Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Multi-omics at single-cell resolution: comparison of experimental and data fusion approaches

Karolis Leonavicius, Juozas Nainys, Dalius Kuciauskas and Linas Mazutis

Biological samples are inherently heterogeneous and complex. Tackling this complexity requires innovative technological and analytical solutions. Recent advances in high-throughput single-cell isolation and nucleic acid barcoding methods are rapidly changing the technological landscape of biological sciences and now make it possible to measure the (epi) genomic, transcriptomic, or proteomic state of individual cells. In addition, few experimental approaches enable multi-omics measurements of the same cell. However, merging-omics data collected from different experiments remains a considerable challenge. Although several strategies for merging transcriptomics datasets have recently been introduced, cell-to-cell variability and heterogeneity remains one of the confounding factors limiting data fusion and integration. Here, we focus our discussion on the latest single-cell technological and analytical solutions to achieve high data dimensionality and resolution. Obtaining datasets with a wealth of multi-omics information will undoubtedly provide new avenues for researchers to unravel the complexity of biological samples encountered in modern biological research and molecular diagnostics.

Address
Institute of Biotechnology, Life Sciences Center, Vilnius University, Sauletekio Av. 7, Vilnius LT-10257, Lithuania

Corresponding author: Mazutis, Linas (linas.mazutis@bti.vu.lt)

Current Opinion in Genetics & Development 2019, **55**:159–166

This review comes from a themed issue on **Analytical biotechnology**

Edited by **Saulius Klimasauskas** and **Linas Mazutis**

For a complete overview see the Issue and the Editorial

Available online 24th October 2018

https://doi.org/10.1016/j.copbio.2018.09.012

Introduction

Diverse types of molecules compose living systems. These molecules work together interdependently to manifest the form and function of living organisms. Historically, biological analysis has focused on examining selected aspects of the system (e.g. proteins, nucleic acids, metabolites, etc.) and, thus, providing only partial understanding of the biological structures and functions. However, modern tools made ultra high-throughputs possible, which provide the data diversity needed for addressing complex problems. Single-cell techniques enable researchers to analyse thousands of diverse analytes and provide new insights into cellular diversity. For biological samples, which are inherently defined by heterogeneity, the single-cell-omic tools and methods revolutionised exploratory analyses. However, the challenge for higher quality data and diversity remains, particularly in the context of complex clinical samples.

The motivation for integrating the different types of -omics information is clear—by simultaneously understanding the different layers of biological information we can arrive at a more detailed and systematic understanding the fundamental principles of living systems [1–3]. Even with integrated multi-omic information, sample heterogeneity and variability still pose major problems, hence the ultimate goal has become to achieve multi-omic datasets at the resolution of the single-cell [4•,5,6]. Merging the information from different-omics experiments and techniques into valid datasets presents one of the most exciting frontiers in molecular biology. Such a goal is embodied in large scale initiatives, like the Human cell Atlas [7•], Mouse cell Atlas [8], and the BRAIN initiative [9], as well as precision and personalised medicine [10]. To understand the emerging fundamental shift towards the multi-omic approaches, herein we discuss the engineering and molecular biology breakthroughs that led to the current modern single-cell, single-omic methods used for genomics [11], transcriptomics [12], epigenomics [13] and proteomics [14].

Fluorescence and mass spectrometry-based single-cell techniques

Fluorescence-based microscopy and flow cytometry are the most common techniques used in single cell biology research [15]. Particularly, fluorescence activated cell sorting (FACS) is among the very few methods that can sort and enrich desirable cells based on specific, typically antibody-based, staining. As a result, marrying FACS instruments with 96-well plates and biochemical protocols produced a variety of useful single-cell sequencing methods [16].

Although fluorescence-based tools were the first to achieve single-cell resolution, they were limited to only a few molecular targets that could be analysed. Modern

mass-spectrometry has evolved to solve this problem by using heavy metal isotope tagged antibodies (CyTOF [14] and Multiplexed Ion Beam Imaging (MIBI) [17,18]). MIBI also enables one to simultaneously analyse large panels of protein and mRNA molecules at the single-cell level and relate the acquired information to spatial cell context within tissues [19].

Single-cell genomics, transcriptomics and epigenomics

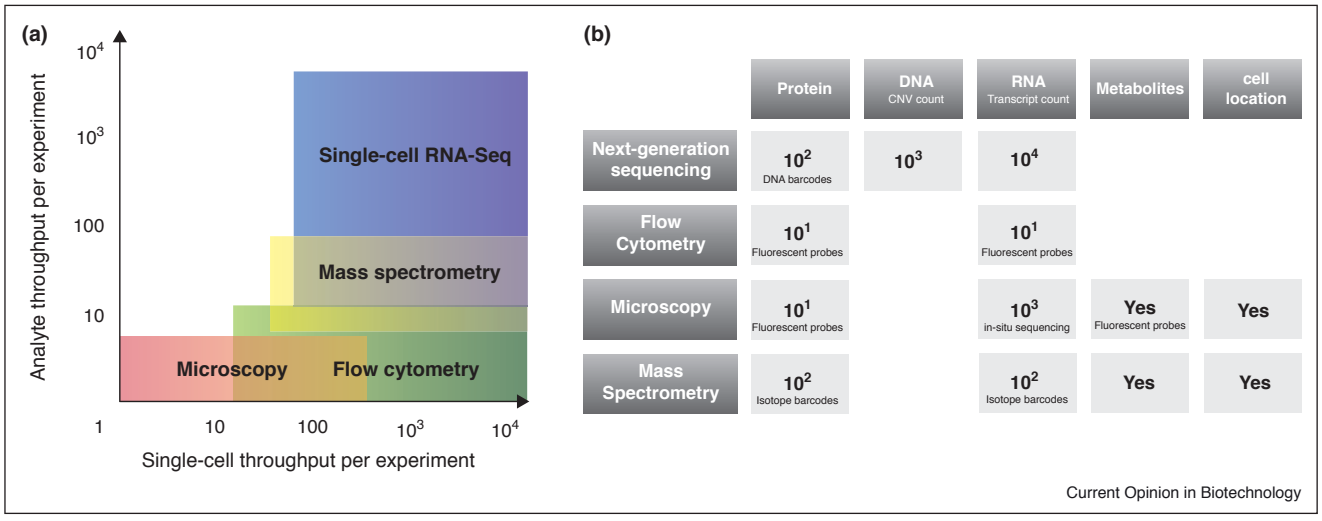
Single-cell genomics has evolved from two distinct perspectives [9]. Historically, one major use was the sequencing of unicellular and diverse organisms [20,21•], while the second attempted to identify genetic mosaicism of cells composing multicellular organisms. Because of the large size of chromosomes, the genome-wide studies are currently feasible on only a small number of individual cells [22]. The main challenge remains in achieving an unbiased and uniform whole genome amplification (WGA) from vanishingly small amounts of genetic material. Several interesting solutions to this problem, such as DOP-PCR [23,24], MDA [25], MALBAC [26] and more recently LIANTI [27], and others, might prove to be crucial in advancing the field of single-cell genomics. Yet, despite these advances the most feasible option, at least economically, for studying large sets of cells continue to be targeted sequencing of single nucleotide variants (SNV) and copy number variants (CNV) [28].

The development of single-cell DNA sequencing ultimately led to single-cell transcriptomics, which has arguably become one of the most active and productive fields. High-throughput single-cell RNA sequencing (scRNA-Seq) techniques have seen breakneck growth since they first appeared almost a decade ago [29] and are now broadly accessible. Innovations in droplet microfluidics [30•,31•] and nano-well systems [32] have significantly reduced the cost of scRNA-Seq approaches facilitating widespread adoption. In a nutshell, modern scRNA-Seq methods rely on mRNA capture and cDNA synthesis using barcoded poly(dT) oligonucleotides, followed by next-generation sequencing [33]. The barcode information is then used to assign cDNA molecules to individual cells and build their transcriptomic profiles. The main advantage of scRNA-Seq lies in its unbiased and, more importantly, *de novo* identification of cell states and types based on their gene expression profiles, rather than DNA sequence [34–36]. The utility of single-cell transcriptomics almost guarantees its wide and long-term adoption. However, the current method complexity and cost may be reduced as exemplified by direct cell split-pool barcoding [37•].

In contrast to the direct analysis of genetic information, epigenetic sequencing has proven more difficult to perform at the single-cell level. Yet, the data is particularly useful because of its diversity [13]. DNA methylation studies have been the most facile because of well-establish techniques used to identify epigenetic marks such as 5-methylcytosine (5mC) [38], 5-hydroxymethylcytosine (5hmC) [39], 5-formylcytosine (5fC) [40]. Other protocols, based on DNase I [41] or Tn5 transposase [42], provide diverse information on open chromatin.

Individual single-cell technologies have distinct advantages and disadvantages in terms of throughput and analyte types (Figure 1). Moreover, most techniques were designed with a single, specific analyte type in mind (protein, RNA, DNA, etc.). Therefore, obtaining

Figure 1



The throughput landscape of the latest single-cell analytical methods.

single-cell multi-omic data requires new innovative solutions. To address this need, two novel, broad strategies have emerged that are discussed in the following section.

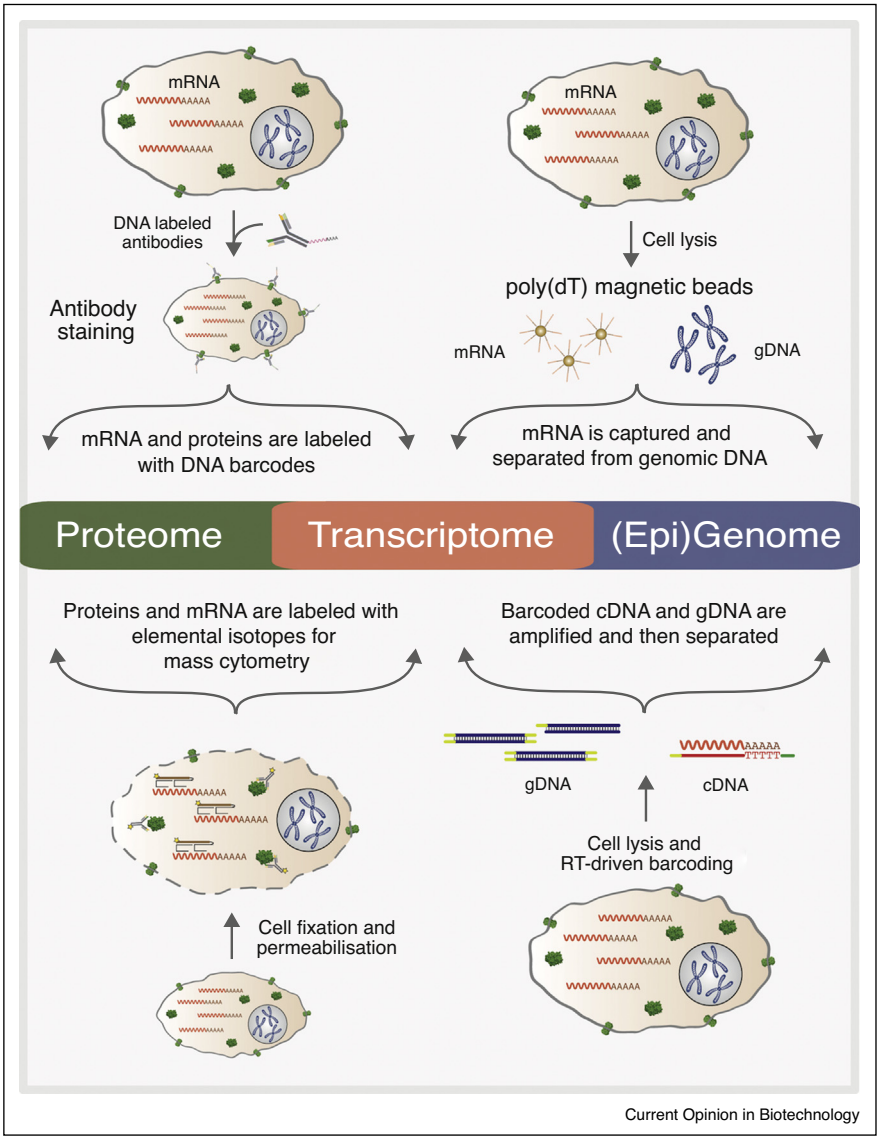
Single-cell multi-omics

Currently there are two dominant approaches for multi-omics research at the single cell level. The 'wet approach' aims to develop new, or modify existing, single-cell techniques to incorporate diverse molecular analytes for simultaneous capture and analysis (Figure 2 and Table 1). The statistical approach, in turn, focuses on analysis workflows to integrate separate experiments into single-cell multi-omics datasets (Figure 3 and Table 2).

Multiple measurements from individual cells

Naturally, the first experiments that achieved multi-omic measurements were based on the same underlying

Figure 2



Brief overview of current experimental strategies for generating single-cell multi-omics datasets.
In REAP-Seq and CITE-Seq approaches (top left) the simultaneous capture of whole transcriptome and a panel of proteins is performed by tagging mRNA and Ab-DNA bound proteins with barcoded poly(dT) primers during reverse transcription step. Alternatively, mRNA and selected proteins can be labelled with elemental isotopes and recorded by mass spectrometry (bottom left). Simultaneous transcriptome and (epi)genome measurements can be performed by separating the cell cytoplasm and nuclei fractions, and then barcoding nucleic acids of each fraction separately as exemplified in scMT-Seq and G&T-Seq methods (top right). Alternatively, in DR-Seq approach the reverse transcription is used to barcode the transcriptome followed by quasi-linear amplification to capture and amplify the genomic DNA and cDNA of the same cells (bottom right).

Table 1

Overview of existing multi-omic techniques focusing on combined cell library analysis

	Types of analytes	Throughput	Infrastructure requirements	Original application
G&T-Seq	**Genome/transcriptome**	172 Cells, 4k–11k transcripts per cell	Next-gen sequencer, cell-sorting and isolation capabilities, barcoded beads	Relating DNA CNV and chromosomal abnormalities to transcriptional differences
DR-seq	**Genome/transcriptome**	33 Cells, 9k transcripts per cell	Next-gen sequencer, cell-sorting and isolation capabilities	Demonstrating that CNVs can drive transcriptional variability among cancer cells
scM&T-seq	**Methylome/transcriptome**	76 and 16 embryonic stem cells. 4k–8k transcripts per cell	Next-gen sequencer, cell-sorting and isolation capabilities, barcoded beads	Demonstrated, that methylation negatively controls expression of Esrrb transcription factor
sc-GEM	**Genotyping methylome/transcriptome**	48-96 cells/ chip (Fluidigm)	Next-gen sequencer, microfluidic cell capture, isolation and processing equipment	Investigated links between differential promoter methylation, transcription, and ES cell pluripotency
scTrio-seq	**Copy number variations methylome/transcriptome**	25 Cells	Next-gen sequencer	Identifying two cancer cell subpopulations based on collective multi-omic information
PEA	**Protein/transcript information**	96 Cells, 38 proteins, 96 qPCR assays	Microfluidic individual cell handling system	Investigating combined proteomic and transcriptomic response of adenocarcinoma to chemical perturbation
REAP-SEQ	**Proteomics/transcriptomics**	7k cells, 20k transcripts, 82 proteins	Next-gen sequencer, droplet microfluidics systems, barcoded beads	Identifying and characterising an unknown cell type
PLAYR	**Proteomics/transcriptomics**	10k cells, 40 transcripts and proteins	Mass cytometry or fluorescence activated cell sorting instruments	Characterising the relationship between transcription and translation at the single cell level

technique (e.g. next-generation sequencing). For example, using physical separation of cell cytoplasm and nuclear fractions into separate genetically barcoded well plates it becomes possible to analyse multiple types of analytes. This strategy is used to combine single-cell epigenomic and transcriptomic measurements [43], as well as genomic and transcriptomic information [44]. Another approach named DR-Seq also enables simultaneous genomic and transcriptomic analysis of individual cells, but without using physical separation of the mRNA on poly-T beads [45]. More recently, integrated multi-omics approaches enabling genotyping, CNVs, methylome as well as transcriptome analysis at single-cell resolution have also been reported [46]. These methods open exciting possibilities to investigate functional consequences of *de novo* structural variant formation, DNA rearrangements, or epigenetic marks of the genome and correlate these changes to transcriptional cell state.

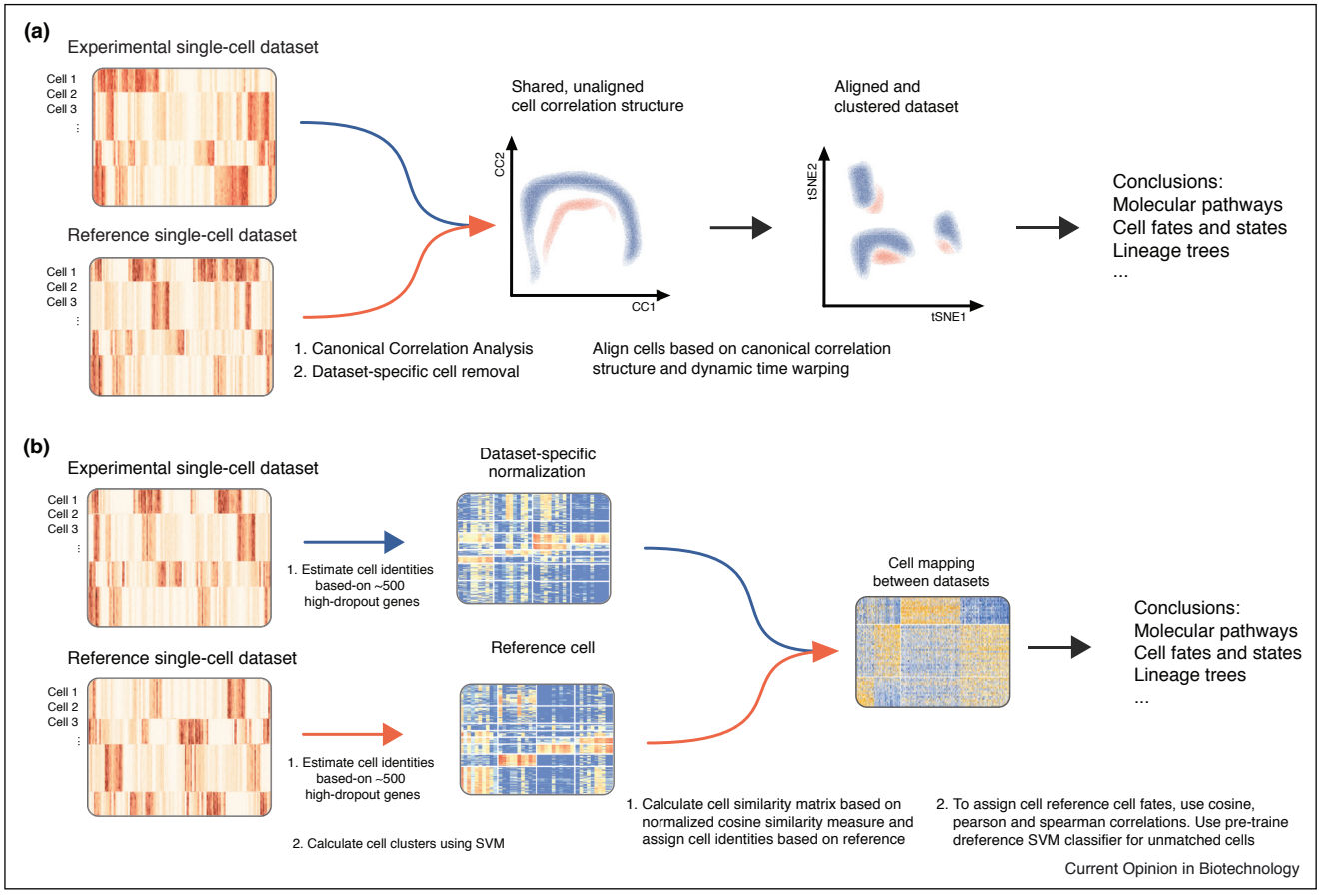
On the other hand, they also suffer from biases, such as mis-incorporation of bases during combined RNA and DNA amplification, limited sensitivity of immune precipitation or tagging of single-cell chromatin. To date, these techniques have been limited to analysis of a few dozen cells. However, as these and other multi-omics techniques are continuously being improved we are likely to witness their wider use to investigate how the genetic and epigenetic factors regulate gene expression in individual cells.

Given the rapid rise of nucleic acid analytical techniques, single cell proteomics has also explored the next-generation sequencing capabilities by taking advantage of protein tagging (labelling) approaches. The most successful attempts to integrate protein-level information with transcriptomic data into single-cell multi-omics datasets have relied on Antibody-DNA (Ab-DNA) conjugates to analyse either intracellular or membrane-associated proteins. To quantify intracellular protein levels Ab-DNA conjugates were used in combination with proximity extension assay and qPCR [47]. To quantify surface proteins, the cells treated with Ab-DNA conjugates can be directly subjected to the reverse transcription reaction, which will simultaneously extend the conjugated DNA primer on Ab with the poly(dT) primers and synthesise complementary DNA from mRNA [48•]. In combination with droplet microfluidics technology, the Ab-DNA approach can reach the throughput of thousands of single-cells. Furthermore, considering the commercial availability of Ab-DNA reagents and their compatibility with different technological platforms protein tagging with Ab-DNA will likely to play a larger role in immunology and cancer research that traditionally relied on flow cytometry.

Data fusion of diverse single-cell datasets

The development of single-cell techniques, and particularly high throughput RNA-seq, created new and

Figure 3



Contrasting two most prominent computational data fusion techniques.
(a) The approach assumes that two corresponding experiments share the same underlying lower dimensional data structure. It then attempts to identify it by considering a canonical correlation analysis. This common data structure can then be used to align the cells among two different experiments and enables a direct comparison between cell types and states, which otherwise would not be possible to achieve due to experimental variability [50•]. **(b)** The second approach relies on statistical cell identification using 500 highest dropout genes, which unique represent cell types. It then clusters the cells and uses cosine similarity measure to find the closest matching cell in a reference dataset [49•], thereby avoiding the underlying biases from different experimental conditions.

interesting opportunities for data integration as well as new challenges (Figure 3). While traditional 'omic' techniques measure the levels of thousands of analytes, single cell datasets contain a lot of missing information because of the diluted analysis depth. Integrating zero-inflated data from different experiments poses major computational difficulties even if the experiments are for the same -omics technique. Integrating scRNA-Seq datasets from isolated experiments poses a major challenge for international collaborative projects, such as the atlases for Human [7•] and Mouse [8] cells. The challenge of data integration has been taken up by a couple of innovative computational approaches: sc-map [49•] and Seurat [50•] to name a few. The methods typically focus on dimensionality reduction to internally normalise different datasets uncovering the underlying low-dimensional manifolds [51], which are conserved in biological samples. The general concept typically uses a form of data imputation followed by dimensionality reduction [52] or pseudo-time mapping [53].

Matching data at the single cell level from multiple omics experiments is only starting to be put into practice, as more diverse and high-quality datasets become available for algorithm development. MATCHER alignment framework [54•] for example, demonstrated how diverse data from multi-omic scM&T-seq and sc-GEM experiments could be integrated without prior knowledge of cell identities. The tool used common modes of variation between epigenome and transcriptome data to create an equivalent pseudo-time representation between the two datasets, which could be used for data alignment. This idea could potentially be adapted towards more universal multi-omic data analysis techniques. It would

Table 2

Overview of existing multi-omic techniques based on data fusion from isolated experiments

	Types of analytes	Principle	Performance	Original applications
sc-map	**Transcriptome/ transcriptome**	Unsupervised transcriptome clustering	75–95% cell alignment accuracy	scRNA-Seq dataset merging
Seurat	**Transcriptome/ transcriptome**	Diagonal canonical correlation analysis	75–95% cell alignment success rate	scRNA-Seq dataset merging
MATCHER	**Epigenome/ transcriptome**	Manifold alignment	Deviation between true and observed correlations 0.16	Transcription and chromatin state changes in mESCs

rely on finding common modes of variation among the individual-omics datasets first, which would them to be merged and generalised into multi-omics datasets (Figure 2).

Such hierarchical analysis presents the most generalisable strategy for combining diverse datasets based on single cells. While some of the most prominent current bulk multi-omic data merging strategies are based on network and biological pathway analysis [55], single-cell datasets also present an opportunity to merge data based on approximated cell identities [49•], or underlying a shared gene correlation structure [50•]. Altogether, combining existing multi-omic data merging strategies with single cell matching techniques presents exciting opportunities for discovering new biological phenomena without lengthy integrated experimental strategy development. The computational methods for data fusion are still in their infancy, and we are only beginning to build the relationships between different analyte types.

Conclusions

Modern single-cell techniques and molecular biology tools present a fundamental shift in how biological complexity and heterogeneity can be approached and resolved. This, in turn, further drives the need for better experimental and analytical approaches, which are naturally turning towards integrated multi-omic analyses. So far, two distinct approaches are starting to crystalise. The first is based on fusing different single-omics techniques to experimentally generate single-cell derived multi-omics data sets. Predominantly, it takes advantage of the nucleic acid barcodes to label the biomolecules of interest (e.g. RNA, DNA or proteins) followed by next-gen DNA sequencing. The second, fuses individual single-cell datasets acquired from separate experiments into one multi-omics atlas. However, even for scRNA-Seq experiments, data integration has been a long standing challenge, and researchers are only beginning to find solutions [49•,50•]. Creating similar tools for data integration across multi-omic experiments presents an even greater challenge, because of the requirement for preliminary understanding of the underlying biological relationships between different types of analytes. As the result, the first such datasets will most likely come from experiments with multi-omic measurements on single cells, which would make it possible to uncover biological relationships between diverse molecule types at single-cell resolution. However, because their lower cost and experimental complexity, computational data fusion approaches will probably take the lead eventually for providing multi-omic datasets from standardised single-cell analysis techniques. As the Human [7•] and Mouse [8] cell atlases continue to generate high-quality transcriptomic data and single-cell proteomic and epigenomic information increasingly become available, the need for experimental and data fusion approaches for comparing cell types and states will only grow in demand and relevance.

Conflict of interest statement

Nothing declared.

Acknowledgements

Funding: this work was supported by the European Social Fund according to the activity 'Improvement of researchers' qualification by implementing world-class R&D projects' of Measure [grant number 09.3.3-LMT-K-712-01-0056. K.L. is supported by a Research Council of Lithuania postdoctoral award [grant number 09.3.3-LMT-K-712-02-0067].

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest

1. Higdon R, Earl RK, Stanberry L, Hudac CM, Montague E, Stewart E, Janko I, Choiniere J, Broomall W, Kolker N *et al.*: **The promise of multi-omics and clinical data integration to identify and target personalized healthcare approaches in autism spectrum disorders**. *OMICS* 2015, **19**:197-208.

2. Argelaguet R, Velten B, Arnol D, Dietrich S, Zenz T, Marioni JC, Huber W, Buettner F, Stegle O: **Multi-omics factor analysis—a framework for unsupervised integration of multi-omic data sets**. *Mol Syst Biol* 2018, **14**:e8124.

3. Tini G, Marchetti L, Priami C, Scott-Boyer M-P: **Multi-omics integration—a comparison of unsupervised clustering methodologies**. *Brief Bioinf* 2017, **167**:1-11 http://dx.doi.org/10.1093/bib/bbx167.

4. • Macaulay IC, Ponting CP, Voet T: **Single-cell multiomics: multiple measurements from single cells**. *Trends Genet* 2017, **33**:155-168.

A comprehensive review on how the development of single-cell analysis techniques is leading towards multiple 'omic' measurements and development of 'Omniscience' analysis.

5. Bock C, Farlik M, Sheffield NC: **Multi-omics of single cells: strategies and applications**. *Trends Biotechnol* 2016, **34**:605-608.

6. Colomé-Tatché M, Theis FJ: **Statistical single cell multi-omics integration**. *Curr Opin Syst Biol* 2018, **7**:54-59.

7. • Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, Bodenmiller B, Campbell P, Carninci P, Clatworthy M *et al.*: **The human cell atlas**. *eLife* 2017, **6**:1-30.
The Human Cell Atlas is one of the most exciting global scientific projects and this is among the best examples of comprehensive single-cell organ-specific datasets containing diverse information, which could be used for future reference.

8. Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, Saadatpour A, Zhou Z, Chen H, Ye F *et al.*: **Mapping the mouse cell atlas by microwell-seq**. *Cell* 2018, **172**:1091-1097.e17.

9. Alivisatos AP, Chun M, Church GM, Greenspan RJ, Roukes ML, Yuste R: **The brain activity map project and the challenge of functional connectomics**. *Neuron* 2012, **74**:970-974.

10. Collins FS, Varmus H: **A new initiative on precision medicine**. *NEJM* 2015, **372**:793-795.

11. Gawad C, Koh W, Quake SR: **Single-cell genome sequencing: current state of the science**. *Nat Rev Genet* 2016, **17**:175-188.

12. Ziegenhain C, Vieth B, Parekh S, Hellmann I, Enard W: **Quantitative single-cell transcriptomics**. *Brief Funct Genomics* 2018, **17**:220-232 http://dx.doi.org/10.1093/bfgp/ely009.

13. Kelsey G, Stegle O, Reik W: **Single-cell epigenomics: recording the past and predicting the future**. *Science* 2017, **358**:69-75.

14. Spitzer MH, Nolan GP: **Mass cytometry: single cells, many features**. *Cell* 2016, **165**:780-791.

15. Adan A, Alizada G, Kiraz Y, Baran Y, Nalbant A: **Flow cytometry: basic principles and applications**. *Crit Rev Biotechnol* 2017, **37**:163-176.

16. Baran-Gale J, Chandra T, Kirschner K: **Experimental design for single-cell RNA sequencing**. *Brief Funct Genomics* 2017, **17**:233-239 http://dx.doi.org/10.1093/bfgp/elx035.

17. Angelo M, Bendall SC, Finck R, Hale MB, Hitzman C, Borowsky AD, Levenson RM, Lowe JB, Liu SD, Zhao S *et al.*: **Multiplexed ion beam imaging of human breast tumors**. *Nat Med* 2014, **20**:436-442.

18. Schulz D, Zanotelli VRT, Fischer JR, Schapiro D, Engler S, Lun XK, Jackson HW, Bodenmiller B: **Simultaneous multiplexed imaging of mRNA and proteins with subcellular resolution in breast cancer tissue samples by mass cytometry**. *Cell Syst* 2018, **6**:25-36.e5.

19. Bendall SC, Simonds EF, Qiu P, Amir E-aD, Krutzik PO, Finck R, Bruggner RV, Melamed R, Trejo A, Ornatsky OI *et al.*: **Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum**. *Science* 2011, **332**:687-696.

20. Lasken RS, McLean JS: **Recent advances in genomic DNA sequencing of microbial species from single cells**. *Nat Rev Genet* 2014, **15**:577-584.

21. • Lan F, Demaree B, Ahmed N, Abate AR: **Single-cell genome sequencing at ultra-high-throughput with microfluidic droplet barcoding**. *Nat Biotechnol* 2017, **35**:640-646.
The ultra high-throughput analysis of bacterial genomes at shalow sequencing depth.

22. Lodato MA, Woodworth MB, Lee S, Evrony GD, Mehta BK, Karger A, Lee S, Chittenden TW, D'Gama AM, Cai X *et al.*: **Somatic mutation in single human neurons tracks developmental and transcriptional history**. *Science* 2015, **350**:94-98.

23. Telenius H, Carter NP, Bebb CE, Nordenskjöld M, Ponder BAJ, Tunnacliffe A: **Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer**. *Genomics* 1992, **13**:718-725.

24. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D *et al.*: **Tumour evolution inferred by single-cell sequencing**. *Nature* 2011, **472**:90-95.

25. Fu Y, Li C, Lu S, Zhou W, Tang F, Xie XS, Huang Y: **Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification**. *Proc Natl Acad Sci U S A* 2015, **112**:11923-11928.

26. Zong C, Lu S, Chapman AR, Xie X: **Genome-wide detection of single-nucleotide and copy-numbernvariations of a single human cell**. *Science* 2012, **338**:1622-1626.

27. Chen C, Xing D, Tan L, Li H, Zhou G, Huang L, Xie XS: **Single-cell whole-genome analyses by Linear Amplification via Transposon Insertion (LIANTI)**. *Science* 2017, **356**:189-194.

28. Potter NE, Ermini L, Papaemmanuil E, Cazzaniga G, Vijayaraghavan G, Titley I, Ford A, Campbell P, Kearney L, Greaves M: **Single-cell mutational profiling and clonal phylogeny in cancer**. *Genome Res* 2013, **23**:2115-2125.

29. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A *et al.*: **mRNA-Seq whole-transcriptome analysis of a single cell**. *Nat Methods* 2009, **6**:377-382.

30. • Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, Peshkin L, Weitz DA, Kirschner MW: **Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells**. *Cell* 2015, **161**:1187-1201.
A foundational primary research paper that led to the development of modern high-throughput single cell RNA sequencing protocols and development of commercial platforms relying on droplet microfluidics.

31. • Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM *et al.*: **Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets**. *Cell* 2015, **161**:1202-1214.
Another research article describing high-throughput single cell RNA sequencing using droplet microfluidics.

32. Gierahn TM, Wadsworth MH, Hughes TK, Bryson BD, Butler A, Satija R, Fortune S, Christopher Love J, Shalek AK: **Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput**. *Nat Methods* 2017, **14**:395-398.

33. Svensson V, Natarajan KN, Ly LH, Miragaia RJ, Labalette C, Macaulay IC, Cvejic A, Teichmann SA: **Power analysis of single-cell RNA-sequencing experiments**. *Nat Methods* 2017, **14**:381-387.

34. Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA: **The technology and biology of single-cell RNA sequencing**. *Mol Cell* 2015, **58**:610-620.

35. Villani A-C, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, Griesbeck M, Butler A, Zheng S, Lazo S *et al.*: **Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors**. *Science* 2017, **356** eaah4573.

36. Fiers MWEJ, Minnoye L, Aibar S, Bravo González-Blas C, Kalender Atak Z, Aerts S: **Mapping gene regulatory networks from single-cell omics data**. *Brief Funct Genomics* 2018, **17**:246-254 http://dx.doi.org/10.1093/bfgp/elx046.

37. • Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, Graybuck LT, Peeler DJ, Mukherjee S, Chen W *et al.*: **Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding**. *Science* 2018, **360**:176-182.
The works presents an exciting technique that enables processing large numbers of cells using split-and-pool approach. The method can potentially reduce the price of scRNA-Seq, making it more widespread and economical for clinical use.

38. Mulqueen RM, Pokholok D, Norberg S, Fields AJ, Sun D, Torkenczy KA, Shendure J, Trapnell C, Roak BJO, Xia Z *et al.*: **Highly scalable generation of DNA methylation profiles in single cells**. *Nat Biotechnol* 2018, **36**:428-431.

39. Mooijman D, Dey SS, Boisset J-C, Crosetto N, van Oudenaarden A: **Single-cell 5hmC sequencing reveals chromosome-wide cell-to-cell variability and enables lineage reconstruction**. *Nat Biotechnol* 2016, **34**:852-856.

40. Zhu C, Gao Y, Guo H, Xia B, Song J, Wu X, Zeng H, Kee K, Tang F, Yi C: **Single-cell 5-formylcytosine landscapes of mammalian early embryos and ESCs at single-base resolution**. *Cell Stem Cell* 2017, **20**:720-731.e5.

41. Jin W, Tang Q, Wan M, Cui K, Zhang Y, Ren G, Ni B, Sklar J, Przytycka TM, Childs R *et al.*: **Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples**. *Nature* 2015, **528**:142-146.

42. Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, Steemers FJ, Trapnell C, Shendure J: **Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing**. *Science* 2015, **348**:910-914.

43. Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, Krueger F, Smallwood SA, Ponting CP, Voet T *et al.*: **Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity**. *Nat Methods* 2016, **13**:229-232.

44. Macaulay IC, Haerty W, Kumar P, Li YI, Hu TX, Teng MJ, Goolam M, Saurat N, Coupland P, Shirley LM *et al.*: **G&T-seq: parallel sequencing of single-cell genomes and transcriptomes**. *Nat Methods* 2015, **12**:519-522.

45. Dey SS, Kester L, Spanjaard B, Bienko M, van Oudenaarden A: **Integrated genome and transcriptome sequencing of the same cell**. *Nat Biotechnol* 2015, **33**:285-289.

46. Hou Y, Guo H, Cao C, Li X, Hu B, Zhu P, Wu X, Wen L, Tang F, Huang Y *et al.*: **Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas**. *Cell Res* 2016, **26**:304-319.

47. Genshaft AS, Li S, Gallant CJ, Darmanis S, Prakadan SM, Ziegler CGK, Lundberg M, Fredriksson S, Hong J, Regev A *et al.*: **Multiplexed, targeted profiling of single-cell proteomes and transcriptomes in a single reaction**. *Genome Biol* 2016, **17**:1-15.

48. • Peterson VM, Zhang KX, Kumar N, Wong J, Li L, Wilson DC, Moore R, Mcclanahan TK, Sadekova S, Klappenbach JA: **Multiplexed quantification of proteins and transcripts in single cells**. *Nat Biotechnol* 2017, **35**:936-939.
One of the primary examples demonstrating how multiomic information could be integrated into diverse datasets. The method uses DNA primer—antibody conjugates to tag proteins of interest with DNA barcodes for subsequent analysis during scRNA-Seq.

49. • Kiselev VY, Yiu A, Hemberg M: **scmap: projection of single-cell RNA-seq data across data sets**. *Nat Methods* 2018, **15**:359-362.
An innovative data fusion approach for merging single cell RNA seq datasets using cluster projection. This allows to align cells between different experiments making it possible to use reference data from large scale initiatives such as Human Cell Atlas.

50. • Butler A, Hoffman P, Smibert P, Papalexi E, Satija R: **Integrating single-cell transcriptomic data across different conditions, technologies, and species**. *Nat Biotechnol* 2018, **36**:411-420.
The interesting report focusing on data fusion methods build on the same type of information (transcriptomes) across unrelated experiments. Although this is not a multi-omic integration tool, this remains one of the best examples of successful data fusion.

51. Moon KR, Stanley JS, Burkhardt D, van Dijk Dv, Wolf G, Krishnaswamy S: **Manifold learning-based methods for analyzing single-cell RNA-sequencing data**. *Curr Opin Syst Biol* 2018, **7**:36-46.

52. Van Der Maaten LJP, Hinton GE: **Visualizing high-dimensional data using t-sne**. *J Mach Learn Res* 2008, **9**:2579-2605.

53. Haghverdi L, Büttner M, Wolf FA, Buettner F, Theis FJ: **Diffusion pseudotime robustly reconstructs lineage branching**. *Nat Methods* 2016, **13**:845-848.

54. • Welch JD, Hartemink AJ, Prins JF: **MATCHER: manifold alignment reveals correspondence between single cell transcriptome and epigenome dynamics**. *Genome Biol* 2017, **18**:1-19.
One of the first methods for integrating data from multi-omic single-cell experiments. Successfully aligned transcriptomic and epigenomic data based on common underlying manifolds. This made it possible to look at how transcription is controlled by epigenetic.

55. Bersanelli M, Mosca E, Remondini D, Giampieri E, Sala C, Castellani G, Milanesi L: **Methods for the integration of multi-omics data: mathematical aspects**. *BMC Bioinf* 2016, **17(Suppl. 2)**:S15.