

Single-cell technologies to study the immune system

Valentina Proserpio^{1,2} and Bidesh Mahata^{1,2}

¹Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, and ²European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge, UK

doi:10.1111/imm.12553

Received 31 August 2015; revised 29

October 2015; accepted 1 November 2015.

Correspondence: Valentina Proserpio and Bidesh Mahata, EMBL-European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK. Emails: vproserp@ebi.ac.uk and bmahata@ebi.ac.uk

Senior author: Valentina Proserpio

Summary

The immune system is composed of a variety of cells that act in a coordinated fashion to protect the organism against a multitude of different pathogens. The great variability of existing pathogens corresponds to a similar high heterogeneity of the immune cells. The study of individual immune cells, the fundamental unit of immunity, has recently transformed from a qualitative microscopic imaging to a nearly complete quantitative transcriptomic analysis. This shift has been driven by the rapid development of multiple single-cell technologies. These new advances are expected to boost the detection of less frequent cell types and transient or intermediate cell states. They will highlight the individuality of each single cell and greatly expand the resolution of current available classifications and differentiation trajectories. In this review we discuss the recent advancement and application of single-cell technologies, their limitations and future applications to study the immune system.

Keywords: CD4⁺ T helper cells; immune cells; single-cell RNA-sequencing; single-cell technology.

Introduction

200 years of cell history

In the first half of the nineteenth century, Schleiden, Schwann and Virchow developed the cell theory stating that 'the cell' is the structural and functional unit of all living things. For the next 170 years scientists have been looking at the biological processes from a cell-population point of view, separating and categorizing cells into subclasses and subpopulations based on their morphology and phenotype. In the past century, this approach of analysing cells at a population level in which all cells are assumed to behave in a constant manner, allowed scientists to study and characterize all the fundamental biological processes that are the basis of our current knowledge of life in a top-down manner. One such example is the process of blood and immune cell differentiation – haematopoiesis. However, the aspect that has been neglected is the contribution that each single cell makes within a population. This is mainly a result of the lack of tools available to address questions from a single-cell perspective.

In the last few years, there has been a rapid development in single-cell technologies, which have revealed huge variability among cells traditionally assigned to the same category.^{1–3}

In this review we analyse the impact and potential of these innovations, with special emphasis on single-cell RNA sequencing (scRNA-seq) as applied to the immune system.

The immune system

The immune system is not only responsible for the defence of the organism from a plethora of diverse infections ranging from bacteria to viruses, but also protects us by healing wounds and clearing cancerous cells. The efficiency of the immune response is dependent on the coordinated and balanced behaviour of a multitude of different cells involved in each stage of the process. This includes pathogen recognition, the initiation of the signalling cascade that leads to recruitment of other effector cells, and the final clearance of the infection.

The classification of all the cells involved in this process has progressively been amplified over the years. This is partly a result of the continuous development and application of more advanced technologies. Starting from the discovery of red blood cells in 1695 and the identification of white blood cells in 1843, every advance in technology has added a new layer of complexity and more and more subcategories in the blood composition tree (Fig. 1).⁴ With the recent growth of high-throughput single-cell technologies we are now realizing that even within a

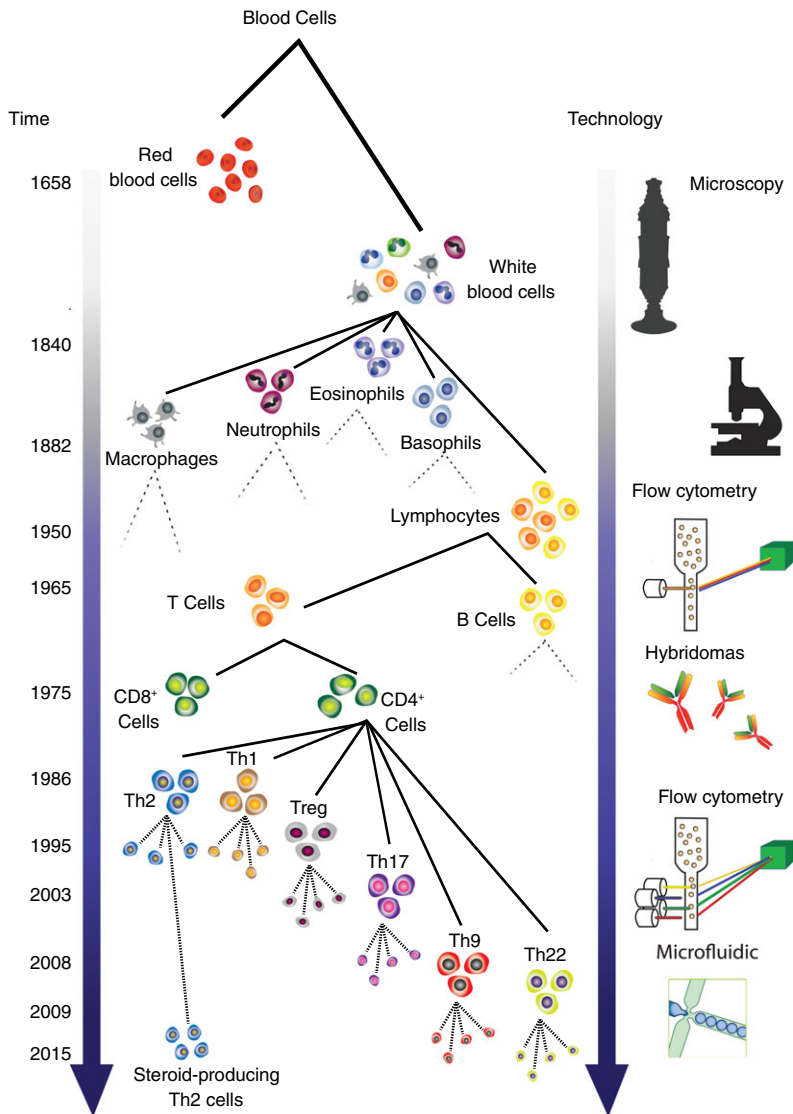


Figure 1. The complexity of the blood cell populations has grown in parallel with the development of always more sophisticated technology. From the discovery of red blood cells in 1658 by the Dutch naturalist, Jan Swammerdam, almost 200 years passed until the identification of leucocytes (1843) by two independent physicians from England and France establishing the beginning of haematology as a new field in medicine. The molecular characterization of the leucocytes required the advent of flow cytometry (1960) and monoclonal antibodies (1975). The latter were a crucial tool for the discrimination of CD4⁺ and CD8⁺ T helper cells. In the next decades the scenario of CD4⁺ T helper cells became more and more complicated with the discovery of distinct subclasses. In 1986 Mosman and Coffman revealed the existence of two functional subsets, termed T helper 1 (Th1) and T helper 2 (Th2). In 1995 Dr Sakaguchi^{60,61} discovered another specific subpopulation of T cells, named regulatory T (Treg) cells, that were specialized for immunosuppression. More recently other subsets have been isolated named Th17 (2005),^{62,63} Th9 (2008)^{64,65} and Th22 (2009).⁴⁰ Finally, single-cell RNA-sequencing has revealed the existence of a subpopulation of steroid-producing cells within the Th2 compartment.²³

well-defined subgroup, there is significant structural and functional heterogeneity. The high throughput study at single-cell level would allow us to investigate the immune cell population in a bottom-up way, and unravel this heterogeneity in a quantitative manner.³

Recent development of single-cell techniques

In the past few years the development of a multitude of single-cell technologies has allowed scientists to dissect different cellular scenarios from a new single-cell perspective. These techniques can be subdivided into four main categories according to which cellular component is under study: the whole cell, the protein content, the DNA and the RNA (Table 1). Sequencing of RNA, DNA and bisulphite are independent of any prior knowledge about the population analysed and are ‘OMICS’ techniques that will give a comprehensive picture of the single-cell state.

Table 1. Different available methods for single-cell sequencing

Method	Principle	Strand-specific?	Positional bias?	Ref
Tang <i>et al.</i>	PolyA tailing	No	3' (weak)	6,14
STRT	5' selection	Yes	5' (strong)	7,11
SMART-seq	Template switching	No	3' (weak)	8,12,13
CEL-seq/ MARS-seq	<i>In vitro</i> transcription	Yes	3' (strong)	9
Quartz-seq	PolyA tailing	No	3' (weak)	10

All the other technologies share the limitations of being restricted to the analysis of fewer, pre-identified markers. Relying on fewer genes has the potential drawback that results can be biased to some extent and it generally fails to give a global picture of the process analysed.

Moreover, some of these new techniques that require little input material are essential when focusing on transcriptomic/genomic profiles of rare cells, as for example with circulating tumor cells, fetal circulating cells, haematopoietic stem cells in peripheral blood and in the bone marrow, and antigen-specific T or B cells that occur at very low frequencies in the population. The gene expression profiles of these rare cells were difficult, if not impossible, to study with the previous technologies and only now are we able to explore the heterogeneity within these specific subclasses of cells.

Single-cell RNA-seq, the good and the bad of a powerful technique

Scientists have made great strides in the 25 years after Iscove and co-workers⁵ succeeded in the purification and amplification of mRNA from one single cell in 1990. Two decades of technological and computational progress have recently resulted in the advent of scRNA-seq technology that allows the simultaneous quantification of the expression of most genes in one single cell.

Even though it seems obvious, we need to underline that single cells do not have a perfect biological replicate as each cell is, by definition, a unique unit. The high variability observed between two different cells could be either biological or purely technical. The technical variation is mainly the result of the extremely low amount of input material, which can in turn cause a substantial loss during cDNA synthesis and promote

amplification biases during the multiple rounds of PCR that are required before the sequencing can be performed.

In the past 6 years, five main methods have been developed and optimized to reverse transcribe the mRNA and amplify the cDNA from one single cell to achieve a better coverage and a lower cost per cell^{6–14} (Table 2). A parallel development of multiple algorithms has taken place in order to deal with the huge amount of data these new experiments have produced.¹⁵ These computational methods become crucial for data interpretation because this new technology generates an incredible amount of data, which require faster and more standardized computational methods. The data are also ‘corrupted’ by numerous confounding factors and biases that need to be corrected for, using automated methods.^{16–20}

To overcome these limitations, different technical strategies have been adopted to calculate the inherent technical variability; for instance, the addition of artificial nucleic acid DNA and RNA standards or spike-in External RNA Controls Consortium (ERCC) molecules that are assumed to be the same across the samples. Alternatively, the count of the mRNA molecules per cell, in which each molecule is individually labelled with random DNA sequences (Unique Molecular Identifiers²¹). According to the sample preparation method, different computational approaches can be used to calculate gene expression level.^{16,19,22}

After data normalization and gene expression level calculation, unsupervised clustering approaches (principal

Table 2. Different available techniques for analysing single cells at Cell/Protein/DNA and RNA level

	Technique	Pros	Cons
Cellular level	Live imaging	Easy to use, available in many laboratories. Live cells	Laborious, long data processing. Restricted to few genes of interest
	Reporter cells	Easy to use, available in many laboratories. Live cells	Restricted to few genes of interest
	Lineage tracing	Useful for developmental studies	Laborious, long data processing. Restricted to few genes of interest
Protein level	Flow cytometry	Well-established technique. Easy to use and available in many laboratories	Restricted to few genes of interest (up to 17). Limited to surface markers for live cells
	Immunofluorescence	Well-established technique. Easy to use, available in many laboratories	Restricted to few genes of interest. Manual data analysis
	CytOF	Up to 40 different proteins analysed in single cells. No compensation required	Costly and specific machine is required
DNA level	Amnis	Automatic data analysis	Restricted to few genes of interest
	Single cell (Sc)-genome Sc-bisulphite sequencing	No pre-knowledge required	Costly
RNA level	Sc-quantitative PCR	Quick results	Costly
	Sc-sequencing	Global profiling, no need of pre-knowledge	Costly Slower than quantitative PCR
	Single molecule RNA-fluorescence	Absolute mRNA count	Laborious, long data processing. Restricted to few genes of interest
	<i>in situ</i> hybridization		

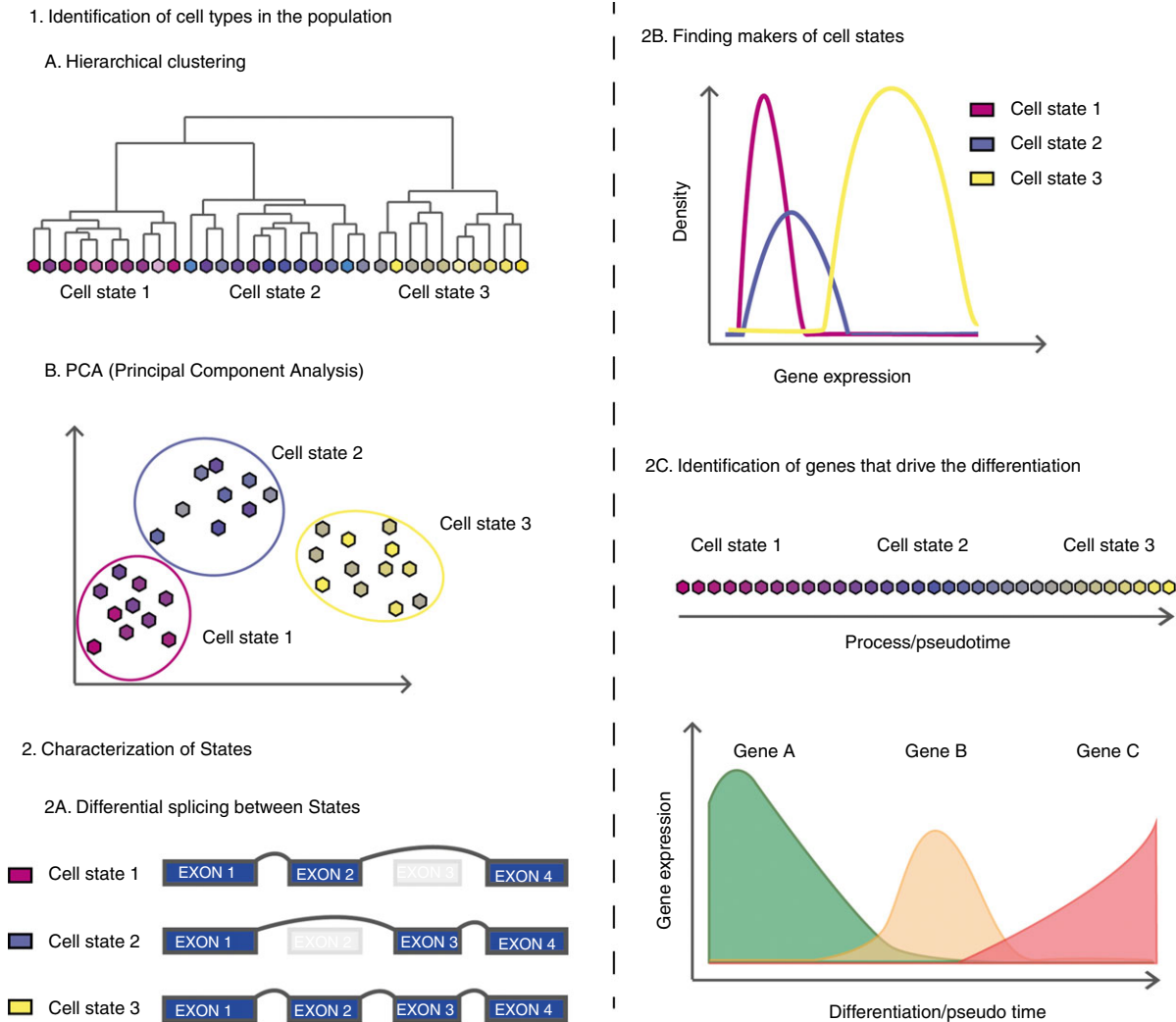


Figure 2. Identification and characterisation of novel immune cell types and cell states (1) Identification of novel immune cell populations or distinct cell states can be performed using hierarchical clustering (1A) or principle component analysis (PCA) for example (1B). (2) Analysis of differential splicing: specific splice variants may associate with a subpopulation of immune cells or cell state because of their differential function (2A). Example of different approaches to characterise novel cell states. Find markers of cell types by analysing differential expression between different groups of cells (2B), identification of genes that show particular pattern during differentiation such as during developmental maturation of immune cells or in response to immunogenic stimuli: genes that either increase, decrease or are transiently expressed (2C).

component analysis and hierarchical clustering) can help in identifying new distinct subpopulations (Fig. 2)^{23,24}, but the discrimination between cell state and cell type still needs to be further validated experimentally. In other words, the distinction between physiological fluctuations of gene expressions without phenotypic changes, and different cells types cannot be made solely by analysis of gene expression pattern. This is especially true in the context of the immune system where cells are known to frequently shift from an inactive to an active state. One of the main problems that we face when analysing single cells is that each cell can be in a different cell cycle phase and can therefore display variation not only in terms of gene expression abundance, but also in terms of size and

RNA content. This difference, that in bulk data is completely averaged out, may account for most of the variation we observe in single-cell data and needs to be corrected for.¹⁷ The same is also true for stress conditions and other factors that might introduce another level of variability within the cell population analysed.

Heterogeneity in the immune system

The heterogeneity displayed by cells in the immune system reflects the extreme flexibility and plasticity that makes the immune system efficient in fighting different pathogens. At the same time, it poses a big challenge that scientists face when studying these cells. The word

'heterogeneity' has been used with regards to the immune system in immune cells since the early 1970s,²⁵ and the level of complexity and diversification has grown since then. The diversity includes diversity at the DNA level as well as the RNA and protein levels. Regarding the DNA level, it is worth noting that, whereas the human genome contains roughly 30 000 genes, the number of T-cell receptors (TCRs) is estimated to be in the order of 10^7 and the same is true for the B-cell receptors.^{26,27} This variability is achieved by somatic recombination of the DNA, a process that is restricted to B and T cells only. This extreme situation is indicative of the importance of the heterogeneity in the immune system, as evolution has selected a way to guarantee the recognition of a multitude of pathogens through the variability of those repertoires. The same happens at the RNA and protein levels. For instance the secretion of specific signalling cytokines by immune cells displays a high rate of diversity as quantified by flow cytometry.

Early reports from 1995 and 1996^{28,29} underlined how the level of synthesis of different cytokines was different under different stimuli in T helper type 1 (Th1) compared with Th2 cells. It was also clear that, even when looking at just two cytokines at a time, cells display enormous cell-to-cell variability. With improvements in flow cytometry technology we are now able to analyse 17 different parameters at the same time per cell.³⁰

The CD4⁺ T-cell lesson

From the discovery of CD4⁺ T cells (T helper cells) in the early 1980s^{31,32} with the advent of monoclonal antibody technology and the initial distinction between Th1 and Th2 from Tim Mosmann and Bob Coffman^{33,34} the T helper subtypes organization has become far more complex. Starting from the two main Th1 and Th2 groups, five other new members have now joined the CD4⁺ T-cell panorama, such as Th17,^{35,36} the regulatory T cells,³⁷ the follicular helper T cells³⁸ and, most recently, the Th9³⁹ and Th22⁴⁰ subsets. This is a very simplified categorization, but, the overall picture, rather than becoming clearer over time, has actually become more blurred (Fig. 3). This subsequent complication is mainly caused by the fact that the distinction among different subtypes is defined by the expression of few specific genes⁴¹ and the intermediate cell types (expressing markers of different subtypes at the same time) has not yet been completely understood and hence incorporated.

Our knowledge of CD4⁺ T cells has significantly expanded over the last few years by employing single-cell technologies. For instance, single molecule RNA-fluorescence *in situ* hybridization combined with immunostaining has been employed to study the interplay between extracellular cytokines and intracellular transcription factors during the early phase of CD4⁺ T helper cell differ-

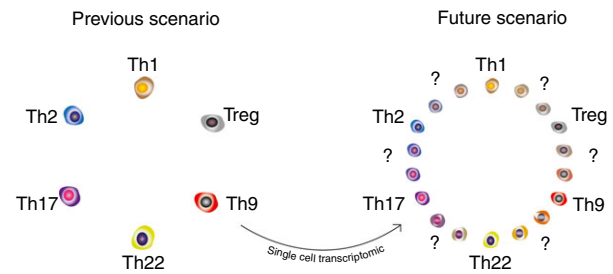


Figure 3. The static and discrete view of the CD4⁺ T helper cell population composition might be completely revised thanks to single-cell technology. In the new scenario many more intermediate subtypes as well as new subpopulations can be introduced by the whole transcriptomic profiles of single immune cells. In this scheme, CD4⁺ T helper cells are used as an example that can be applied to many different cell types.

entiation.⁴² The same process has been investigated at single-cell resolution⁴³ with conventional flow cytometry staining integrated with mathematical modelling. The integration of the results from both these works showed that the CD4⁺ T-cell differentiation scenario is more variegated than was thought and stochasticity seems to play a very important role in determination of cell fate. The original theory of the two mutually exclusive master regulators Gata3 and Tbx21 in instructing the Th2 versus the Th1 lineage commitment has now been substituted by the co-existence of a mixed continuum of T helper cells with a weak intracellular network that is strongly affected by a robust extracellular cytokine signal. In an article by Peine and co-workers,⁴⁴ the authors demonstrated the existence of an *in vivo* intermediate Th1/Th2 cell type. These cells, that were conventionally thought to be unstable, are shown to be highly stable and cannot be considered a mere undecided precursor. The same approach applied to other differentiation pathways might lead to similar conclusions.

A recently published paper from Mahata *et al.* revealed the existence of a Th2 subpopulation identified by the single-cell sequencing technologies. This specific subset of Th2 cells is distinguished from the rest of the population by the expression of a specific enzyme (*Cyp11a1*) that is at the basis of steroid biosynthesis.²³ The comprehensive data obtained by the single-cell transcriptomic approach allowed not only identification, but also purification of these cells based on new markers. Cells of interest were isolated from the whole population using antibodies against a surface molecule, the expression of which correlated with the expression of the cytoplasmic enzyme of interest, allowing a broad *ex vivo* functional validation of the new cell type discovered. This is the first example of how the single-cell sequencing approach can ultimately shed light on the variegated compositions of the standard T helper cell subclasses, not only in terms of expression profile but also in terms of functional characterization of the cells.

TCR sequencing coupled with functional analysis of T cells

T-cell receptor repertoire studies have greatly benefited from the new single-cell sequencing approaches. Pioneer studies on TCR repertoire at single-cell level^{45–48} (revised in refs. 49,50) were limited to the analysis of TCR- α or TCR- β chains (sometimes one of the two only) without any information on the expression of other genes characteristic of different subpopulations of T cells.

A recently published work⁵¹ has finally coupled the analysis of α -chain and β -chain sequences with the expression of a panel of genes characteristic of differentiated T cells. This analysis showed how T cells bearing the same TCR- α and TCR- β sequences can exhibit substantial differences in the expression of cytokines and transcription factors, demonstrating for the first time that T cells derived from the same progenitor can actually differentiate towards different mature T cells.

Future directions

Dissection of immune cell populations

The advantage of scRNA-seq is the identification of cellular sub-populations in an unbiased way. Redetermination of immune cell types taking each cell's transcriptome into account in a bottom-up manner would reveal a new dimension. Immune responses often rely on sequential changes in cell state from inactive to active. Single-cell transcriptomic analysis of the transition between cellular states can reveal new insights into regulatory mechanisms. Whether the transitions between states are binary or graded is still unresolved. Understanding the nature of the immune cell activation process and possible intermediate states can lead to the identification of key genes that act as switches and drivers of these processes. Hence the clustering of cells into groups based on their single-cell transcriptomes is an important future task.

For example, can we regroup tumour-infiltrated immune cells (e.g. tumour-associated macrophages)? Tumour-associated macrophages are structurally plastic and heterogeneous, and functionally crucial for determining the fate of a tumour.⁵² These cells are widely accepted to be a cause of anti-tumour immune suppression. A single-cell study on tumour-infiltrated lymphoid and myeloid cells may lead to new insights and ways to identify cancer drug targets.

An alternative to the clustering approach is to use principle component analysis to identify cell types. Cellular transitions can be studied by defining cell states using hierarchical clustering or principal component analysis-like methods. The approach has been applied to show how cells change gradually along the developmental pathway from zygote to the late blastocyst.⁵³ In future, a simi-

lar approach can be applied to the *in vivo* immune cell activation/inactivation dynamics and their dysfunction.

Promoterwise gene expression kinetics

Variability in the gene's expression levels across cells provides information on how tightly its expression is controlled. Genes with higher transcriptional bursts and lower frequency are noisier than genes that are expressed in small frequent bursts.⁵⁴ An scRNA-seq based study of bone marrow-derived dendritic cell activation revealed that certain genes had a bimodal pattern of gene expression.⁵⁵ Subsequent studies showed that paracrine signalling of a subset of fast-responding dendritic cells affects the whole population.⁵⁶ Future studies on other specific groups of immune cell populations (e.g. T cells) following a similar approach would reveal further insights.

Differential splicing: a mode of immune regulation

Some scRNA-seq methods (e.g. Smart-Seq) that provide full transcript coverage can allow us to determine and quantify the alternative splice forms. Immune cell heterogeneity due to the differential expression of splice variants is still underexplored (Fig. 2). The report from Shalek *et al.*^{55,56} showed that the predominant isoform of several genes differs between different dendritic cell populations. It remains to be seen whether this paradigm holds more generally in other immune cells and plays a functional role in building an effective immune response.

Gene–gene correlation studies and gene-regulatory networks

Cell-to-cell variability in gene expression can be used to infer gene regulatory interactions and gene regulatory networks.^{57,58} More specifically, by using scRNA-seq data, gene-regulatory modules (sets of genes that are co-regulated) can be inferred by calculating gene-to-gene correlations or by clustering genes based on gene expression profiles across cells (Fig. 2). Such an approach is expected to reveal novel genes correlated with ones already known to play an important role in the immune response. We have applied such an approach to identify cell surface receptors that were specifically enriched in the steroid-producing Th2 cells.²³ Further application of this approach is anticipated to be beneficial for immunological studies.

Concluding remarks

Assessment of bulk immune cell populations using classic immunological techniques provided an imperfect systems level view that greatly underestimated the constitutive and functional diversity of the immune system. This is

due to the averaging effect and contributions of minor but functionally important populations. Both the huge cell-to-cell variability and the amount of data these new technologies have brought to light constitute a great resource and at the same time a challenge for scientists. The parallel development of more sophisticated algorithms, faster computational approaches and new data visualization methods has already allowed scientists to gain new insights into immune system diversity.^{18,23,55,56}

Despite this incredible progress, there are still some issues that need to be solved and most of the works involving single-cell sequencing are methodological. Few mechanistic questions have been addressed with this technology so far. This limitation is partly due to the high complexity of the data and the inability to compare different data sets due to batch effect variation.

The need for a standardized, more consistent and globally approved method for data normalization and data analysis will allow us to merge and compare the multitude of data sets that scientists from around the world are producing, enabling a deeper, single-cell-centred characterization of the architecture of different organs, and ultimately the whole body.⁵⁹

Though significant progress has been made capturing the single-cell transcriptome, obtaining the proteome at single-cell resolution is far from a reality. The most ambitious goal of single-cell-level studies would be the integration of transcriptomic and proteomic profiles from a single cell. The exact correlation between protein and RNA content at the single-cell level would be an interesting area of research specifically in immunobiology.

Acknowledgements

This work was supported by European Research Council grant ThSWITCH (grant number: 260507). The authors thank Dr Sarah Teichmann for critical reading of the manuscript and Aleksandra Kolodziejczyk for her input on Figure 2.

Disclosures

The authors declare that they have no competing interests.

References

- Pollen AA, Nowakowski TJ, Shuga J, Wang X, Leyrat AA, Lui JH, et al. Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat Biotechnol* 2014; **32**:1053–8.
- Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 2014; **344**:1396–401.
- Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretzky I, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* 2014; **343**:776–9.
- Hajdu SI. A note from history: the discovery of blood cells. *Ann Clin Lab Sci* 2003; **33**:237–8.

- Brady G, Barbara M, Iscove NN. Representative *in vitro* cDNA amplification from individual hemopoietic cells and colonies. *Methods Mol Cell Biol* 1990; **2**:17–25.
- Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 2009; **6**:377–82.
- Islam S, Kjällquist U, Moliner A, Zajac P, Fan JB, Lönnerberg P, et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res* 2011; **21**:1160–7.
- Ramskold D, Luo S, Wang YC, Li R, Deng Q, Faridani OR, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol* 2012; **30**:777–82.
- Hashimshony T, Wagner F, Sher N, Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep* 2012; **2**:666–73.
- Sasagawa Y, Nikaïdo I, Hayashi T, Danno H, Uno KD, Imai T, et al. Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol* 2013; **14**:R31.
- Islam S, Kjällquist U, Moliner A, Zajac P, Fan JB, Lönnerberg P, et al. Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nat Protoc* 2012; **7**:813–28.
- Picelli S, Björklund Å, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods* 2013; **10**:1096–8.
- Picelli S, Faridani OR, Björklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* 2014; **9**:171–81.
- Tang F, Barbacioru C, Nordman E, Li B, Xu N, Bashkirov VI, et al. RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat Protoc* 2010; **5**:516–35.
- Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet* 2015; **16**:133–45.
- Brennecke P, Anders S, Kim JK, Kolodziejczyk AA, Zhang X, Proserpio V, et al. Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* 2013; **10**:1093–5.
- Buettner F, Natarajan KN, Casale FP, Proserpio V, Scialdone A, Theis FJ, et al. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat Biotechnol* 2015; **33**:155–60.
- Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 2014; **32**:381–6.
- Reiter M, Kirchner B, Müller H, Holzhauser C, Mann W, Pfaffl MW. Quantification noise in single cell experiments. *Nucleic Acids Res* 2011; **39**:e124.
- Wilson NK, Kent DG, Buettner F, Shehata M, Macaulay IC, Calero-Nieto FJ, et al. Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. *Cell Stem Cell* 2015; **16**:712–24.
- Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods* 2014; **11**:163–6.
- Ding B, Zheng L, Zhu Y, Li N, Jia H, Ai R, et al. Normalization and noise reduction for single cell RNA-seq experiments. *Bioinformatics* 2015; **31**:2225–2227.
- Mahata B, Zhang X, Kolodziejczyk AA, Proserpio V, Haim-Vilmovsky L, Taylor AE, et al. Single-cell RNA sequencing reveals T helper cells synthesizing steroids *de novo* to contribute to immune homeostasis. *Cell Rep* 2014; **7**:1130–42.
- Treutlein B, Brownfield DG, Wu AR, Neff NF, Mantalas GL, Espinoza FH, et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* 2014; **509**:371–5.
- Talal N. Lymphocyte heterogeneity and function. *Arthritis Rheum* 1973; **16**:422–5.
- Nikolich-Zugich J, Slifka MK, Messaoudi I. The many important facets of T-cell repertoire diversity. *Nat Rev Immunol* 2004; **4**:123–32.
- Schroeder HW. Similarity and divergence in the development and expression of the mouse and human antibody repertoires. *Dev Comp Immunol* 2006; **30**:119–35.
- Openshaw P, Murphy EE, Hosken NA, Maino V, Davis K, Murphy K, et al. Heterogeneity of intracellular cytokine synthesis at the single-cell level in polarized T helper 1 and T helper 2 populations. *J Exp Med* 1995; **182**:1357–67.
- Hussell T, Spender LC, Georgiou A, O'Garra A, Openshaw PJ. Th1 and Th2 cytokine induction in pulmonary T cells during infection with respiratory syncytial virus. *J Gen Virol* 1996; **77**:2447–55.
- Perfetto SP, Chattopadhyay PK, Roederer M. Seventeen-colour flow cytometry: unravelling the immune system. *Nat Rev Immunol* 2004; **4**:648–55.
- Reinherz EL, Schlossman SF. The differentiation and function of human T lymphocytes. *Cell* 1980; **19**:821–7.
- Engleman EG, Benike CJ, Grumet FC, Evans RL. Activation of human T lymphocyte subsets: helper and suppressor/cytotoxic T cells recognize and respond to distinct histocompatibility antigens. *J Immunol* 1981; **127**:2124–9.
- Mosmann TR, Chervinski H, Bond MW, Giedlin MA, Coffman RL. Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. *J Immunol* 1986; **136**:2348–57.
- Mosmann TR. T lymphocyte subsets, cytokines, and effector functions. *Ann N Y Acad Sci* 1992; **664**:89–92.

V. Proserpio and B. Mahata

- 35 Park H, Li Z, Yang XO, Chang SH, Nurieva R, Wang YH, *et al.* A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. *Nat Immunol* 2005; **6**:1133–41.
- 36 Harrington LE, Hatton RD, Mangan PR, Turner H, Murphy TL, Murphy KM, *et al.* Interleukin 17-producing CD4⁺ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. *Nat Immunol* 2005; **6**:1123–32.
- 37 Josefowicz SZ, Rudensky A. Control of regulatory T cell lineage commitment and maintenance. *Immunity* 2009; **30**:616–25.
- 38 Crotty S. Follicular helper CD4 T cells (TFH). *Annu Rev Immunol* 2011; **29**:621–63.
- 39 Schmitt E, Klein M, Bopp T. Th9 cells, new players in adaptive immunity. *Trends Immunol* 2014; **35**:61–8.
- 40 Eyerich S, Eyerich K, Pennino D, Carbone T, Nasorri F, Pallotta S, *et al.* Th22 cells represent a distinct human T cell subset involved in epidermal immunity and remodeling. *J Clin Invest* 2009; **119**:3573–85.
- 41 Luckheeram RV, Zhou R, Verma AD, Xia B. CD4⁺ T cells: differentiation and functions. *Clin Dev Immunol* 2012; **2012**:925135.
- 42 Fang M, Xie H, Dougan SK, Ploegh H, van Oudenaarden A. Stochastic cytokine expression induces mixed T helper cell States. *PLoS Biol* 2013; **11**:e1001618.
- 43 Antebi YE, Reich-Zeliger S, Hart Y, Mayo A, Eizenberg I, Rimer J, *et al.* Mapping differentiation under mixed culture conditions reveals a tunable continuum of T cell fates. *PLoS Biol* 2013; **11**:e1001616.
- 44 Peine M, Rausch S, Helmstetter C, Fröhlich A, Hegazy AN, Kühl AA, *et al.* Stable T-bet⁺ GATA-3⁺ Th1/Th2 hybrid cells arise *in vivo*, can develop directly from naive precursors, and limit immunopathologic inflammation. *PLoS Biol* 2013; **11**:e1001633.
- 45 Dash P, McClaren JL, Oguin TH, Rothwell W, Todd B, Morris MY, *et al.* Paired analysis of TCR α and TCR β chains at the single-cell level in mice. *J Clin Invest* 2011; **121**:288–95.
- 46 Douek DC, Betts MR, Brenchley JM, Hill BJ, Ambrozak DR, Ngai KL, *et al.* A novel approach to the analysis of specificity, clonality, and frequency of HIV-specific T cell responses reveals a potential mechanism for control of viral escape. *J Immunol* 2002; **168**:3099–104.
- 47 Zhou J, Dudley ME, Rosenberg SA, Robbins PF. Selective growth, *in vitro* and *in vivo*, of individual T cell clones from tumor-infiltrating lymphocytes obtained from patients with melanoma. *J Immunol* 2004; **173**:7622–9.
- 48 Kim SM, Bhonsle L, Besgen P, Nickel J, Backes A, Held K, *et al.* Analysis of the paired TCR α - and β -chains of single human T cells. *PLoS One* 2012; **7**:e37338.
- 49 Miconnet I. Probing the T-cell receptor repertoire with deep sequencing. *Curr Opin HIV AIDS* 2012; **7**:64–70.
- 50 Newell EW, Davis MM. Beyond model antigens: high-dimensional methods for the analysis of antigen-specific T cells. *Nat Biotechnol* 2014; **32**:149–57.
- 51 Han A, Glanville J, Hansmann L, Davis MM. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat Biotechnol* 2014; **32**:684–92.
- 52 Noy R, Pollard JW. Tumor-associated macrophages: from mechanisms to therapy. *Immunity* 2014; **41**:49–61.
- 53 Deng Q, Ramskold D, Reinius B, Sandberg R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* 2014; **343**:193–6.
- 54 Kim JK, Marioni JC. Inferring the kinetics of stochastic gene expression from single-cell RNA-sequencing data. *Genome Biol* 2013; **14**:R7.
- 55 Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R, *et al.* Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* 2013; **498**:236–40.
- 56 Shalek AK, Satija R, Shuga J, Trombetta JJ, Gennert D, Lu D, *et al.* Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* 2014; **510**:363–9.
- 57 Moignard V, Woodhouse S, Haghverdi L, Lilly AJ, Tanaka Y, Wilkinson AC, *et al.* Decoding the regulatory network of early blood development from single-cell gene expression measurements. *Nat Biotechnol* 2015; **33**:269–76.
- 58 Kolodziejczyk AA, Kim JK, Tsang JC, Ilicic T, Henriksson J, Natarajan KN, *et al.* Single cell RNA-sequencing of pluripotent states unlocks modular transcriptional variation. *Cell Stem Cell* 2015; **17**:471–85.
- 59 Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet* 2013; **14**:618–30.
- 60 Sakaguchi S, Fukuma K, Kuribayashi K, Masuda T. Organ-specific autoimmune diseases induced in mice by elimination of T cell subset. I. Evidence for the active participation of T cells in natural self-tolerance; deficit of a T cell subset as a possible cause of autoimmune disease. *J Exp Med* 1985; **161**:72–87.
- 61 Sakaguchi S, Sakaguchi N, Asano M, Itoh M, Toda M. Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor α -chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. *J Immunol* 1995; **155**:1151–64.
- 62 Cua DJ, Sherlock J, Chen Y, Murphy CA, Joyce B, Seymour B, *et al.* Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. *Nature* 2003; **421**:744–8.
- 63 Langrish CL, Chen Y, Blumenschein WM, Mattson J, Basham B, Sedgwick JD, *et al.* IL-23 drives a pathogenic T cell population that induces autoimmune inflammation. *J Exp Med* 2005; **201**:233–40.
- 64 Veldhoen M, Uytendhove C, van Snick J, Helmbly H, Westendorf A, Buer J, *et al.* Transforming growth factor- β 'reprograms' the differentiation of T helper 2 cells and promotes an interleukin 9-producing subset. *Nat Immunol* 2008; **9**:1341–6.
- 65 Dardalhon V, Awasthi A, Kwon H, Galileos G, Gao W, Sobel RA, *et al.* IL-4 inhibits TGF- β -induced Foxp3⁺ T cells and together with TGF- β , generates IL-9⁺ IL-10⁺ Foxp3⁺ effector T cells. *Nat Immunol* 2008; **9**:1347–55.