell surface markers that can be used to further differentiate these HSPC functional subsets.

High-throughput drug screening. Although single-cell methods can provide valuable information, most are limited to analysing only one to a few samples at a time. However, cellular barcoding techniques are evolving to remove this limitation[18,116]. For mass cytometry, the basic

idea is that cells are separated into a multiwell plate and barcoded with a unique combination of mass signatures that identify a given cell with its well location and the experimental conditions (that is, a specific dose of a specific drug) applied to that location. The cells are then analysed simultaneously to enable the parallel capture of many experimental conditions. For example, Bodenmiller and co-workers[116] used 7 mass-labelled barcodes to provide up to 2^7 barcoding capacity (128 possible addresses). This method was applied to a 96-well plate format to explore the kinetic and/or dosing influences of 27 inhibitors on 14 distinct peripheral blood mononuclear cell phenotypes (defined by 10 cell surface markers) via the monitoring of 14 phosphorylation sites per cell. From these data, half-maximal inhibitory concentration (IC_{50}) values and percentage inhibition of the phosphorylation levels for all phosphorylated sites were extracted[116].

As described above, single-cell, multiplex phosphoproteomic assays yield both the levels of the assayed proteins and the protein–protein correlations. Thus, a major advantage and distinguishing feature of this high-throughput screening approach is that it permits an analysis of how both on-target and off-target drug interactions influence the signalling networks, rather than just the relevant protein levels. A major challenge going forward will be to expand the multiplexing of these types of assays to capture more complete pictures of the phosphoprotein signalling networks, as well as additional networks associated with cellular proliferation, apoptosis and metabolism.

Single-cell analysis in oncology

The single-cell analytical methods discussed in this Review are being applied towards addressing several fundamental biomedical problems, particularly in cancer biology and clinical oncology. Below, we discuss how single-cell analysis is being applied to two key areas of cancer research and drug discovery: cancer immunotherapy and tumour heterogeneity.

Cancer immunotherapy. The prototype model for our understanding of cellular differentiation and diversification in humans is the haematopoietic system. In fact, this knowledge has provided a scientific cornerstone underlying the recent and remarkable advances in cancer immunotherapy[117,118]. Single-cell technologies have emerged as a critical set of tools for advancing this knowledge, often in a dramatic manner.

For cancer immunotherapies, single-cell analytical tools are beginning to provide critical guidance across multiple levels of biological information. Whether the immunotherapy is based on dendritic cell vaccines[119], adoptive cell transfer[120] or checkpoint inhibitors[121–123], or some combination thereof, the primary tumour cell killers are T cells. Some of the most important biomarkers are the kinetic persistence and functional behaviours of specific antitumour T cell phenotypes across the course of a given patient's therapeutic regimen. For cell-based therapies, the importance of designing clinical protocols that account for T cell differentiation has emerged as a key consideration[124]. In addition, a few years ago it was suggested that patient-specific mutant epitopes[125] (called neoantigens) may be a potentially important factor for understanding, or perhaps controlling, the antitumour specificity of an immunotherapy; this theory has recently been confirmed[105,126–130]. Closely associated factors are the TCRα and TCRβ chain sequences that recognize the specific expressed neoantigens with high avidity[53].

Much of this work has involved highly multiplex flow and mass cytometry methods[7,9,131,132] and associated reagent development[133–138]. These techniques enable the phenotypic characterization of immune cells and also the study of intracellular signalling pathways. With the realization that T cell responses to cancer can lead to unprecedented levels of durable tumour responses in several types of cancer (for example, melanoma, lung, bladder, lymphoma, leukaemia)[139], there is a need for further characterization of such responses to facilitate the increased refinement of therapeutic approaches and continued improvements in patient care.

To understand immune responses to cancer, it is of significant interest to develop approaches that can match the TCR genes with their specific (or cognate) antigen, which usually result from nonsynonymous somatic mutations specific for each cancer[105,133,135–138,140,141]. As each T cell has two TCR chains, it is important that they are defined from individual T cells to allow their correct pairing. Single-cell analysis platforms coupled with DNA sequencing for TCR chains and paired neoantigens have the potential to revolutionize our knowledge about this critical interaction, guiding the success of cancer immunotherapy strategies[17,53]. With the increased knowledge, it is easy to envision that in the near future the definition of TCR chains that specifically recognize neoantigens in cancers may be translated into truly personalized cancer immunotherapy approaches for patients.

Once the recognition elements of T cells are fully defined, a next question is which T cell subsets are empowered to fight cancer. Although these T cell subsets are each governed by specific transcription factors and can be identified by a series of surface molecules[142], a specific subset can also exhibit a broad range of functional phenotypes, ranging from antitumour to immunomodulatory[99]. Single-cell assays provide an unparalleled quantitative assessment of the different T cell subtypes, their progenitors and their functional capabilities[9,12,37,99,103]. These assays are being applied to the characterization of T cell responses to cancer induced by several immunotherapy approaches[99,103]. Such methods are helping to define how patients respond to or resist immunotherapy approaches, such as checkpoint blockade therapy, and may help guide the next generation of combination therapy studies that will be designed based on understanding what is lacking in patients whose immune systems do not respond to these therapies.

Advances in cell therapy manufacturing for adoptive cell transfer approaches, in which a large army of T cells are manufactured in the laboratory and re-infused back into patients, are being supported by new biotechnology approaches designed to guide higher level T cell characterizations[25,53,105,130]. The understanding of TCR specificity and the generation of chimeric antigen receptors to genetically redirect T cell specificity to cancer enables the manufacture of autologous cell therapies[143]. By applying

Neoantigens
Small peptide fragments that contain a genetic mutation. These fragments may be recognized by T cells during an antitumour immune response.

highly multiplexed single-cell analyses, the different T cell subsets can be surveyed before and after infusion of these cell therapies to patients to define which approaches improve their long-term functionality to attack cancer. It has become clear that less mature cells that have long-term repopulation ability (for example, naive, T stem cell and long-term memory cells) are preferred in these adoptive cell transfer approaches[124]; more mature T effector cells have short-term functionality and cancer may regrow after their infusion[99,144].

The next wave of advances in cancer immunotherapy are likely to rely on the characterization of large numbers of single immune cells at the DNA, RNA and protein levels to deconvolute the complexity of immune responses to cancer and guide further therapeutic strategies. Lower order analyses fail to provide the necessary knowledge to understand immune responses to cancer and cannot explain the heterogeneity in patient responses.

Understanding tumour heterogeneity. Intratumoural heterogeneity is increasingly being recognized as a central hallmark of human cancer[145], and encompasses three main types of variability: variations in mutational patterns among tumours of the same histological type; variations in histological patterns within a tumour; and intratumoural mutational polyclonality, that is, variations in the mutational complement within individual cells of a tumour[146]. In addition to mutational polyclonality, single cells within a tumour will intrinsically vary in the activity of their signalling[10,11] and metabolic[23] networks, thus influencing the biological properties and therapeutic vulnerabilities of distinct tumour cell subpopulations. Due to a lack of suitable genetic tools, the impact of intratumoural mutational polyclonality and heterogeneity of signalling and biochemical networks on treatment and resistance are not currently well understood.

Tumours develop into a complex heterogeneous tumour mass, primarily through the intertwined forces of spontaneous somatic mutation coupled to clonal sequential selection for aggressive subclones[145,147–149]. As tumours progress, new mutations are produced with an ever increasing frequency, accelerating the extent of intratumoural mutational polyclonality and confounding treatment strategies[146,147]. Intratumoural mutational polyclonality is enhanced in cancers that are associated with a causal environmental insult that directly damages DNA[150], as well as by the progressive loss of key tumour suppressor proteins, and mutations in genes that sense and repair damaged DNA[151].

The local microenvironment also provides a critical non-genetic force. Autocrine and paracrine interactions among inflammatory, stromal, endothelial and tumour cells are just a few factors that can influence the process of selection, and may yield cells bearing different mutations within different parts of a tumour or its distant metastases[152,153]. Treatments also provide a source of non-genetic heterogeneity, expanding or collapsing tumour cell subpopulations depending on the treatment[145,146,149].

Exome and genome-wide surveys have provided an atlas of driver mutations and a compelling road map for guiding the implementation of precision and personalized cancer medicine. However, intratumoural heterogeneity presents a serious challenge to this paradigm[154]. Fortunately, single-cell technologies are poised to address this challenge.

Glioblastoma, the most common and lethal form of primary brain cancer, provides an illustrative example. Glioblastoma was one of the first cancers sequenced by The Cancer Genome Atlas[155,156]. In that survey, 57% of tumours contained epidermal growth factor (*EGFR*) amplification and/or gain-of-function mutations, including *EGFR* variant III (*EGFRvIII*)[157,158]. *EGFRvIII* is oncogenic in mouse models when introduced in association with cyclin-dependent kinase inhibitor 2 (CDKN2A) loss[159]; a combination that commonly occurs in patients[146]. However, EGFRvIII protein expression varies dramatically among cells within a glioblastoma, and single-cell DNA sequencing[50], RNA sequencing[17], as well as bulk analysis of DNA and RNA extracted from different regions of a tumour[160,161] demonstrate considerable DNA, transcript and protein heterogeneity, including of EGFRvIII. Importantly, recent research suggests that the widespread variability of gene, transcript and protein levels of EGFRvIII within individual cells of a glioblastoma may contribute to the resistance to EGFR-targeted therapies that is currently seen in the clinic[110]. In addition, single-cell barcode proteomics and metabolomics assays point to considerable variability in the signalling and metabolic networks of individual glioblastoma cells[23] within an *EGFRvIII*$^+$ tumour, potentially shedding new light on mechanisms of resistance (either pre-existing and/or adaptive) to targeted therapies that could be used to guide more effective combination treatments.

Challenges, limitations and outlook

The emergence of single-cell omics tools over the past 5 years has happened at a lightning pace, and the potential for their use in the discovery and development of broad classes of therapies and therapeutic strategies is high. The resultant datasets do not simply provide deeper views of biology that are already measured using existing methods, but also provide a fundamentally different view that is not masked by the intrinsic heterogeneity of a cell population. However, the novelty of these single-cell techniques also implies various limitations. For example, most methods discussed in this Review have just recently emerged from academic laboratories and therefore require significant skill sets and cross-disciplinary infrastructure that may be new to those in the drug discovery and development community. As cases in point, REFS. 102,103,116 (discussed above) have, on average, 13 authors representing 5 different departments or institutions, which demonstrate the need to effectively integrate experts in technology, biology and computational analysis. A second caveat is that algorithms for the in-depth analysis of single-cell data are even less mature than the experimental platforms, and effective visualization and interpretation of what are increasingly large datasets remain challenging, with techniques that vary across research groups. However, as methods mature, the experimental protocols, the reagents and the computational analysis routines will become more standardized. This standardization has,

of course, largely occurred for multicolour flow cytometry, and it is beginning to happen for CyTOF and single-cell RNA-seq, but even these methods are rapidly evolving and all involve dedicated user facilities.

Much of modern biological practice is designed around extracting correlations and associated statistical trends from biological systems that are intrinsically heterogeneous and therefore noisy. The promise of single-cell biology is to resolve and make sense of this confounding heterogeneity. Several studies highlighted in this Review provide hints of the resultant clarity that can be achieved. It is likely that, as the tools of the field increase in terms of quantification, throughput and ease of use, the impact will be to fundamentally change the practice of biology, as well as the associated applied sciences, including drug discovery and development.

1. Sakmann, B. & Neher, E. Patch clamp techniques for studying ionic channels in excitable membranes. *Annu. Rev. Physiol.* **46**, 455–472 (1984).
2. Amann, R. & Fuch, B. M. Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. *Nat. Rev. Microbiol.* **6**, 339–348 (2008).
3. Langer-Safer, P. R., Levine, M. & Ward, D. C. Immunological method for mapping genes on *Drosophila* polytene chromosomes. *Proc. Natl Acad. Sci. USA* **79**, 4381–4385 (1982).
4. Herzenberg, L. A. *et al.* The history of the fluorescence activated cell sorter and flow cytometry: a view from Stanford. *Clin. Chem.* **48**, 1819–1827 (2002).
5. Herzenberg, L. A., Julius, M. H. & Masuda, T. Demonstration that antigen-binding cells are precursors of antibody-producing cells after purification with a fluorescence-activated cell sorter. *Proc. Natl Acad. Sci. USA* **69**, 1934–1938 (1972).
6. Czerkinsky, C., Nilsson, L., Nygren, H., Ouchterlony, O. & Tarkowski, A. A solid-phase enzyme-linked immunospot (ELISPOT) assay for enumeration of specific antibody-secreting cells. *Immunol. Methods* **65**, 109–121 (1983).
7. Perfetto, S. P., Chattopadhyay, P. K. & Roederer, M. Seventeen-colour flow cytometry: unravelling the immune system. *Nat. Rev. Immunol.* **4**, 648–655 (2004).
An illustration of the state of the art of multiplex flow cytometry.
8. Lu, Y. *et al.* Highly multiplexed profiling of single-cell effector functions reveals deep functional heterogeneity in response to pathogenic ligands. *Proc. Natl Acad. Sci. USA* **112**, E607–E615 (2015).
9. Bendall, S. C. *et al.* Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687–696 (2011).
10. Shi, Q. *et al.* Single cell proteomic chip for profiling intracellular signaling pathways in single tumor cells. *Proc. Natl Acad. Sci. USA* **109**, 419–425 (2012).
An illustration of quantitative and multiplex single-cell proteomics.
11. Irish, J. M. *et al.* Single cell profiling of potentiated phospho-protein networks in cancer cells. *Cell* **118**, 217–228 (2004).
12. Navin, N. *et al.* Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90–94 (2011).
13. de Bourcy, C. F. *et al.* A quantitative compariso of single-cell whole genome amplification methods. *PLoS ONE* **9**, e105585 (2014).
14. Wang, J., Fan, H. C., Behr, B. & Quake, S. R. Genome-wide single-cell analysis of recombination activity and *de novo* mutation rates in human sperm. *Cell* **150**, 402–412 (2012).
15. Tay, S. *et al.* Single-cell NF-κB dynamics reveal digital activation and analogue information processing. *Nature* **466**, 267–271 (2010).
16. Treutlein, B. *et al.* Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* **509**, 371–375 (2014).
17. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1400 (2014).
18. Fan, H. C., Fu, G. K. & Fodor, S. P. A. Combinatorial labeling of single cells for gene expression cytometry. *Science* **347**, 6222 (2015).
Describes the CytoSeq method.
19. Han, L. *et al.* Co-detection and sequencing of genes and transcripts from the same single cells facilitated by a microfluidics platform. *Sci. Rep.* **4**, 6485 (2014).
20. Mazumder, A., Tummler, K., Bathe, M. & Samson, L. D. Single-cell analysis of ribonucleotide reductase transcriptional and translational response to DNA damage. *Mol. Cell. Biol.* **33**, 635–642 (2013).
21. Porichis, F. *et al.* High-throughput detection of miRNAs and gene-specific mRNA at the single-cell level by flow cytometry. *Nat. Commun.* **5**, 5641 (2014).
22. Stahlberg, A., Thomsen, C., Ruff, D. & Aman, P. Quantitative PCR analysis of DNA, RNAs and proteins in the same single cell. *Clin. Chem.* **58**, 1682–1691 (2012).
23. Xue, M. *et al.* Chemical methods for the simultaneous quantitation of metabolites and proteins from single cells. *J. Am. Chem. Soc.* **137**, 4066–4069 (2015).
24. Wang, J. *et al.* Quantitating cell–cell interaction functions, with applications to glioblastoma multiforme cancer cells. *Nano Lett.* **12**, 6101–6106 (2012).
25. Liadi, I. *et al.* Individual motile $CD4^+$ T cells can participate in efficient multikilling through conjugation to multiple tumor cells. *Cancer Immunol. Res.* **3**, 473–482 (2015).
26. Elitas, M., Brower, K., Lu, Y., Chen, J. J. & Fan, R. A microchip platform for interrogating tumormacrophage paracrine signaling at the single-cell level. *Lab Chip* **14**, 3582–3588 (2014).
27. Wei, W. *et al.* Hypoxia induces a phase transition within a kinase signaling network in cancer cells. *Proc. Natl Acad. Sci. USA* **110**, e1352–e1360 (2013).
28. Mehling, M., Frank, T., Albayrak, C. & Tay, S. Real-time tracking, retrieval and gene expression analysis of migrating human T cells. *Lab Chip* **15**, 1276–1283 (2015).
29. Brouzes, E. *et al.* Droplet microfluidic technology for single-cell high-throughput screening. *Proc. Natl Acad. Sci. USA* **106**, 14195–14200 (2009).
30. Sendra, V. G., Lie, A., Romain, G., Agarwal, S. K. & Varadarajan, N. Detection and isolation of auto-reactive human antibodies from primary B cells. *Methods* **64**, 153–159 (2013).
Describes the microengraving technique used to identify cells with desirable proteomic signatures and then separate those cells for further analysis.
31. Love, J. C., Ronan, J. L., Grotenbreg, G. M., Van der Veen, A. G. & Ploegh, H. L. A microengraving method for rapid selection of single cells producing antigen specific antibodies. *Nat. Biotechnol.* **24**, 703–707 (2006).
32. Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M. & Long, C. Single-cell *in situ* RNA profiling by sequential hybridization. *Nat. Methods* **11**, 360–361 (2014).
33. Yang, B. *et al.* Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell* **158**, 945–958 (2014).
34. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, 6233 (2015).
Reports a single-cell transcriptomics method applied to cells in a bulk environment.
35. Angelo, M. *et al.* Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* **20**, 436–442 (2014).
36. Shin, Y. S. *et al.* Chemistries for patterning robust DNA microbarcodes enable multiplex assays of cytoplasm proteins from single cancer cells. *ChemPhysChem* **11**, 3063 (2010).
37. Amir, E. D. *et al.* viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nat. Biotechnol.* **31**, 545–548 (2013).
38. Shin, Y. S. *et al.* Protein signaling networks from single cell fluctuations and information theory profiling. *Biophys. J.* **100**, 2378–2386 (2011).
39. Jaitin, D. A. *et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* **343**, 776–779 (2014).
40. Kravchenko-Balasha, N., Wang, J., Remacle, F., Levine, R. D. & Heath, J. R. Glioblastoma cellular architectures are predicted through the characterization of two-cell interactions. *Proc. Natl Acad. Sci. USA* **111**, 6521–6526 (2014).
41. Bruggner, R. V., Bodenmiller, B., Dill, D. L., Tibshirani, R. J. & Nolan, G. P. Automated identification of stratifying signatures in cellular subpopulations. *Proc. Natl Acad. Sci. USA* **111**, E2770–E2777 (2014).
42. Shapiro, E., Biezuner, T. & Linnarsson, S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat. Rev. Genet.* **14**, 618–630 (2013).
43. Blainey, P. C. & Quake, S. R. Dissecting genomic diversity, one cell at a time. *Nat. Methods* **11**, 19–21 (2014).
44. Zhang, L. *et al.* Whole genome amplification from a single cell: implications for genetic analysis. *Proc. Natl Acad. Sci. USA* **89**, 5847–5851 (1992).
45. Acinas, S. G., Sarma-Rupavtarm, R., Klepac-Ceraj, V. & Polz, M. F. PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Appl. Environ. Microbiol.* **71**, 8966–8969 (2005).
46. Dean, F. B., Nelson, J. R., Giesler, T. L. & Lasken, R. S. Rapid amplification of plasmid and phage DNA using Phi29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* **11**, 1095–1099 (2001).
47. Dago, A. E. *et al.* Rapid phenotypic and genomic change in response to therapeutic pressure in prostate cancer inferred by high content analysis of single circulating tumor cells. *PLoS ONE* **9**, e101777 (2014).
48. Zong, C., Lu, S., Chapman, A. R. & Xie, X. S. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* **338**, 1622–1626 (2012).
Reports the MALBAC technique.
49. Ni, X. *et al.* Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. *Proc. Natl Acad. Sci. USA* **110**, 21083–21088 (2013).
50. Francis, J. M. *et al.* EGFR variant heterogeneity in glioblastoma resolved through single-nucleus sequencing. *Cancer Discov.* **4**, 956–971 (2014).
51. Dash, P. *et al.* Paired analysis of TCRα and TCRβ chains at the single-cell level in mice. *J. Clin. Invest.* **121**, 288–295 (2011).
52. Kim, S.-M. *et al.* Analysis of the paired TCR α- and β-chains of single human T cells. *PLoS ONE* **7**, e37338 (2013).
53. Han, A., Glanville, J., Hansmann, L. & Davis, M. M. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat. Biotechnol.* **32**, 684–692 (2014).
Illustrates an important method that is relevant to various classes of immunotherapies.
54. Ng, S. B. *et al.* Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272–276 (2009).
55. Mamanova, L. *et al.* Target-enrichment strategies for next-generation sequencing. *Nat. Methods* **7**, 111–118 (2010).
56. Hodges, E. *et al.* Genome-wide *in situ* exon capture for selective resequencing. *Nat. Genet.* **39**, 1522–1527 (2007).
57. Hou, Y. *et al.* Single-cell exome sequencing and monoclonal evolution of a *JAK2*-negative myeloproliferative neoplasm. *Cell* **148**, 873–885 (2012).
58. Xu, X. *et al.* Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* **148**, 886–895 (2012).

59. Lohr, J. G. *et al.* Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. *Nat. Biotechnol.* **32**, 479–484 (2014).
60. Eberwine, J. *et al.* Analysis of gene expression in single live neurons. *Proc. Natl Acad. Sci. USA* **89**, 3010–3014 (1992).
61. Tang, F. *et al.* mRNA-seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377–382 (2009).
62. Islam, S. *et al.* Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat. Methods* **11**, 163–166 (2014).
63. Wu, A. R. *et al.* Quantitative assessment of single-cell RNA-sequencing methods. *Nat. Methods* **11**, 41–46 (2014).
64. Shalek, A. K. *et al.* Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* **498**, 236–240 (2013).
65. Pollen, A. A. *et al.* Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* **32**, 1053–1058 (2014).
66. Ting, D. T. *et al.* Single-cell RNA sequencing identifies extracellular matrix gene expression by pancreatic circulating tumor cells. *Cell Rep.* **8**, 1905–1918 (2014).
67. Luo, Y. *et al.* Single-cell transcriptome analyses reveal signals to activate dormant neural stem cells. *Cell* **161**, 1175–1186 (2015).
68. Klein, A. M. *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).
Reports on the InDrop single-cell transcriptomics method.
69. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
70. Kivioja, T. *et al.* Counting absolute numbers of molecules using unique molecular identifiers. *Nat. Methods* **9**, 72–74 (2012).
A description of molecular barcoding for quantitative biology.
71. Grun, D., Kester, L. & van Oudenaarden, A. Validation of noise models for single-cell transcriptomics. *Nat. Methods* **11**, 637–640 (2014).
A description of the limits and possibilities of absolute quantitation from single-cell transcriptome analysis.
72. Polz, M. F. & Cavanaugh, C. M. Bias in template-to-product ratios in multitemplate PCR. *Appl. Environ. Microbiol.* **64**, 3724–3730 (1998).
73. Schloss, P. D., Gevers, D. & Westcott, S. L. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS ONE* **6**, e27310 (2011).
74. Suzuki, M. T. & Giovannoni, S. J. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl. Environ. Microbiol.* **62**, 625–630 (1996).
75. Brennecke, P. *et al.* Accounting for technical noise in single-cell RNA-seq experiments. *Nat. Methods* **10**, 1093–1095 (2013).
76. Fu, G. K., Wilhelmy, J., Stern, D., Fan, H. C. & Fodor, S. P. A. Digital encoding of cellular mRNAs enabling precise and absolute gene expression measurement by single-molecule counting. *Anal. Chem.* **86**, 2867–2870 (2014).
77. Fu, G. K. *et al.* Molecular indexing enables quantitative targeted RNA sequencing and reveals poor efficiencies in standard library preparations. *Proc. Natl Acad. Sci. USA* **111**, 1891–1896 (2014).
78. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
79. Yates, J. R., Ruse, C. I. & Nakorchevsky, A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev. Biomed. Eng.* **11**, 49–79 (2009).
80. Picotti, P. & Aebersold, R. Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat. Methods* **9**, 555–566 (2012).
81. Torres, A. J., Contento, R. L., Gordo, S., Wucherpfennig, K. W. & Love, C. L. Functional single-cell analysis of T-cell activation by supported bilayer-tethered ligands on arrays of nanowells. *Lab Chip* **13**, 90–99 (2013).
82. Hughes, A. J. *et al.* Single-cell western blotting. *Nat. Methods* **11**, 749–755 (2014).
83. Guo, M. T., Rotem, A., Heyman, J. A. & Weitz, D. A. Droplet microfluidics for high-throughput biological assays. *Lab Chip* **12**, 2146–2155 (2012).
84. Huebner, A. *et al.* Static microdroplet arrays: a microfluidic device for droplet trapping, incubation and release for enzymatic and cell-based assays. *Lab Chip* **9**, 692–698 (2009).
85. Kintses, B., van Vliet, L. D., Devenish, S. R. A. & Hollfelder, F. Microfluidic droplets: new integrated workflows for biological experiments. *Curr. Opin. Chem. Biol.* **14**, 548–555 (2010).
86. Mazutis, L. *et al.* Single-cell analysis and sorting using droplet-based microfluidics. *Nat. Protoc.* **8**, 870–891 (2013).
87. Yu, J. *et al.* Microfluidics-based single-cell functional proteomics for fundamental and applied biomedical applications. *Annu. Rev. Anal. Chem.* **7**, 275–295 (2014).
An overview of nanodroplet microfluidics methods for single-cell analysis.
88. McKelvey-Martin, V. J. *et al.* The single cell gel electrophoresis assay (comet assay): a European review. *Mut. Res.* **288**, 47–63 (1993).
89. Weingeist, D. M. *et al.* Single-cell microarray enables high-throughput evaluation of DNA double-strand breaks and DNA repair inhibitors. *Cell Cycle* **12**, 907–915 (2013).
90. Han, Q. *et al.* Polyfunctional responses by human T cells result from sequential release of cytokines. *Proc. Natl Acad. Sci. USA* **109**, 1607–1612 (2012).
91. Lu, Y. *et al.* High-throughput secretomic analysis of single cells to assess functional cellular heterogeneity. *Anal. Chem.* **85**, 2548–2556 (2013).
92. Romain, G. *et al.* Antibody Fc-engineering improves frequency and promotes kinetic boosting of serial killing mediated by NK cells. *Blood* **124**, 3241–3249 (2014).
93. Marcon, E. *et al.* Assessment of a method to characterize antibody selectivity and specificity for use in immunoprecipitation. *Nat. Methods* **12**, 725–731 (2015).
Describes the limitations of antibodies for protein detection assays.
94. Towbin, H., Staehelin, T. & Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl Acad. Sci. USA* **76**, 4350–4354 (1979).
95. Ma, C. *et al.* A clinical microchip for evaluation of single immune cells reveals high functional heterogeneity in phenotypically similar T cells. *Nat. Med.* **17**, 738–744 (2011).
96. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
An excellent illustration of how single cells and bulk cell populations differ.
97. Kaern, M., Elston, T. C., Blake, W. J. & Collins, J. J. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.* **6**, 451–464 (2005).
98. Bengtsson, M., Stahlberg, A., Rorsman, P. & Kubista, M. Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels. *Genome Res.* **15**, 1388–1392 (2005).
99. Ma, C. *et al.* Multifunctional T-cell analyses to study response and progression in adoptive cell transfer immunotherapy. *Cancer Discov.* **3**, 418–429 (2013).
Illustrates the value of single-cell functional proteomics for cancer immunotherapy monitoring in patients.
100. Powell, A. A. *et al.* Single cell profiling of circulating tumor cells: transcriptional heterogeneity and diversity from breast cancer cell lines. *PLoS ONE* **7**, e33788 (2012).
101. Halo, T. L. *et al.* NanoFlares for the detection, isolation, and culture of live tumor cells from human blood. *Proc. Natl Acad. Sci. USA* **111**, 17104–17109 (2014).
102. Dalerba, P. *et al.* Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat. Biotechnol.* **29**, 1120–1127 (2011).
An example of lineage tracing via single-cell transcriptomics.
103. Zhao, J. L. *et al.* Conversion of danger signals into cytokine signals by hematopoietic stem and progenitor cells for regulation of stress-induced hematopoiesis. *Cell Stem Cell* **14**, 1–15 (2014).
104. Lin, L. *et al.* Human natural killer cells licensed by killer Ig receptor genes have an altered cytokine program that modifies $CD4^+$ T cell function. *J. Immunol.* **193**, 940–949 (2014).
105. Van Buuren, M. M., Calis, J. J. & Schumacher, T. N. High sensitivity of cancer exome-based CD8 T cell neo-antigen identification. *Oncoimmunology* **3**, e28836 (2014).
106. Kellogg, R. A., Gómez-Sjöberg, R., Leyrat, A. A. & Tay, S. High-throughput microfluidic single-cell analysis pipeline for studies of signaling dynamics. *Nat. Protoc.* **9**, 1713–1726 (2014).
107. Kellogg, R. A. & Tay, S. Noise facilitates transcriptional control under dynamic inputs. *Cell* **160**, 381–392 (2015).
108. Yamanaka, Y. J. *et al.* Single-cell analysis of the dynamics and functional outcomes of interactions between human natural killer cells and target cells. *Integr. Biol.* **4**, 1175–1184 (2012).
109. Gupta, P. B. *et al.* Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633–644 (2011).
110. Nathanson, D. *et al.* Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant *EGFR* DNA. *Science* **343**, 72–76 (2014).
A mechanism of adaptive resistance to a targeted inhibitor is revealed by single-cell analysis.
111. Jaiswal, S. & Weissman, I. L. Hematopoietic stem and progenitor cells and the inflammatory response. *Ann. NY Acad. Sci.* **1174**, 118–121 (2009).
112. Baldridge, M. T., King, K. Y., Boles, N. C., Weksberg, D. C. & Goodell, M. A. Quiescent haematopoietic stem cells are activated by IFN-γ in response to chronic infection. *Nature* **465**, 793–797 (2010).
113. Baldridge, M. T., King, K. Y. & Goodell, M. A. Inflammatory signals regulate hematopoietic stem cells. *Trends Immunol.* **32**, 57–65 (2011).
114. King, K. Y. & Goodell, M. A. Inflammatory modulation of HSCs: viewing the HSC as a foundation for the immune response. *Nat. Rev. Immunol.* **11**, 685–692 (2011).
115. Nagai, Y. *et al.* Toll-like receptors on hematopoietic progenitor cells stimulate innate immune system replenishment. *Immunity* **24**, 801–812 (2006).
116. Bodenmiller, B. *et al.* Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nat. Biotechnol.* **30**, 858–867 (2012).
An illustration of CyTOF, combined with mass-label barcoding, for single-cell analysis and drug screening.
117. Couzin-Frankel, J. Cancer immunotherapy. *Science* **343**, 1432 (2013).
118. Ledford, H. Cancer treatment: the killer within. *Nature* **508**, 24–26 (2014).
119. Palucka, K. & Banchereau, J. Cancer immunotherapy via dendritic cells. *Nat. Rev. Cancer* **12**, 265–277 (2012).
120. Rosenberg, S. A., Restifo, N. P., Yang, J. C., Morgan, R. A. & Dudley, M. E. Adoptive cell transfer: a clinical path to effective cancer immunotherapy. *Nat. Rev. Cancer* **8**, 299–308 (2008).
121. Leach, D. R., Krummel, M. F. & Allison, J. P. Enhancement of antitumor immunity by CTLA-4 blockade. *Science* **271**, 1734–1736 (1996).
122. Ribas, A. & Tumeh, P. C. The future of cancer therapy: selecting patients who respond to PD-1/L1 blockade. *Clin. Cancer Res.* **20**, 4982–4984 (2014).
123. Iwai, Y. *et al.* Involvement of PD-L1 on tumor cells in the escape from host immune system and tumor immunotherapy by PD-L1 blockade. *Proc. Natl Acad. Sci. USA* **99**, 12293–12297 (2002).
124. Crompton, J. G., Sukumar, M. & Restifo, N. P. Uncoupling T-cell expansion from effector differentiation in cell-based immunotherapy. *Immunol. Rev.* **257**, 264–276 (2014).
125. Segal, N. H. *et al.* Epitope landscape in breast and colorectal cancer. *Cancer Res.* **68**, 889–892 (2008).
126. Coulie, P. G., Van den Eynde, B. J., van der Bruggen, P. & Boon, T. Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. *Nat. Rev. Cancer* **14**, 135–146 (2014).
127. DuPage, M., Mazumdar, C., Schmidt, L. M., Cheung, A. F. & Jacks, T. Expression of tumour-specific antigens underlies cancer immunoediting. *Nature* **482**, 405–409 (2012).
128. Matsushita, H. *et al.* Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* **482**, 400–404 (2012).
129. Robbins, P. F. *et al.* Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells. *Nat. Med.* **19**, 747–752 (2013).
130. van Rooij, N. *et al.* Tumor exome analysis reveals neoantigen-specific T-cell reactivity in an ipilimumab-responsive melanoma. *J. Clin. Oncol.* **31**, e439–e442 (2013).

131. De Rosa, S. C., Herzenberg, L. A., Herzeberg, L. A. & Roederer, M. 11-color, 13-parameter flow cytometry: identification of human naive T cells by phenotype, function, and T-cell receptor diversity. *Nat. Med.* **7**, 245–248 (2001).
132. Lee, P. P. *et al.* Characterization of circulating T cells specific for tumor-associated antigens in melanoma patients. *Nat. Med.* **5**, 677–685 (1999).
133. Altman, J. D. *et al.* Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94–96 (1996).
134. Ramachandiran, V. *et al.* A robust method for production of MHC tetramers with small molecule fluorophores. *J. Immunol. Methods* **319**, 13–20 (2007).
135. Bakker, A. H. *et al.* Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7. *Proc. Natl Acad. Sci. USA* **105**, 3825–3830 (2008).
136. Celie, P. H. *et al.* UV-induced ligand exchange in MHC class I protein crystals. *J. Am. Chem. Soc.* **131**, 12298 (2009).
137. Kwong, G. A. *et al.* Modular nucleic acid assembled p/MHC microarrays for multiplexed sorting of antigen-specific lymphophocytes. *J. Am. Chem. Soc.* **131**, 9695–9703 (2009).
138. Rodenko, B. *et al.* Generation of peptide–MHC class I complexes through UV-mediated ligand exchange. *Nat. Protoc.* **1**, 1120–1132 (2006).
139. Bordon, Y. Immunotherapy: checkpoint parley. *Nat. Rev. Cancer* **15**, 3 (2015).
140. Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* **348**, 69–74 (2015).
141. Rosenberg, S. A. & Restifo, N. P. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science* **348**, 62–68 (2015).
142. Gattinoni, L., Klebanoff, C. A. & Restifo, N. P. Paths to stemness: building the ultimate antitumour T cell. *Nat. Rev. Cancer* **12**, 671–684 (2012).
143. Stroncek, D. F. *et al.* New directions in cellular therapy of cancer: a summary of the summit on cellular therapy for cancer. *J. Transl. Med.* **10**, 48–52 (2012).
144. Restifo, N. P. & Gattinoni, L. Lineage relationship of effector and memory T cells. *Curr. Opin. Immunol.* **25**, 556–563 (2013).
145. Polyak, K. Tumor heterogeneity confounds and illuminates: a case for Darwinian tumor evolution. *Nat. Med.* **20**, 344–346 (2014).
146. Furnari, F. B., Cloughesy, T. F., Cavenee, W. K. & Mischel, P. S. Heterogeneity of epidermal growth factor receptor signalling networks in glioblastoma. *Nat. Rev. Cancer* **15**, 302–310 (2015).
147. Nowell, P. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
148. Aktipis, C. A., Boddy, A. M., Gatenby, R. A., Brown, J. S. & Maley, C. C. Life history trade-offs in cancer evolution. *Nat. Rev. Cancer* **13**, 883–892 (2013).
149. Almendro, V., Marusyk, A. & Polyak, K. Cellular heterogeneity and molecular evolution in cancer. *Annu. Rev. Pathol.* **8**, 277–302 (2013).
150. Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495–501 (2014).
151. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
152. Quail, D. F. & Joyce, J. A. Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* **19**, 1423–1437 (2013).
153. Semenza, G. L. Cancer–stromal cell interactions mediated by hypoxia-inducible factors promote angiogenesis, lymphangiogenesis, and metastasis. *Oncogene* **32**, 4057–4063 (2013).
154. Yaffe, M. B. The scientific drunk at the lamppost: massive sequencing efforts in cancer discovery and treatment. *Sci. Signal.* **6**, e13 (2013).
155. Brennan, C. W. *et al.* The somatic genetic landscape of glioblastoma. *Cell* **155**, 462–477 (2013).
156. Chin, L. *et al.* Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061–1068 (2008).
157. Inda, M.-d.-M. *et al.* Tumor heterogeneity is an active process maintained by a mutant EGFR-induced cytokine circuit in glioblastoma. *Genes Dev.* **24**, 1731–1745 (2010).
158. Zadeh, G., Bhat, K. P. L. & Aldape, K. EGFR and EGFRvIII in glioblastoma: partners in crime. *Cancer Cell* **24**, 403–404 (2013).
159. Bachoo, R. M. *et al.* Epidermal growth factor receptor and Ink4a/Arf: convergent mechanisms governing terminal differentiation and transformation along the neural stem cell to astrocyte axis. *Cancer Cell* **1**, 269–277 (2002).
160. Sottoriva, A. *et al.* Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc. Natl Acad. Sci. USA* **110**, 4009–4014 (2013).
161. Gill, B. J. *et al.* MRI-localized biopsies reveal subtype-specific differences in molecular and cellular composition at the margins of glioblastoma. *Proc. Natl Acad. Sci. USA* **111**, 12550–12555 (2014).
162. Klages, R. *et al. Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond* (Wiley, 2013).
163. Losick, R. & Desplan, C. Stochasticity and cell fate. *Science* **320**, 65–68 (2008).
164. Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L. & Leibler, S. Bacterial persistence as a phenotypic switch. *Science* **305**, 1622–1625 (2004).
165. Aguirre-Ghiso, J. A. Models, mechanisms and clinical evidence for cancer dormancy. *Nat. Rev. Cancer* **7**, 834–846 (2007).
166. Kholodenko, B., Yaffe, M. B. & Kolch, W. Computational approaches for analyzing information flow in biological networks. *Sci. Signal.* **5**, re1 (2012).

Acknowledgements
The authors acknowledge the following funding agencies and grants for support of some of the work described in this Review: The National Cancer Institute (5R01CA170689 to J.R.H. as the principal investigator (PI) and A.R. as the co-PI, and 5U54 CA119347 to J.R.H. as the PI); Stand up to Cancer Foundation (to A.R. and J.R.H.); the Cancer Research Institute (to A.R. and J.R.H.); The Ben and Catherine Ivy Foundation (to P.S.M. and J.R.H.); and the Jean Perkins Foundation (to J.R.H. as the PI).

Competing interests statement
The authors declare competing interests: see Web version for details.