



# Single cell transcriptomics comes of age

Sarah Aldridge <sup>1</sup> & Sarah A. Teichmann <sup>1,2</sup> 

Single cell transcriptomics technologies have vast potential in advancing our understanding of biology and disease. Here, Sarah Aldridge and Sarah Teichmann review the last decade of technological advancements in single-cell transcriptomics and highlight some of the recent discoveries enabled by this technology.

The past decade has seen a revolution in single-cell transcriptomics. Here we describe advances made and the transformation that this powerful approach has on our ability to build detailed cellular maps of tissues, gaining unique insights in health and disease.

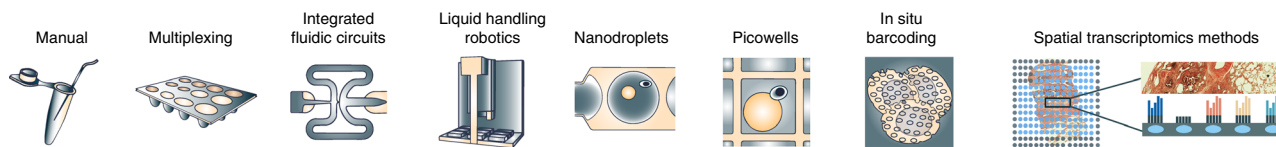
## Development of single-cell transcriptomics technologies over the past decade

The history of single-cell transcriptomics connects to precursor methods, such as single-cell qPCR, performed for the first time on a small number of genes by several groups in the early 90 s (e.g. refs. <sup>1,2</sup>). To our knowledge, the first example of bona fide single-cell transcriptomics is the study of a handful of mouse primordial germ cells by Tang et al. in 2009<sup>3</sup>. While this work was motivated by the sparsity of very rare cells that arise during murine development, key to driving forwards single-cell transcriptomics has been the scaling of technologies to profile large numbers of cells in parallel. The ability to perform unbiased single-cell transcriptome-wide analysis coupled with exponential scaling in cell numbers that can be analysed simultaneously has almost followed a “Moore’s Law” of single-cell genomics.

Following Tang et al., other single-cell transcriptomics protocols were developed including tag sequencing methods such as STRT-seq, and CEL-seq, MARS-seq and full length protocols such as SMART-seq/SMART-seq2, as well as other methods (Fig. 1). These protocols differed in their amplification technology and transcript coverage, as well as in the extent of robotisation of liquid handling in plates. This was followed by development of nanodroplets, picowell technologies and in situ barcoding which have made it possible to sequence tens of thousands of cells in parallel (reviewed in ref. <sup>4</sup>). In recent years multi-modal single-cell methodologies that measure and integrate genomics readouts from different molecules (RNA, DNA and protein)<sup>5</sup>, enabled dissection of the complex regulatory and cell-cell communication networks that drive cell identity to a greater degree than ever before.

Computational methods for low-level processing, quality control and ultimately sophisticated data interpretation were developed rapidly alongside the first single-cell transcriptomics data sets<sup>6</sup>. These methods are crucial for defining cell states and types from data, revealing marker genes and identifying developmental trajectories that relate cells to each other. The more recent multi-omics methods are catalysing new computational methods<sup>5</sup> as are spatial methods discussed below.

<sup>1</sup> Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB1 1SA, UK. <sup>2</sup> Theory of Condensed Matter Group, Cavendish Laboratory/Department of Physics, University of Cambridge, JJ Thomson Ave, Cambridge CB3 0HE, UK. ✉email: [st9@sanger.ac.uk](mailto:st9@sanger.ac.uk)



**Fig. 1 Development of single-cell technologies.** Significant technology developments have enabled the profiling of large numbers of cells in parallel by single-cell transcriptomics. Initial manual methods allowed single-cell transcriptomic analysis on only a few cells. The development of integrated fluidic circuits and the introduction of liquid handling robots into the process brought cell numbers to several thousand and then further increased with nanodroplet and picowell technologies. The introduction of in situ barcoding has increased throughput even further to hundreds of thousands of cells with the latest developments in spatial methods integrating the spatial location of transcriptomic information within tissue sections.

Methods for whole transcriptome spatial mapping, characterisation of gene expression profiles whilst retaining information on spatial context of expression, are fast emerging. smFISH has been used for a number of years for analysing gene expression whilst retaining spatial information in tissue. Methodologies based on this include low-plex that analyse only a small number of genes in parallel, up to highly multiplexed that interrogate hundreds or thousands of mRNA targets in parallel, for example MERFISH, in situ sequencing, seqFISH, that have emerged over the past five years or so. Spatial transcriptomics methods, first developed by Stahl et al., can be used to interrogate the whole transcriptome coupled with spatial location, with most recent methods beginning to reach single-cell resolution. Combining single-cell transcriptomics with computational and spatial methodologies will increasingly yield data-rich cell-atlases that map the spatial location of all cellular gene activities at high resolution.

### Single-cell transcriptomics is transforming our understanding of health and disease

As mentioned above, single-cell transcriptomic approaches were first applied to model organism, and continue to be instrumental in driving forwards the understanding of cells and tissues in the mouse and other organisms such as zebrafish, worm etc. In 2018, two high-throughput compendia of murine tissues profiled by single-cell transcriptomics were published (e.g. refs. <sup>7,8</sup>). These compendia provide data across many tissues using relatively uniform protocols, offering a rich biological resource. Focusing in on specific organs, a modified version of smFISH (osmFISH) was used to build an atlas of the mouse brain, giving a detailed view of cell identify and exact position within the tissue<sup>9</sup>.

The scaling of single-cell technologies is ultimately applicable to the entire human body. Therefore, catalysed by this single-cell revolution, the Human Cell Atlas (HCA) global consortium was co-founded in 2016, led by one of us (Sarah Teichmann) and Aviv Regev (Broad Institute) and including colleagues excited by the vision. The HCA is using high-throughput technologies at single-cell resolution to map cells in human tissue, analogous to a “Google maps” of the human body. The value of high-resolution transcriptomic data of individual cells in their tissue context is profound and is generating fundamental knowledge of each human tissue studied by the HCA community. A healthy reference atlas of both human tissues and model organisms is a basic prerequisite to assessing changes in ageing, disease and in response to therapeutic treatment (Fig. 2).

Examples of new understanding into basic biology from cell atlasing of human tissues has been highlighted as especially powerful by studies of human development. The first single-cell atlas of the maternal-fetal interface in early pregnancy (first trimester placenta and decidua) revealed new cell states (e.g. NK and stromal cells), and a more detailed cellular and molecular understanding of the layered architecture of this tissue and its role in maternal immune tolerance of paternal antigens<sup>10</sup>. This study revealed mechanisms behind healthy pregnancy and increases our

understanding of how these could change resulting in conditions such as pre-eclampsia.

Ground-breaking single-cell-atlases of development have been further extended to fundamental organs and systems including for example brain, heart, liver, thymus, gut and kidney. For example analysing thymic development at single-cell resolution has deepened understanding of new stromal and epithelial cell states and novel T cell types and indicates new principles of naïve T cell V(D)J repertoire formation, revealing that the human naïve T cell repertoire is biased in its receptor gene segments prior to pathogen exposure<sup>11</sup>.

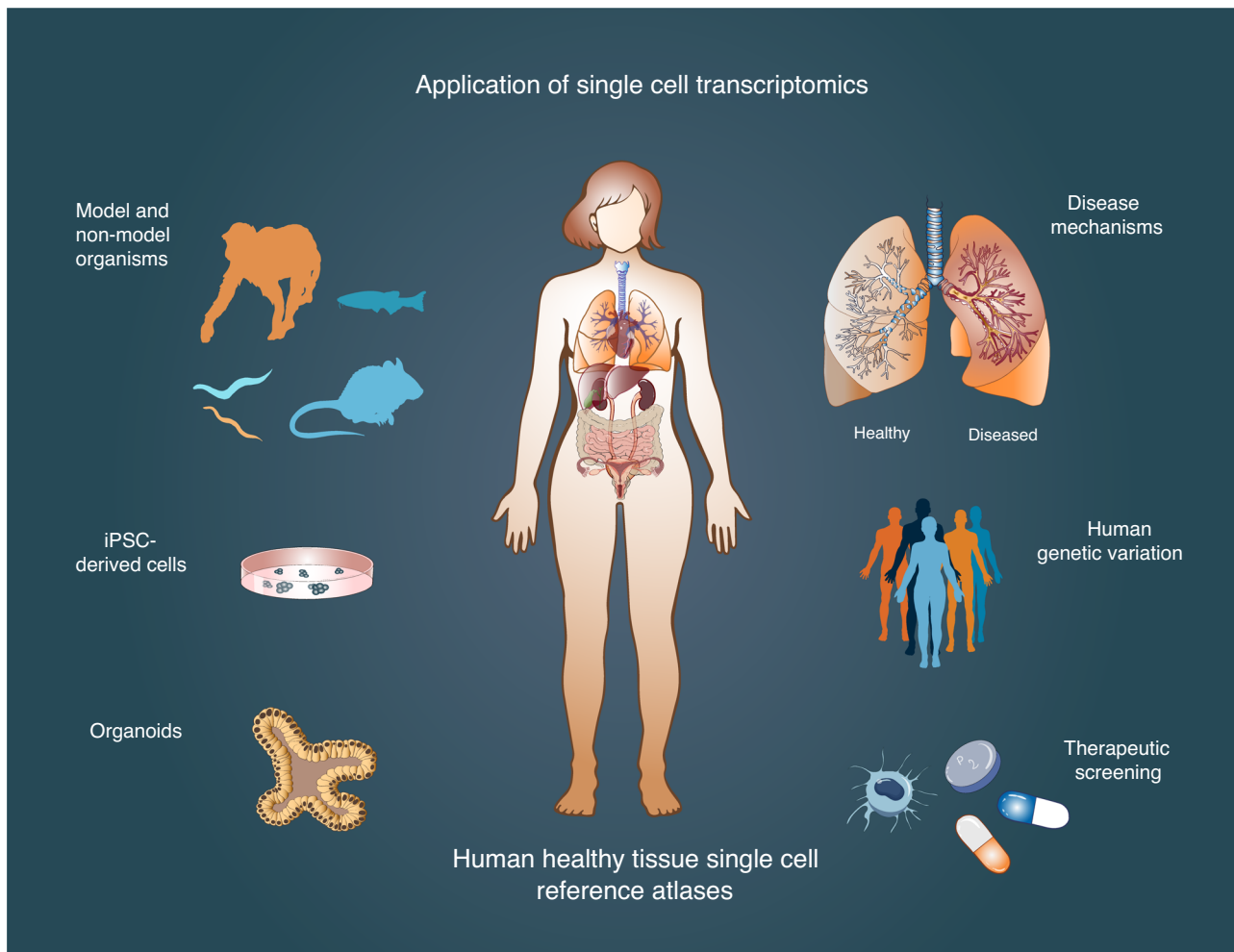
Single-cell transcriptomics have been applied to a spectrum of diseases including cancer and auto-inflammatory disease. The airways are a good example of this, where comparison of healthy and asthmatic lung at single-cell level uncovered novel mucosal ciliated cell states, and CD4+ T cell states in health (tissue migratory CD4+ T cell) and disease (pathogenic Th2 cells)<sup>12</sup>. This revealed potential new therapeutic targets and provides a framework for the use of single-cell transcriptomics in target discovery. Use of scRNA-seq and murine lineage tracing have been used to analyse the cellular composition of the tracheal airway, revealing the cystic fibrosis gene CFTR is expressed in a rare ionocyte cell population at a higher level than in abundant ciliated cells<sup>13</sup>. Insights such as these are transforming our understanding of diseases with genetic components such as cystic fibrosis and asthma.

In recent months single-cell transcriptomics has been harnessed to further our biological understanding of COVID-19 infection in response to the current global pandemic. Single-cell expression analysis of the coronavirus receptor ACE2, viral entry associated protease TMPRSS2 and infection associated genes in healthy human individuals and non-human primates across a range of tissues gives insights into potential sites of viral transmission. Two specific cell types within the nose look likely to be sites of infection and could explain high rates of transmission (reviewed in ref. <sup>14</sup>). This provides a valuable platform for investigating infection, transmission and developing clinical strategies for prevention and therapy with further single-cell studies underway.

### Future outlook

An exciting area is single-cell transcriptomics for improved engineering of in vitro cellular models such as organoids. It is becoming clear that single-cell transcriptomics can be applied across a breadth of medically important and competitive areas, including cell-based models, cell therapies, regenerative medicine and target discovery.

The utility of single-cell transcriptomics for in-depth validation of organoid models is demonstrated with single-cell RNA-seq applied to models of the dorsal forebrain, showing reliable recapitulation of the variety of cell types seen in the human cerebral cortex<sup>15</sup>. HCA data provides a high-dimensional description of cell states, including developmental pathways beginning from early first trimester human embryogenesis. This provides a



**Fig. 2 Application of single-cell technologies.** Single-cell transcriptomics are being used to create reference maps of healthy human tissues, organs and systems at single-cell resolution. These approaches are also being applied to understand non-human organisms, including mice and non-human primates. Analysing healthy versus diseased tissues and genetic variation between individuals is also a valuable platform for the understanding of disease mechanisms. Factors identified from in vivo single-cell transcriptomic studies can also be applied to generate improved in vitro models and responses to therapeutic screening can be assayed at the single-cell level.

precise template for in vitro models, and by comparing in vivo and in vitro data, we can predict the factors required to generate improved in vitro systems such as iPSC-derived cells and organoids. Insights from organ cell-atlases, exemplified by the precise characterisation of CFTR expressing cells in the lung, can be applied to analysis of gene expression in specific cell types to inform on the most accurate targeting of gene editing methodologies such as CRISPR-based editing.

As spatial transcriptomic technologies reach single-cell resolution they hold a great deal of promise for study of cells in tissues with complex morphology, such as regions of the brain and endometrium. The field is also tackling the challenge of single-cell spatial transcriptomics of whole organs and this will be hugely beneficial for the understanding of fundamental processes such as development. Increasing development of technologies compatible with FFPE tissue positions these methods as next generation diagnostic and pathology tools.

The surge of success in single-cell technologies has developed through the combination of academic research together with commercial initiatives to make robust, scalable and affordable technologies. Underpinning this has been the continual development of computational methods for data processing, analysis

and integration. Happy tenth birthday single-cell transcriptomics! May the next decade be even more exciting than the last!

Received: 30 April 2020; Accepted: 4 August 2020;  
Published online: 27 August 2020

## References

1. Eberwine, J. et al. Analysis of gene expression in single live neurons. *Proc. Natl Acad. Sci. USA* **89**, 3010–3014 (1992).
2. Lamboldz, B. et al. AMPA receptor subunits expressed by single purkinje cells. *Neuron* **9**, 247–258 (1992).
3. Tang, F. et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377–382 (2009).
4. Svensson, V., Vento-Tormo, R. & Teichmann, S. A. Exponential scaling of single-cell RNA-seq in the past decade. *Nat. Protoc.* **13**, 599–604 (2018).
5. Efreanova, M. & Teichmann, S. A. Computational methods for single-cell omics across modalities. *Nat. Methods* **17**, 14–17 (2020).
6. Stegle, O., Teichmann, S. A. & Marioni, J. C. Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* **16**, 133–145 (2015).
7. Tabula Muris Consortium. et al. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* **562**, 367–372 (2018).

8. Xiaoping, H. et al. Mapping the mouse cell atlas by microwell-seq. *Cell* **172**, 1091–1107 (2018).
9. Codeluppi, S. et al. Spatial organization of the somatosensory cortex revealed by osmFISH. *Nat. Methods* **15**, 932–935 (2018).
10. Vento-Tormo, R. et al. Single-cell reconstruction of the early maternal-fetal interface in humans. *Nature* **563**, 347–353 (2018).
11. Park, J.-E. et al. A cell atlas of human thymic development defines T cell repertoire formation. *Science* **367**, eaay3224 <https://doi.org/10.1126/science.aay3224> (2020).
12. Vieira Braga, F. A. et al. A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat. Med.* **25**, 1153–1163 (2019).
13. Montoro, D. T. et al. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* **560**, 319–324 (2018).
14. Teichmann, S. & Regev, A. The network effect: studying COVID-19 pathology with the Human Cell Atlas. *Nat. Rev. Mol. Cell Biol.* <https://doi.org/10.1038/s41580-020-0267-315> (2020).
15. Velasco, S. et al. Individual brain organoids reproducibly form cell diversity of the human cerebral cortex. *Nature* **570**, 523–527 (2019).

### Acknowledgements

We gratefully acknowledge Jana Eliasova for the graphical images and all authors of single-cell transcriptomics papers that we have not been able to cite due to space constraints.

### Author contributions

The authors S.A.T. and S.A. contributed equally to all aspects of this article.

### Competing interests

In the last 3 years, S.A.T. has consulted for Genentech and Roche, and is a member of SABs of Biogen, GlaxoSmithKline and Foresite Labs. The remaining author declares no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to S.A.T.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020