



REVIEW

Decoding Human Biology and Disease Using Single-cell Omics Technologies



Qiang Shi^{1,#}, Xueyan Chen^{1,#}, Zemin Zhang^{1,2,3,*}

¹ Biomedical Pioneering Innovation Center, School of Life Sciences, Peking University, Beijing 100871, China

² Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

³ Changping Laboratory, Beijing 102206, China

Received 19 January 2023; revised 22 May 2023; accepted 8 June 2023

Available online 20 September 2023

Handled by Xiang Chen

KEYWORDS

Single-cell omics;
 Computational method;
 Cellular heterogeneity;
 Disease;
 Cancer research

Abstract Over the past decade, advances in **single-cell omics** (SCO) technologies have enabled the investigation of **cellular heterogeneity** at an unprecedented resolution and scale, opening a new avenue for understanding human biology and **disease**. In this review, we summarize the developments of sequencing-based SCO technologies and **computational methods**, and focus on considerable insights acquired from SCO sequencing studies to understand normal and diseased properties, with a particular emphasis on **cancer research**. We also discuss the technological improvements of SCO and its possible contribution to fundamental research of the human, as well as its great potential in clinical diagnoses and personalized therapies of human disease.

Introduction

The cell acts as the fundamental unit of life. A single zygote gives rise to the entire human body, in which approximately 37 trillion cells of diverse types are highly orchestrated into a variety of tissues, organs, and systems. Traditionally, distinct cell types have been defined in terms of cellular morphology, location, or expression levels of a small number of proteins,

which grossly neglects differences in additional molecular layers across cells within a population. The large heterogeneity of cells underlies functional diversity in human biology [1]. Notably, the characteristics of a cell are pertinent not only to its own state, size, or ancestor, but also to its unique niche around and how the cell interacts with adjacent or distant cells [2,3]. For example, although found in most organs to synthesize the extracellular matrix of connective tissue by producing collagen, fibroblasts perform specialized functions depending on the specific contexts across a broad range of tissues and disease conditions [4]. However, conventional bulk sample sequencing technologies mask the diversity of cells, as exemplified by RNA sequencing, which derives average measurements of gene expression for all cells within an experimental sample [5]. Therefore, understanding human biology and disease at single-cell resolution is imperative.

* Corresponding author.

E-mail: zemin@pku.edu.cn (Zhang Z).

Equal contribution.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformatics and Genetics Society of China.

<https://doi.org/10.1016/j.gpb.2023.06.003>

1672-0229 © 2023 The Authors. Published by Elsevier B.V. and Science Press on behalf of Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformatics and Genetics Society of China.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Since the first single-cell RNA sequencing (scRNA-seq) method arising in 2009 [6], numerous single-cell omics (SCO) sequencing technologies have been developed to characterize cellular properties at different molecular layers, including the genome, epigenome, transcriptome, and proteome [6]. Single-cell multimodal omics sequencing represents the state-of-the-art technology, which can simultaneously depict multiple characteristics of one cell [7]. The revolution in SCO sequencing technology has dramatically expanded our toolbox for investigating biomedical systems in which cells develop along their fates, transition between different states, vary across individuals, and fail in disease [8,9]. Importantly, SCO sequencing has given rise to high-throughput measurements of linkages between intrinsic genotypes and extrinsic phenotypes at the cellular, tissue, organ, and individual levels [10]. These advances have led to many significant insights in the fields of cancer, development, immunity, regenerative medicine, and plant research. Because of its rapid development and enormous potential, SCO sequencing has twice been selected as the Method of the Year by the journal *Nature Methods* [11,12].

In this review, we will summarize the developments of SCO sequencing technologies and computational tools, and highlight the representative knowledge brought by SCO sequencing, especially in cancer research. Finally, we will provide concrete prospects for SCO technologies in fundamental research and clinical applications over the next few years.

Development of SCO sequencing technologies

A primary purpose of SCO sequencing technologies is to disentangle the tremendous cell-to-cell heterogeneity driven by intrinsic programs and extrinsic factors. Essentially, all SCO sequencing technologies aim to decode underlying information surrounding the DNA, RNA, and protein that are the core molecules in the genetic central dogma of molecular biology [13] (**Figure 1**). In this section, we briefly review the rationales of SCO sequencing technologies at different molecular layers, as their technical details have been recently and elaborately reviewed [14].

Genetic variants detected by the single-cell genome sequencing

First of all, genomic DNA sequences represent the genetic information of an organism. Diploid cells have only two copies of genomic DNA, posing the main challenge for single-cell whole-genome sequencing. In the last decade, several methods including MALBAC [15], eMDA [16], LIANTI [17], SISSOR [18], and META-CS [19], have been established to improve the uniformity, efficiency, accuracy, and coverage of the whole genome amplification at single-cell resolution. Although most cells contain the same two copies of genomic DNA, plentiful genetic variants of different types, such as single nucleotide variants (SNVs), small indels, copy number variations (CNVs), and structural variations (SVs) caused by stochastic mutations, emerge in low frequency yet gradually accumulate during the development, aging, as well as disease progression across multiple tissue types of the human body [20–22]. Therefore, these genetic variants detected by single-cell DNA sequencing can be intrinsic markers tracing the historical trajectory of a cell in the human body. In addition, as another part of the cellular genome, mitochondrial DNA (mtDNA) has hundreds of

copies in an individual cell and 10–100-fold higher mutation rates than nuclear DNA. Single-cell genomic assays can be used to detect somatic mtDNA mutations and to track cellular clones via routine sequencing [23]. Recently, the SMOOTH-seq, a novel single-cell genome sequencing method based on the single-molecule real-time DNA sequencing platform, has been reported [24]. Compared with methods based on the short-read sequencing platforms, SMOOTH-seq performs better for the detection of SVs and extrachromosomal circular DNA in individual cells, but shows lower accuracy for CNVs and SNVs.

Revealing the complexity of epigenetic regulations at single-cell resolution

The epigenetic regulation including chromatin states, chromosomal conformations, and DNA or histone modifications, is characterized by heritable variations independent of DNA sequences, representing a crucial molecular mechanism associating the genetic information with its functional output [25]. In eukaryotes, nuclear chromatin is composed of basic repeating structural and functional subunits called nucleosomes, which consist of approximately 146 base pairs of DNA wound around the eight histones. The chromatin accessibility indicating active or repressed states of genomic regions is highly correlated with the dynamics of the gene regulatory network (GRN) [26]. Notably, although only a small fraction (about 2%–3%) of the nuclear genome is kept free from nucleosomes, the accessible genome captures over 90% of regions bound by transcription factors (TFs), and enriches a large number of regulatory DNA elements such as promoters, enhancers, silencers, insulators, and genetic variants associated with diseases [27]. In recent years, multiple SCO technologies have been developed to measure chromatin accessibility by quantifying the susceptibility of chromatin to the cleavage of its constituent DNA via enzymes such as Tn5 transposase [28–33], DNase I [34], or MNase [35]. Among these strategies, single-cell assay for transposase-accessible chromatin using sequencing (ATAC-seq) and its adaptations have been the most widely used, as they leverage hyperactive Tn5 transposons to simultaneously insert, fragment, and add Illumina sequencing adaptors to accessible chromatin regions in individual cells, achieving low cost and high accessibility for users [26].

The three-dimension genomics seeks to figure out how the chromatin of two meters in length is spatially organized into high-order structures within the micron-level nucleus, and how these architectures mediate gene expression modulations by regulatory elements [36]. At present, the main strategies for identifying single-cell genome structures are microscopy and single-cell high-throughput chromosome conformation capture (scHi-C). The microscope is able to detect a broad range of genomic interactions in a single cell, but is generally limited in terms of coverage and overall throughput [37]. In contrast, scHi-C provides the capability to interrogate genome-wide nuclear structures within a cell. In the past decade, several scHi-C methods have been developed to identify the genome-wide interactions by coupling proximity-based ligation followed by massively parallel sequencing [38–45]. Additionally, a method called single-cell SPRITE measures higher-resolution, multiway DNA contacts than that can be achieved by scHi-C [46].

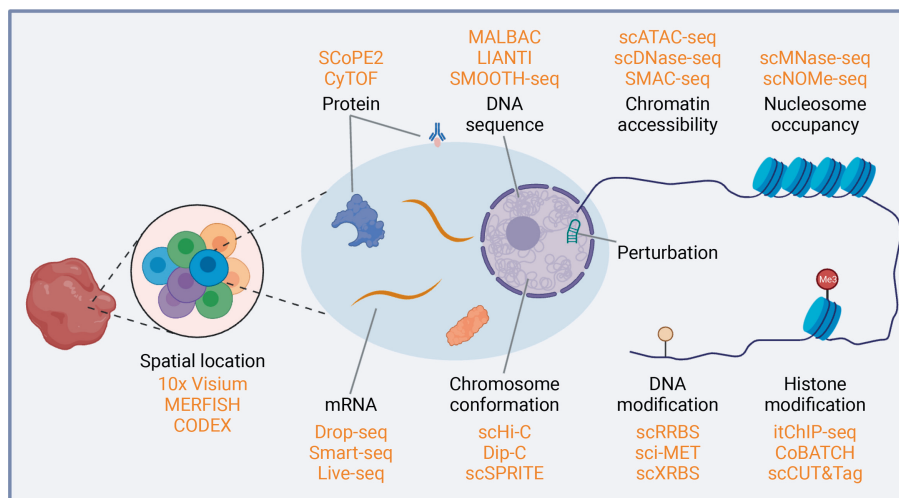


Figure 1 Cellular heterogeneity at different molecular layers

The cell-to-cell heterogeneity is reflected at several distinct molecular layers. Representative methods for profiling each of the individual molecular levels are noted. Single-cell multimodal omics sequencing technologies have been developed to simultaneously profile multiple layers in the same cell. SCoPE, single-cell proteomics; CyTOF, cytometry by time of flight; MALBAC, multiple annealing and looping-based amplification cycles; LIANTI, linear amplification via transposon insertion; SMOOTH-seq, single-molecule real-time sequencing of long fragments amplified through transposon insertion; scATAC-seq, single-cell assay for transposase-accessible chromatin using sequencing; scDNase-seq, single-cell DNase sequencing; SMAC-seq, single-molecule long-read accessible chromatin mapping sequencing; scMNase-seq, single-cell micrococcal nuclease sequencing; scNOME-seq, single-cell nucleosome occupancy and methylome sequencing; MERFISH, multiplexed error robust fluorescence *in situ* hybridization; CODEX, co-detection by indexing; scHi-C, single-cell high-throughput chromosome conformation capture; mRNA, messenger RNA; Smart-seq, switching mechanism at 5' end of the RNA transcript; Dip-C, diploid chromatin conformation capture; scSPRITE, single-cell split-pool recognition of interactions by tag extension; scRRBS, single-cell reduced-representation bisulfite sequencing; sci-MET, single-cell combinatorial indexing for methylation analysis; scXRBS, single-cell extended-representation bisulfite sequencing; itChIP-seq, simultaneous indexing and tagmentation-based chromatin immunoprecipitation with massively parallel DNA sequencing; CoBATCH, combinatorial barcoding and targeted chromatin release; scCUT&Tag, single-cell cleavage under targets and tagmentation.

The histone post-translational modification represents another epigenetic regulation, in which proteins of diverse modifications modulate the behavior of DNA molecules by physical interactions [47]. The chromatin immunoprecipitation (ChIP) is a popular method for detecting genome-wide modifications of histones and bindings of TFs. Although having been applied to identify heterogeneity of chromatin states, single-cell ChIP sequencing assays usually suffer from weak signal-to-noise ratio [48–50]. In recent years, a few non-immunoprecipitation, enzyme-tethering chromatin profiling approaches have been developed to improve the efficiency of epigenomic analysis at the single-cell level. For example, scCUT&Tag [51], CoBATCH [52], ACT-seq [53], and ChIL-seq [54] use the Tn5 transposase that is tethered to protein A binding to the antibody to simplify experimental procedures and alleviate the loss of biological signal. Additionally, the newly reported MulTI-Tag can profile multiple histone modifications simultaneously in single cells [55].

The process of DNA methylation is the covalent addition of a methyl (CH₃) group to the 5'-carbon of the cytosine to form 5-methylcytosine. The DNA methylation occurs almost exclusively in the context of CpG dinucleotides of the genome and plays an essential role in gene regulation via recruiting proteins repressing transcription or inhibiting the binding of TFs [56]. Single-cell genome-wide DNA methylation sequencing methods mainly apply two strategies: reduced-representation bisulfite sequencing [57,58] or whole-genome

bisulfite sequencing [59–61]. Recently, a single-cell extended-representation bisulfite sequencing technology struck a balance between the coverage and enrichment of regulatory elements [62]. Importantly, it has been revealed that DNA modifications may also have long-term noninvasive lineage-tracking potential for their inheritance [25]. In addition, the dynamics of DNA methylation have shown great association with mammalian aging, and specific CpG loci can predict the biological ages of cells, tissues, as well as organisms [63,64].

Transcriptome sequencing at the center stage of SCO

With tremendous advances made in the throughput, accuracy, automation, and commercialization, scRNA-seq has been the most widely used single-cell option because of its low cost and high availability for most researchers. Two complementary strategies are frequently used [65]. Plate-based methods usually capture full-length messenger RNA (mRNA), and can detect more genes and transcripts including low abundance transcripts as well as alternative splicing events in individual cells [66–69]. In contrast, droplet-based approaches [70,71] are more likely to detect rare cell types given their properties of high throughput. For example, the 10X Genomics system can partition up to 10,000 cells per channel [72]. However, this kind of method is usually subjected to more dropout events since only a fraction of the transcriptome is captured in

individual cells [73]. Additionally, most scRNA-seq approaches only provide snapshots of cellular transcriptomes rather than temporal dynamics. Several methods have come a long way to distinguish newly transcribed and pre-existing mRNAs in the same cell [74–77]. The newly reported Live-seq enables the single-cell transcriptome profiling while keeping the cell alive and functional [78].

All in one: single-cell multimodal omics

Despite great advances in profiling cell-to-cell heterogeneity at unprecedented resolution and scale, these technologies can only investigate one characteristic of cells. In fact, molecular interrogations of single cells encompass not only the genome, epigenome, and transcriptome, but also epitranscriptome [79,80], proteome [81], as well as epiproteome [82], all of which collectively depict comprehensive characteristics of cells. In order to dissect a cell more comprehensively, single-cell multimodal omics sequencing tools [83–147] emerge by measuring multiple modalities at a time in an individual cell (**Table 1**). Importantly, combining SCO sequencing with clustered regularly interspaced short palindromic repeats (CRISPR)-based perturbations can directly associate genotypes with phenotypes. These have greatly accelerated our comprehensive understanding of the complexity of genetic variants, gene expression, intracellular regulatory networks, intercellular crosstalk, and environmental effects on animal cells.

Basic pipeline of scRNA-seq data analysis

Recently, the rapid generation of SCO sequencing data and construction of cell atlases have been both the beneficiary and the driving force of computational advances, formulating a positive feedback loop. Featuring wide applicability and availability, scRNA-seq has been the center of attention for computational biologists. The total number of scRNA-seq analytic tools has reached approximately 1400 [148]. A recent review comprehensively introduced the single-cell analysis across modalities, providing suggestions for the best computational workflows for other omics data analysis including chromatin accessibility, surface protein expression, immune receptor repertoires, and spatial localization patterns [149]. In view of this, here we take scRNA-seq data analysis as an example to illustrate the computational aspect of SCO sequencing.

Although the growing number of tools substantially facilitate the inception of new scientific insights, it makes the standardization of the workflow challenging and complex. Generally, benchmarks provide suggestions for best practices to choose from. Of note, there is not always a “golden rule”. In practice, the optimal choice of computational methods could be different in the context of specific biological backgrounds and analytical goals. Researchers should make special effort to try multiple methods, perform careful parameter tuning, and sometimes tailor the algorithm according to their needs.

Here we present an overview of the basic pipeline of scRNA-seq data analysis, from pre-processing to cell type identification (**Figure 2**). In each section, we summarize the conceptual reasoning, computational challenges, common practices, and recent advances.

Quality control

After count matrices are obtained from initial raw data pre-processing, which can be readily handled by pipelines such as Cell Ranger [72] and kallisto | bustools [150], quality control (QC) is a necessary step to filter out low-quality cells, empty droplets, doublets, or multiplets. Commonly used QC metrics include the detected gene number, the total count, and the percentage of counts attributed to mitochondrial RNA in each cell. Low-quality cells and empty droplets can be identified by too few genes and total counts. High mitochondrial RNA percentage, reflecting plasma mRNA leakage due to compromised cell membrane integrity, is another sign of low-quality or dying cells. Aberrantly high gene numbers and total counts indicate doublets. Doublets can be further verified with computational doublet detection methods [151–153], most of which can be summarized as setting a threshold for the similarity of each droplet to artificial doublets. DoubletFinder [151] is considered to have the best doublet detection accuracy according to a benchmark [154]. The aforementioned QC metrics should be jointly considered and thresholds should be chosen carefully to avoid filtering out biologically informative cells. For example, comparably high mitochondrial RNA may also correspond to cells active in respiration [155].

Normalization

The goal of normalization is to account for unbalanced sequencing depths and to stabilize the variance. Generally, normalization includes two steps, scaling with a size factor to make absolute counts comparable, and transformation to mitigate the skewness in the distribution of gene expression. The most commonly used method is LogNormalize, as implemented in Seurat [156] and SCANPY [157]. It employs a uniform size factor for all cells and performs log transformation. Evidently, it relies on the questionable assumption that all cells have the same amount of mRNA and all genes should apply the same size factor. To address this, SCnorm has been developed, which groups genes with their count–depth relationships calculated by quantile regression, and assigns each gene group a scale factor [158]. Alternatively, the increasingly popular sctransform utilizes a generalized linear model to explain the sequencing depth with counts, and performs transformation using the Pearson residuals in a gene-specific manner, although it is only applicable to unique molecular identifier (UMI)-based scRNA-seq data [159].

Feature selection

Selecting a subset of the most informative and representative genes is a crucial prerequisite for dimensionality reduction. To reflect the variability in the expression profile across cells, the task equals to identifying highly variable genes (HVGs). Feature selection methods differ in the choice of variability measurement, including variance [156,157,160,161], dropout rate [162,163], and Gini index [164]. In the commonly used HVG selection method implemented in Seurat [156], genes are first binned according to mean expression, and HVGs are subsequently selected by normalized dispersion within each bin, thus preventing the under-representation of lowly expressed genes in HVGs.

Table 1 Single-cell multimodal omics sequencing methods

Method	Genome	Chromatin accessibility	DNA methylation	Histone modification	Chromosome conformation	Transcriptome	Proteome	Perturbation	Ref.
DR-seq	✓					✓			[83]
G&T-seq	✓					✓			[85]
SCTG	✓					✓			[84]
sci-L3-RNA/DNA	✓					✓			[116]
scONE-seq	✓					✓			[147]
SIDR-seq	✓					✓			[103]
TARGET-seq	✓					✓			[113]
DNTR-seq	✓					✓			[120]
scNanoATAC-seq	✓	✓							[146]
scGET-seq	✓	✓							[142]
scTrio-seq	✓		✓			✓			[91]
scTrio-seq2	✓		✓			✓			[99]
scCOOL-seq	✓	✓	✓						[95]
iscCOOL-seq	✓	✓	✓						[107]
PHAGE-ATAC	✓	✓					✓		[137]
sci-CAR		✓				✓			[100]
scPCOR-seq		✓				✓			[139]
scTHS-seq		✓				✓			[104]
SNARE-seq		✓				✓			[105]
SNARE-seq2		✓				✓			[127]
scCAT-seq		✓				✓			[110]
Paired-seq		✓				✓			[117]
SHARE-seq		✓				✓			[118]
ASTAR-seq		✓				✓			[119]
ISSAAC-seq		✓				✓			[144]
scNOMe-seq		✓	✓						[97]
nano-CT		✓		✓					[135]
Pi-ATAC		✓					✓		[101]
ASAP-seq		✓					✓		[125]
ICICLE-seq		✓					✓		[130]
scNOMeRe-seq		✓	✓			✓			[131]
scChARM-seq		✓	✓			✓			[133]
snmCAT-seq		✓	✓			✓			[138]
scNMT-seq		✓	✓			✓			[102]
DOGMA-seq		✓				✓	✓		[125]
NEAT-seq		✓				✓	✓		[136]
TEA-seq		✓				✓	✓		[130]
scM&T-seq			✓			✓			[86]
sc-GEM			✓			✓			[87]
scMT-seq			✓			✓			[92]
Smart-RRBS			✓			✓			[122]
sn-m3C-seq			✓			✓			[108]
scMethyl-HiC			✓			✓			[109]
scDam&T-seq				✓		✓			[114]
EpiDamID with scDam&T-seq				✓		✓			[140]
CoTECH				✓		✓			[132]
Paired-Tag				✓		✓			[134]
scSET-seq				✓		✓			[129]
scNTT-seq				✓			✓		[141]
scCUT&Tag-pro				✓			✓		[145]
ORCA					✓	✓			[111]
PLAYR						✓	✓		[89]
PEA/STA						✓	✓		[90]
REAP-seq						✓	✓		[96]
CITE-seq						✓	✓		[98]
inCITE-seq						✓	✓		[121]
RAID						✓	✓		[106]
SCITO-seq						✓	✓		[123]
SPARC						✓	✓		[128]
Prox-seq						✓	✓		[143]
Perturb-seq						✓		✓	[88]

(continued on next page)

Table 1 (continued)

Method	Genome	Chromatin accessibility	DNA methylation	Histone modification	Chromosome conformation	Transcriptome	Proteome	Perturbation	Ref.
CRISP-seq						✓		✓	[93]
CROP-seq						✓		✓	[94]
Perturb-ATAC		✓						✓	[115]
CRISPR-sciATAC		✓						✓	[124]
Spear-ATAC		✓						✓	[126]
ECCITE-seq						✓	✓	✓	[112]

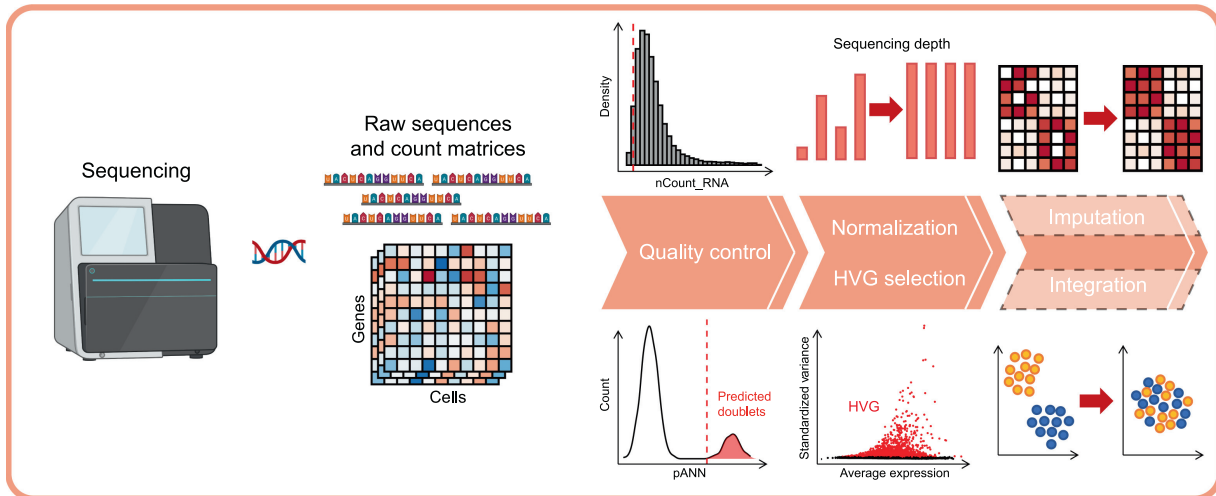
Note: DR-seq, gDNA–mRNA sequencing; gDNA, genomic DNA; mRNA, messenger RNA; G&T-seq, genome and transcriptome sequencing; SCTG, single-cell transcriptogenomics; sci-L3, a single-cell sequencing method that combines combinatorial (3-level) indexing and linear amplification; SIDR-seq, simultaneous isolation of genomic DNA and total RNA sequencing; DNTR-seq, direct nuclear tagmentation and RNA sequencing; ATAC-seq, assay for transposase accessible chromatin with high throughput sequencing; scNanoATAC-seq, single-cell ATAC-seq on Nanopore sequencing platform; scGET-seq, single-cell genome and epigenome by transposases sequencing; scTrio-seq, single-cell triple omics sequencing; scCOOL-seq, single-cell chromatin overall omic-scale landscape sequencing; iscCOOL-seq, improved scCOOL-seq; sci-CAR, single-cell combinatorial indexing-based coassay for chromatin accessibility and mRNA; scPCOR-seq, single-cell profiling of chromatin occupancy and RNAs sequencing; scTHS-seq, single-cell transposome hypersensitive site sequencing; SNARE-seq, single-nucleus chromatin accessibility and mRNA expression sequencing; scCAT-seq, single-cell chromatin accessibility and transcriptome sequencing; Paired-seq, parallel analysis of individual cells for RNA expression and DNA accessibility by sequencing; SHARE-seq, simultaneous high-throughput ATAC and RNA expression with sequencing; ASTAR-seq, assay for single-cell transcriptome and accessibility regions with sequencing; ISSAAC-seq, *in situ* sequencing hetero RNA–DNA-hybrid after assay for transposase-accessible chromatin-sequencing; scNOME-seq, single-cell nucleosome occupancy and methylome-sequencing; nano-CT, nano-CUT&Tag; CUT&Tag, cleavage under targets and tagmentation; Pi-ATAC, protein-indexed assay of transposase accessible chromatin with sequencing; ASAP-seq, ATAC with select antigen profiling by sequencing; ICICLE-seq, integrated cellular indexing of chromatin landscape and epitopes; scNOMERe-seq, single-cell nucleosome occupancy, methylome, and RNA expression sequencing; scChARM-seq, single-cell chromatin accessibility, RNA barcoding, and DNA methylation sequencing; snmCAT-seq, single-nucleus methylcytosine, chromatin accessibility, and transcriptome sequencing; scNMT-seq, single-cell nucleosome, methylation and transcription sequencing; NEAT-seq, sequencing of nuclear protein epitope abundance, chromatin accessibility and the transcriptome in single cells; TEA-seq, assay for transcription, epitopes, and accessibility with sequencing; scM&T-seq, single-cell genome-wide methylome and transcriptome sequencing; sc-GEM, single-cell analysis of genotype, expression and methylation; scMT-seq, single-cell methylome and transcriptome sequencing; RRBS, reduced representation bisulfite sequencing; sn-m3C-seq, single-nucleus methyl-3C sequencing; scMethyl-HiC, single-cell methyl-HiC; HiC, high-throughput chromosome conformation capture; scDam&T-seq, combining single-cell DNA adenine methyltransferase identification (DamID) with messenger RNA sequencing of the same cell; EpiDamID, an extension of DamID to epigenetic chromatin marks; CoTECH, combined assay of transcriptome and enriched chromatin binding; Paired-Tag, parallel analysis of individual cells for RNA expression and DNA from targeted tagmentation by sequencing; SET-seq, same cell epigenome and transcriptome sequencing; scSET-seq, single-cell SET-seq; scNTT-seq, single-cell nanobody-tethered transposition followed by sequencing; scCUT&Tag, single-cell cleavage under targets and tagmentation; ORCA, optical reconstruction of chromatin architecture; PLAYR, proximity ligation assay for RNA; PEA/STA, proximity extension assays/specific (RNA) target amplification; REAP-seq, RNA expression and protein sequencing assay; CITE-seq, cellular indexing of transcriptomes and epitopes by sequencing; inCITE-seq, intranuclear CITE-seq; RAID, RNA and immunodetection; SCITO-seq, single-cell combinatorial indexed cytometry sequencing; SPARC, single-cell protein and RNA co-profiling; Prox-seq, proximity sequencing; CRISPR, clustered regularly interspaced short palindromic repeats; CRISP-seq, an integrated method that combines the resolution of massively parallel single-cell RNA sequencing with the genome editing scale of pooled CRISPR screens; CROP-seq, CRISPR droplet sequencing; CRISPR-sciATAC, CRISPR-based, single-cell combinatorial indexing ATAC; Spear-ATAC, single-cell perturbations with an accessibility read-out using scATAC-seq; ECCITE-seq, expanded CRISPR-compatible cellular indexing of transcriptomes and epitopes by sequencing.

Imputation

It has been reported that scRNA-seq count data often exhibit a high proportion of zero values, sometimes exceeding 90% [165]. Such vast zeros raise concerns about technical noise hindering downstream analysis, for example, obscuring the gene expression correlation [166]. Methods have been invented to impute and denoise the highly sparse count data, most of which not only correct zero values, but also smooth over non-zero values, reasoning that technical noises affect all genes [167]. Some methods like MAGIC [166] treat all zeros as missing data. Conversely, methods like scImpute [168], SAVER [169], and ALRA [170] try to discriminate technical zeros from biological zeros to preserve biologically relevant information. Biological zeros originate from two sources. First, some genes are not expressed in certain cells, and second, the transcription process is not constant but intermittent, in waves of bursts, thus generating transient zero expression [171]. Of note, the

imputation step is not always necessary in the scRNA-seq pre-processing workflow. It has been reported that UMI-based scRNA-seq count data are not zero-inflated, thus nullifying the necessity for imputation [172,173]. In brief, UMI-based counts are less susceptible to amplification bias, greatly lowering the threshold for lowly expressed genes such that they become non-zero, and can be sufficiently modeled with negative binomial distribution without zero-inflation. Of note, over-correction may introduce false signals. An alternative strategy to address the sparsity of scRNA-seq count data is to pool individual cells with similar phenotypes into small groups called metacells, which still offer higher intragroup homogeneity and inter-group granularity than unsupervised clusters, as implemented in MetaCell [174] and SEACells [175]. Nevertheless, one major concern of metacells is the possibility of underrepresentation for rare cell types or cell states, and the metacell-based strategies have not been routinely applied in scRNA-seq data analysis.

Pre-processing



Cell type identification

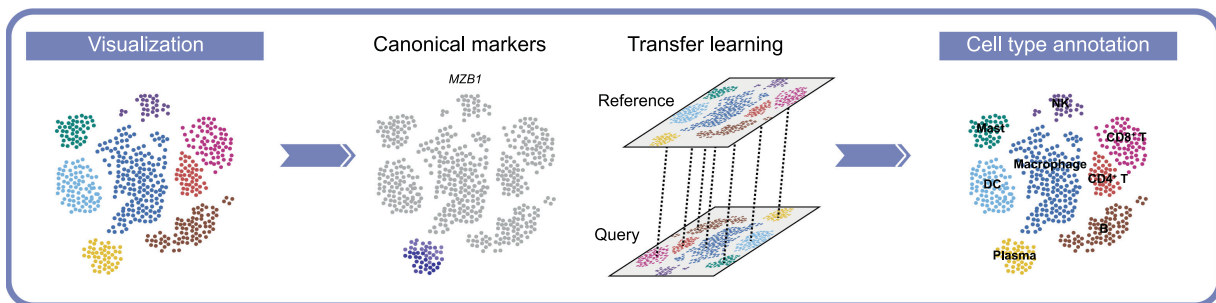


Figure 2 Basic analysis in the scRNA-seq data analysis workflow

scRNA-seq data produced by sequencers undergo pre-processing steps including quality control, normalization, HVG selection, optional imputation, and integration. Dimensionality-reduced data are then visualized and clustered, ready to be assigned cell types with manual or automatic approaches. scRNA-seq, single-cell RNA sequencing; pANN, proportion of artificial nearest neighbors; HVG, highly variable gene; NK, natural killer; DC, dendritic cell.

Integration

Batch effects, which represent technical variation arising from experimental procedures and sometimes include biological factors such as tissues or species [176], can potentially overshadow relevant biological signals. Batch effects almost universally exist in scRNA-seq datasets, and their complexity escalates exponentially when performing atlas-level integration, posing a grand challenge. The goal of integration is to minimize the undesired batch effects, while preserving informative biological variability. Measuring tradeoffs between batch correction and biological variance conservation, scIB [176] compared the overall performance of 16 integration methods in different scenarios, and concluded that a linear embedding model Harmony [177] performs well on datasets with simple and distinct batches, whereas a mutual nearest neighbor matching-based method Scanorama [178] and deep learning-based methods like scVI [161], scANVI [179], and scGen [180], perform best in complex tasks. Deep learning-based methods, most based on autoencoder networks, dominate atlas-level integration, which can be partly attributed to their flexible tunability to reflect complex structures of batch effects, although sometimes at the cost of decreased interpretability and increased time/space complexity.

Cell type identification

For visualization convenience, feature-selected data undergo non-linear dimensionality reduction, either t-distributed stochastic neighbor embedding (t-SNE) [181] or uniform manifold approximation and projection (UMAP) [182], with the latter excelling at time consumption, reproducibility, and global structure preservation [183]. Cells are clustered subsequently, during which process the purity of putative clusters can be evaluated by an entropy-based metric provided by ROGUE [184] for the tuning of clustering resolution. For the subsequent cell type assignment stage, the annotation strategies can be divided into manual and automatic approaches. The manual approach relies on the observation of expression hotspots of canonical markers, which can be obtained from databases such as CellMarker [185] and PanglaoDB [186], or from the literature. Sometimes rounds of sub-clustering are needed to discover rare subsets. Although labor-intensive and subjective, the manual curation process is flexible and tunable in terms of annotation resolution. Conversely, the automatic approaches are time-saving and heavily rely on reference datasets, including marker-based CellAssign [187] and scSorter [188], correlation-based SingleR [189], and supervised classification-based SciBet [190]. Recently, the

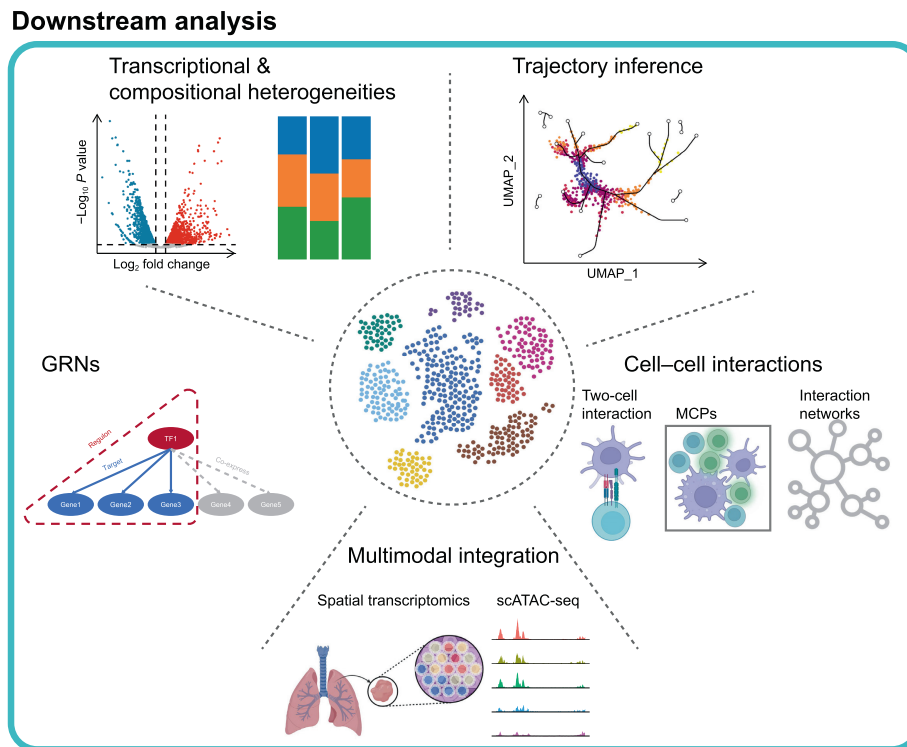


Figure 3 Advanced downstream analysis in the scRNA-seq data analysis workflow

Downstream analysis of scRNA-seq data includes transcriptional and compositional comparative analysis, trajectory inference, GRN reconstruction, cell-cell interaction exploration, and multimodal integration. UMAP, uniform manifold approximation and projection; GRN, gene regulatory network; TF, transcription factor, MCP, multicellular program.

construction of large atlases has inspired the invention of new annotation methods, including CellTypist [191], scArches [192], TOSICA [193], and scBERT [194]. In essence, scArches, TOSICA, and scBERT are deep learning-based, resolving the annotation problem with query-to-reference integration by transfer learning, allowing the easy reuse of annotated consortia without sharing raw data.

Advanced analysis of scRNA-seq data

In contrast to the relatively fixed workflow of basic analysis, advanced downstream analysis of scRNA-seq features considerable flexibility and mutual complementarity, posing even greater challenges for evaluation and standardization. Here we list five popular topics (**Figure 3**) to illustrate, from a computational perspective, how biological questions can be answered by scRNA-seq, and how such methods may evolve.

Transcriptional & compositional heterogeneities

The most basic downstream analysis is a comparative analysis, between subpopulations or between the same subpopulation in different conditions. Transcriptional and compositional heterogeneities are two complementary facets of the comparative analysis, which are occasionally interchangeable depending on the resolution of annotations [195].

The transcriptional heterogeneity can be explored first with differential expression (DE) analysis. DE genes could be fed

into downstream pathway enrichment analysis to gain functional interpretation. A primary principle of DE analysis is to account for the intrinsic biological variability to minimize false positives. Methods designed for bulk DE analysis sometimes perform ironically better than those specifically designed for scRNA-seq [196,197]. Nevertheless, the most popular method is still the simple Wilcoxon rank-sum test, used in 86% of studies of single-cell transcriptomics [197].

One challenge of compositional analysis is that the proportions of different cell types are not independent. Significant expansion of one cell type could result in significant proportion shrinkage of all other cell types. Thus, commonly used univariate tests such as the *t*-test and the Wilcoxon rank-sum test potentially risk false positives. Recent methods scCODA [198] and Cacoa [195] resolve this problem, via joint modeling of all cell types and via isometric log-ratio transformation, respectively, to minimize the dependency among cell type proportions.

Trajectory inference

The lineage trajectory of cells could be inferred from scRNA-seq data, tracing cell transition from one developmental stage or cell state to another, under the assumption that cells along such a trajectory are sufficiently captured and are continuous in expression profiles. Meanwhile, pseudo-temporal ordering of cells along the trajectory empowers the exploration of transition-dependent expression shift. The choice of trajectory inference methods is mainly driven by the expected trajectory

topology, including the choices between connected or disconnected graphs, cycles or trees, and linearity or bifurcation [199]. Commonly used trajectory inference methods include Monocle [200], Slingshot [201], and PAGA [202]. Additional cell transition information could be obtained from the splicing maturation process of mRNA, inspiring the invention of RNA velocity [203], which describes the time derivative of gene expression state, by calculating the ratio between spliced mRNA and unspliced mRNA. The field of trajectory inference has been extended by recent advances. scVelo [203], CellRank [204], and UniTVelo [205] all combine the RNA velocity information with expression profiles to perform trajectory inference or predict cell fates, thus allowing the generation of directed trajectories and nullifying the need for choosing a root node. In addition, researchers have also attempted to align and integrate trajectories from different datasets, leading to the invention of cellAlign [206] and CAPITAL [207].

GRNs

A key notion of gene regulation is that genes do not function in isolation. Instead, genes and TFs are organized in regulons, which constitute complex GRNs, governing and coordinating transcriptional activity of the whole genome. One popular tool for delineation of GRNs is SCENIC [208], which identifies and scores the activity of regulons, followed by the prediction of cell states based on the shared activity of regulatory subnetworks. In terms of performance, methods that do not require pseudo-time-ordered cells as input have been reported to be generally more accurate [209], such as mutual information-based PIDC [210] and random forest-based GENIE3 [211], with the latter applied in SCENIC for the identification of genes co-expressing with TFs. Nevertheless, two benchmarks independently concluded that the performance of all tested methods is less than ideal in accuracy and reproducibility [209,212], calling for future development of better alternatives. Possible improvements have been proposed recently, such as using higher-order moments to distinguish between correlation and regulation [213]. Alternatively, methods leveraging two modalities, transcriptome and epigenome, have been invented to perform integrated regulatory analysis with improved accuracy and reproducibility, including SCENIC+ [214], GRaNIE [215], Pando [216], FigR [217], and MIRA [218]. By directly incorporating TF binding information, these methods allow the inference of enhancer-driven GRNs and the delineation of regulatory circuitry underlying the developmental trajectories of cells.

Cell–cell interactions

Cell–cell interactions are crucial in various cell activities, including differentiation, migration, and apoptosis. Cellular interaction networks could be delineated from scRNA-seq data at different granularity, starting from interactions between two cells as building blocks. Current methods for evaluating cell–cell interactions could be divided into three categories: ligand–receptor expression-based, downstream signaling-derived, and spatial reconstruction-oriented.

Ligand–receptor expression-based methods include the commonly used CellPhoneDB [219] and recently published CellChat [220]. The former boasts the most comprehensive

ligand–receptor database, and the latter employs degree metrics of graph theory to assign signaling roles, thus simplifying complex signaling patterns. In a benchmark, CellChat exhibited the highest consistency with paired spatial information among 16 cell–cell interaction tools [221].

Downstream signaling-derived methods can be represented by NicheNet [222]. NicheNet incorporates intracellular signaling, connecting the binding of a ligand–receptor pair with downstream targets. Complementary to CellPhoneDB, NicheNet empowers reverse thinking, in which pertinent DE genes are first identified and then searched for responsible upstream interactions. CytoSig [223], which measures the cytokine activity, considers downstream signaling for a different reason: the expression of cytokines and their receptors is insufficient to delineate cytokine activity due to its redundant and pleiotropic nature.

Spatial reconstruction-oriented methods like CSOmap [224] are based on the assumption that cells are assembled into spatial structures, where the cell proximity is determined by the ligand–receptor interaction strength. Therefore, CSOmap can reconstruct cell spatial organizations *de novo* from scRNA-seq data. Similarly, NovoSpaRc [225] can also perform spatial reconstruction, although by expression profile similarity instead of interaction strength between cells.

In fact, such interactions occur beyond just between two cells. Under the reasoning that cells in the same spatial niche are exposed to shared cues, eliciting coordinated expression shift, a research group defined multicellular programs (MCPs) as combinations of different cell types and their coordinated expression programs in the tissue, and developed the first method to systematically uncover MCPs, DIALOGUE [226]. We envision that the structural complexity of interaction units will be increasingly represented, leading to the ultimate full depiction of cell–cell interaction networks.

Multimodal integration

The advantage of integration and joint analysis of multimodalities is to complement transcriptomic data with extra information to better delineate regulatory networks and cell–cell interactions.

Spatially resolved transcriptomics assays can provide spatial information lost in the dissociation step of scRNA-seq. The integration of spatial and single-cell transcriptomics includes two aspects, to predict the spatial distribution of undetected RNA transcripts and to deconvolute the cell type composition of each detected spot. As assessed in a benchmark [227], methods based on probabilistic models combined with negative binomial or Poisson distributions, such as gimVI [228], RCTD [229], and Cell2location [230], generally excel at both tasks, and the sparsity of the spatial data is a major determinant of the performance of the aforementioned methods, encouraging increasing depth of sequencing or implementing imputation methods to combat this issue.

ATAC-seq assays can confer the heterogeneity of chromatin accessibility. Although recent sequencing methods like single-nucleus chromatin accessibility and mRNA expression sequencing (SNARE-seq) [105] and simultaneous high-throughput ATAC and RNA expression with sequencing (SHARE-seq) [118] enable simultaneous profiling of transcriptome and epigenome from the same cell, different omics data

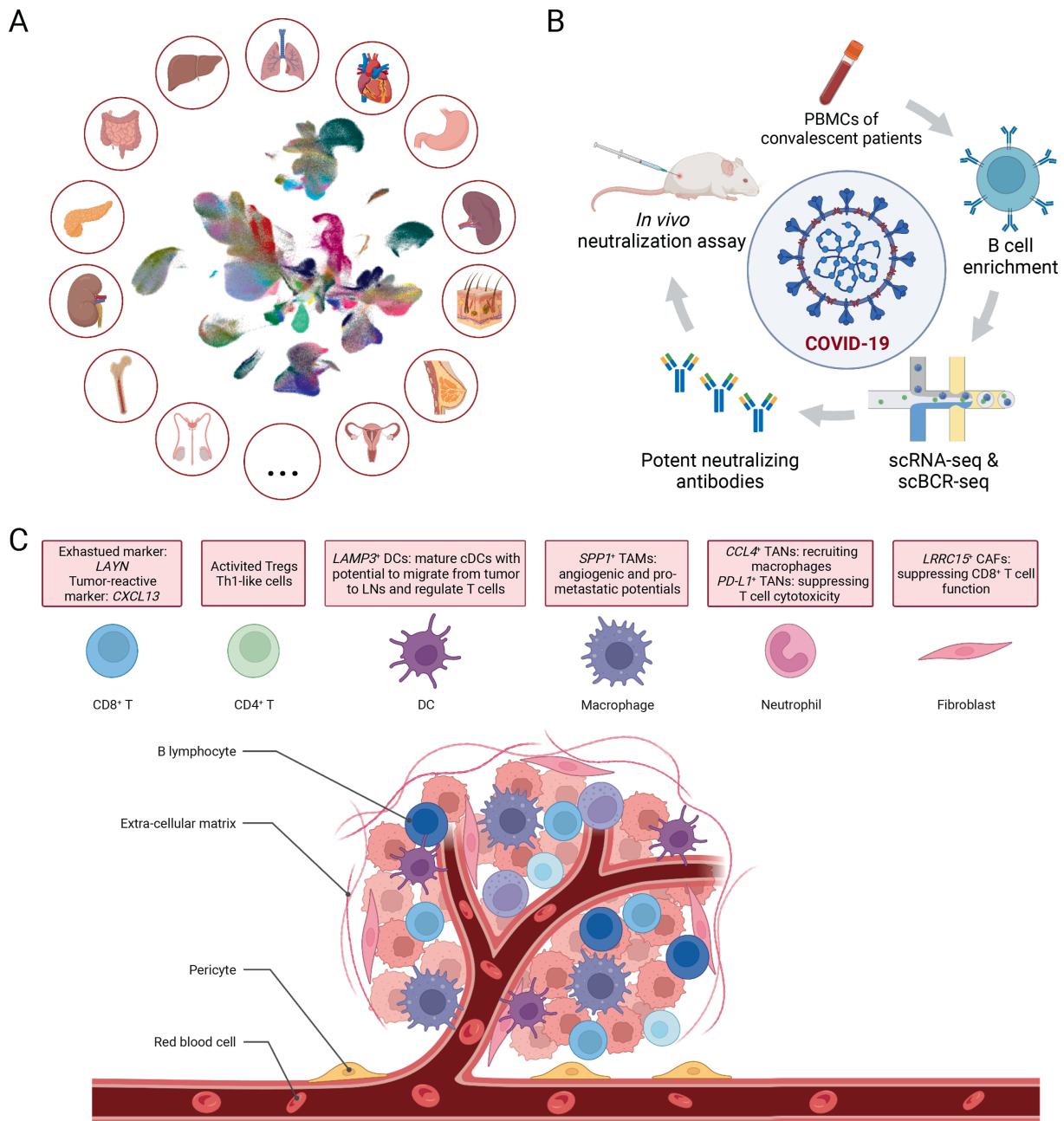


Figure 4 Representative applications of SCO sequencing

A. A single-cell cross-tissue molecular map of the human. **B.** SCO sequencing identifies potent neutralizing antibodies in COVID-19. **C.** Dissection of TMEs using SCO sequencing. SCO, single-cell omics; COVID-19, coronavirus disease 2019; PBMC, peripheral blood mononuclear cell; scBCR-seq, single-cell B cell receptor sequencing; TME, tumor microenvironment; Treg, regulatory T cell; cDC, conventional DC; LN, lymph node; TAM, tumor-associated macrophage; TAN, tumor-associated neutrophil; CAF, cancer-associated fibroblast.

are often unpaired. Methods to integrate unpaired RNA-seq and ATAC-seq data typically aim to align gene expression profiles and chromatin accessibility-derived gene activity scores in a shared space. Canonical correlation analysis (CCA)-based methods include Signac [231], LIGER [232], and bindSC [233]. Among them, bindSC takes a step further beyond traditional CCA by iteratively aligning the matrices of two modalities, thus generating refined co-embeddings. Alternatively, a shared feature representation can be obtained via the reconstruction of a common weighted nearest neighbor (WNN)

graph and the subsequent supervised principal component analysis, as implemented in Seurat [156]. A tough challenge is how to resolve distinct feature spaces of different modalities with minimum information loss. The recently proposed GLUE [234] overcomes this challenge by explicitly modeling regulatory interactions across omics layers and achieves accurate and scalable integration between transcriptome and epigenome data. In contrast to the integration of unpaired multimodal data, MOFA+ [235] and MultiVI [236] are designed for the integration within the same sample space, requiring the

Table 2 Single-cell consortia and data platforms

Platform	Web link
Adult Human Cell Atlas	http://research.gzsums.net:8888
Allen Brain Atlas	https://portal.brain-map.org
Azimuth	https://azimuth.hubmapconsortium.org
Broad Single Cell Portal	https://singlecell.broadinstitute.org/single_cell
Cambridge Cell Atlas	https://www.cambridgecellatlas.org
CancerSEA	http://biocc.hrbmu.edu.cn/CancerSEA
COVID-19 Cell Atlas	https://www.covid19cellatlas.org
Cross-tissue Immune Cell Atlas	https://www.tissueimmuncellatlas.org
Curated Cancer Cell Atlas (3CA)	https://www.weizmann.ac.il/sites/3CA
CZ CELLxGENE Discover	https://cellxgene.cziscience.com
DISCO	https://www.immunesinglecell.org
Expression Atlas	https://www.ebi.ac.uk/gxa
FASTGenomics	https://beta.fastgenomics.org
Gene Expression Omnibus	https://www.ncbi.nlm.nih.gov/geo
GenitoUrinary Development Molecular Anatomy Project	https://www.gudmap.org
Genomic Data Commons Data Portal	https://portal.gdc.cancer.gov
Genotype-Tissue Expression project	https://gtexportal.org/home
Gut Cell Atlas	https://www.gutcellatlas.org
HCA Data Coordination Platform	https://data.humancellatlas.org
Heart Cell Atlas	https://www.heartcellatlas.org
Human BioMolecular Atlas Program	https://portal.hubmapconsortium.org
Human Cell Landscape	http://bis.zju.edu.cn/HCL/
Human Development Cell Atlas	https://developmental.cellatlas.io
Human Protein Atlas	https://www.proteinatlas.org
Human Tumor Atlas Network	https://humantumoratlas.org
HUSCH	http://husch.comp-genomics.org
JingleBells	https://jinglebells.bgu.ac.il
Kidney Cell Atlas	https://www.kidneycellatlas.org
Kidney Precision Medicine Project	https://www.kpmp.org
Liver Cell Atlas	https://livercellatlas.org
Lung Cell Atlas	https://lungcellatlas.org
LungMAP	https://www.lungmap.net
National Genomics Data Center	https://ngdc.cncb.ac.cn/databases
Oral Mucosa Cell Atlas	https://oral.cellatlas.io
PanglaoDB	https://panglaoDB.se
(Re)Building a Kidney	https://rebuildingakidney.com
Reproductive Cell Atlas	https://www.reproductivecellatlas.org
SCPortalen	http://single-cell.clst.riken.jp
scRNASeqDB	https://bioinfo.uth.edu/scrnaseqdb
single-cell eQTLGen Consortium	https://eqtlgen.org/sc/index.html
Tabula Sapiens	https://tabula-sapiens-portal.ds.czbiohub.org
TISH	http://tisch.comp-genomics.org
Tissue Stability Cell Atlas	https://www.tissuestabilitycellatlas.org
UCSC Cell Browser	https://cells.ucsc.edu

measurements of individual modalities performed on the exact same population of cells. Typically, multimodal integration methods are applicable to diverse multimodal integration scenarios, bridging the transcriptome, surface proteome, and epigenome to comprehensively characterize distinct cell states and dynamics.

Mapping cellular heterogeneity using SCO sequencing

Dissecting the cellular diversity of the human body

Although there are numerous advanced SCO sequencing technologies, the most widely used one is scRNA-seq for its low cost and robust performance. In recent years, a large number of studies have delineated transcriptome characteristics of diverse cell types and cell states in both homeostatic and dis-

eased conditions. Herein, we review a few representative insights. Deep profiling of ~ 2400 cells from human blood using Smart-seq2 protocol identified a novel dendritic cell (DC) subset with the ability to potently activate T cells, representing less than 3% of the blood DC populations [237]. These fundamental findings have modified the taxonomy of DCs and monocytes and will facilitate their developmental and functional analysis in health and disease. For neutrophils, despite their short lifetimes and intrinsic poor viability, systematic single-cell analysis of such cells has established the reference model across multiple tissues and highlighted a discrete and definable neutrophil subpopulation expressing interferon stimulating genes [238,239]. Another study reported a fascinating new neutrophil subset that has the ability to improve central nervous system neuron survival and axon regeneration [240]. Additionally, single-cell profiling of hematopoietic cells revealed that lineage commitment is a continuous process, challenging the classical hematopoietic model in which

hematopoietic system had been acknowledged as a collection of discrete hierarchically organized progenitor populations [241].

Furthermore, the ever-increasing throughput has enabled organ-level even organism-level atlases (**Figure 4A**). Sequencing individual nuclei and cells from six anatomical adult heart regions revealed the complexity of the cellular heterogeneity of cardiomyocytes, pericytes, as well as fibroblasts, and highlighted cardiac resident macrophages with protective and inflammatory transcriptional signatures [242]. These results deepen our understanding of the molecular mechanisms underlying cardiovascular diseases and therapeutic strategies. Profiling the spatial and temporal architecture of the developing and mature human kidney has demonstrated that the localization of antibacterial neutrophils and macrophages are well orchestrated by the epithelial-immune crosstalk to the kidney regions which are the most susceptible to infection [243]. Moreover, four back-to-back studies reported pan-tissue single-cell transcriptome atlases characterizing about 500 cell types covering more than a million cells across over 30 human tissues [191,244–246]. Cross-tissue comparison of cell types has provided critical insights into the cell heterogeneity and revealed shared and tissue-specific transcriptional features about organ development and functions [3,247]. Taken together, SCO technologies, especially transcriptome sequencing, have greatly facilitated the identification of rare cell types, the investigation of cellular functions, and the understanding of cell fate determinations.

Deciphering the complexity of human disease at the cellular level

Significantly, disease-oriented applications of the SCO method have been conducted to disentangle multiple human tissues in disorder. For instance, population-scale scRNA-seq analysis of skin and blood samples from healthy controls and patients with scleroderma, a severe autoimmune disease, revealed a previously undefined scleroderma-associated fibroblast, whose perturbations are primarily associated with disease severity and clinical features [248]. A study of pediatric colitis and inflammatory bowel disease via single-cell and risk gene analysis elucidated the common pathogenesis marked by defective cyclic adenosine monophosphate (cAMP) response pathway, and demonstrated that the drug dipyridamole modulating cAMP signaling can restore immune homeostasis and improve clinical symptoms [249]. In another study, single-cell analysis of Crohn's disease patients identified a cellular module that is composed of IgG plasma cells, inflammatory mononuclear phagocytes, activated T cells, and stromal cells, and a subset of patients highly expressing the cellular module in their inflamed tissues exhibited resistance to anti-tumor necrosis factor (TNF) therapy [250].

SCO methods have also been used to dissect the human brain, the most complex organ in the human body. For instance, compared with the normal brain, a distinct transcriptional state that corresponds to the nidus emerges in malformed human brain vasculature [251]. Profiling single-nucleus cortical transcriptomes of 48 Alzheimer's disease individuals with varying degrees of pathology, revealed that the strongest disease-associated changes occur at early states in a cell type-specific manner, whereas genes highly expressed at late stages are common across cell types [252]. Notably, a nas-

cent field, single-cell genetics, is emerging at the intersection of SCO and human genetics [253]. Several studies combined scRNA-seq with genotype data to identify substantial expression quantitative trait loci, most of which show cell type-specific effects on gene expression and some of which are linked to diseases [254–257]. Additionally, corresponding tools have been developed to identify disease-associated cell types [258] or individual cells [259]. Taken together, these efforts have transitioned the understanding of disease biology from disease-causing genes to specific cells, programs, and tissues [260], and elucidation of the genetic basis will have broad implications for the treatment of diseases.

In the past three years, coronavirus disease 2019 (COVID-19) caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has spread rapidly as a global pandemic. Accordingly, peripheral blood and lung biopsies of patients have been subjected to SCO sequencing around the world. The community has endeavored to analyze immune responses during COVID-19 infection to gain insights into the biology of the disease via single-cell methodologies [261,262]. Importantly, virus neutralizing antibodies present in the plasma of convalescent patients, represent a promising therapeutic intervention by providing effective prevention for virus entry into host cells. Using high-throughput scRNA-seq and single-cell B cell receptor sequencing (scBCR-seq), potent neutralizing antibodies were identified from convalescent patients' B cells, and their efficiencies were validated by *in vitro* and *in vivo* SARS-CoV-2 neutralization assays [263] (**Figure 4B**).

Reference data resources of the wide community

In addition to great insights gained in multiple research fields, the single-cell revolution has led to the initiation and pursuit of several global consortia, projects, as well as databases (**Table 2**). The international Human Cell Atlas (HCA) initiative has ambitiously taken place to map all the cells of the human body [264,265]. To date, the HCA community has generated multimodal omics data for more than 36 million cells covering nearly 5000 donors across almost all of human tissues, and all the produced datasets are being processed by uniform pipelines and publicly accessible. Beyond providing molecular cell landscape of the human body, the Human BioMolecular Atlas Program (HuBMAP) aims to build a Human Reference Atlas, which charts three-dimensional organizations of whole organs and provides standard terminologies by anatomical structures, cell types, plus biomarkers (ASCT+B) tables [266,267]. In summary, cutting-edge SCO sequencing technologies and massive multimodal omics sequencing data are transforming our understanding of human disease at the cellular and tissue levels.

SCO sequencing in cancer research

Cancer is a systemic disease in which malignant cells arising from genetic alterations acquire capabilities for escaping from the surveillance of the immune system, proliferating uncontrollably, as well as invading local or distant normal tissues [268]. The tumor microenvironment (TME) is a complex ecosystem in which immune cells interact with heterogeneous malignant cells and stromal cells to mediate tumor progression, metastas-

sis, relapse, and drug resistance. In the last few years, a large number of studies have been conducted to resolve the complexity of the TME at single-cell resolution and brought significant insights into TMEs across many cancer types, including melanoma [269–272]; glioma [273–275]; and breast [50,276–281], colorectal [99,282–288], gastric [289,290], liver [91,291–293], kidney [294], pancreas [295,296], and lung [297–304] cancers. Here we summarize a couple of key aspects.

Cellular diversity in TMEs at baseline

Applications of SCO sequencing technologies have revealed a few tumor-specific cell states which could be potential targets for cancer immunotherapies to enhance anti-tumor abilities of the TME (Figure 4C). Influenced by chronic antigen stimulations, T cells usually reach a dysfunctional state called exhaustion, characterized by inactive cytotoxicity and augmented expression of inhibitory receptors, including PD-1, CTLA-4, TIM-3, TIGIT, and LAG3 [305]. Importantly, integrated analysis of tumor-infiltrating T cells from hepatocellular carcinoma patients has identified *LAYN* as a suppressive marker of exhausted CD8⁺ T cells, which are preferentially enriched and clonally expanded in tumors compared with peripheral blood and adjacent normal tissues [291]. As all T cells from the same clone share identical T cell receptor (TCR) sequences, combined scRNA-seq and single-cell TCR sequencing (scTCR-seq) can associate cellular states with clonal expansion patterns and cellular lineages [306]. For example, CD8⁺ T cells exhibiting states preceding exhaustion have also been observed in treatment-naive non-small-cell lung cancer patients, and the higher ratio of “pre-exhausted” T cells to exhausted T cells indicates better prognosis of lung adenocarcinoma [302]. Notably, combined gene expression and TCR-based lineage tracing revealed that CD8⁺ effector and exhausted T cells are independently connected with tumor-infiltrating CD8⁺ effector memory cells in colorectal cancer, although both exhibiting high clonal expansion [287]. Another player in the tumor immunity, CD4⁺ T cells, play an important role in regulating effective immune responses to cancer cells. Deep single-cell transcriptome profiling based on Smart-seq2 protocol revealed two different *FOXP3*⁺ regulatory T cells (Tregs) which present distinct distribution patterns of *TNFRSF9* (*4-1BB*) indicating activation of antigen-specific Tregs, and those activated Tregs are associated with poor prognosis in lung adenocarcinoma [302]. Notably, two T helper 1 (Th1)-like populations marked by *IFNG* have been identified in colorectal cancer, yet only *CXCL13*⁺*BHLHE40*⁺ Th1-like subset is specifically enriched in microsatellite-unstable patients [287].

In addition, SCO sequencing technologies have also shed important light on the tumor-infiltrating myeloid compartment and their crosstalk with lymphocytes and non-immune cells. Integrated analysis of two scRNA-seq platforms demonstrated that *LAMP3*⁺ DCs, derived from conventional DCs (cDCs), have the potential to migrate from hepatic tumors to local lymph nodes (LNs) [293]. This population was further confirmed in lung cancer (named as “mregDC”) [304]. Single-cell interrogations of TMEs in colorectal cancer patients revealed two distinct subsets of tumor-associated macrophages (TAMs) that show dichotomous functional

phenotypes. Strikingly, the pathway analysis demonstrated that *SPP1*⁺ TAMs are specifically enriched in tumor angiogenesis genes, whereas *CIQC*⁺ TAMs are associated with the complement activation and antigen presentation signaling. Importantly, computational modeling of scRNA-seq and The Cancer Genome Atlas (TCGA) data illustrated that TAMs and cDCs constitute the core components of the cell–cell interaction network in colorectal cancer patients [288]. Recently, a large-scale single-cell atlas of human liver cancer was reported, consisting of 160 samples from 124 treatment-naive patients, all of whom can be stratified into five subtypes based on their composition of immune and stromal cells. Of the five subtypes that harbor distinct TMEs, tumor-associated neutrophils enriched in the myeloid-cell-enriched subtype have been associated with poor prognosis [292].

To understand whether such significant insights gained from a single cancer type can be extended to other cancer types, several studies conducted pan-cancer analysis to comprehensively characterize the degree of similarity and discrimination of tumor-infiltrating immune cells across different tumors. For instance, a pan-cancer study of myeloid cells showed that the proportion of mast cells varies remarkably across different cancer types. Specifically, nasopharyngeal cancer patients harbor the highest proportion of mast cells that are largely absent in many other cancer types such as multiple myeloma and hepatocellular carcinoma, indicating that mast cells might exhibit diverse functions in different tumor contexts [307]. An integrated analysis of 397,810 T cells from 316 patients across 21 cancer types, computationally detected two major developmental paths from naive to exhausted T cells, which pass through the tissue-resident memory T cells and effector memory T cells, respectively [308]. Additionally, a cross-tissue study integrating single-cell transcriptomic data of fibroblasts uncovered the cancer-associated *LRRCL15*⁺ myofibroblasts, playing a role in pro-tumor immunity [4,309]. In summary, distinguishable immune cell signatures have been found across different cancer types, which may lead to the development of more personalized immunotherapies for different patients with distinct clinical parameters.

Insights into cancer immunotherapies

Cancer immunotherapy, attempting to restore the host’s natural defenses to eradicate malignant cells, represents a promising strategy for cancer treatment. Immune checkpoint blockades (ICBs) are designed to inhibit immunoregulatory pathways such as the PD-1/PD-L1 signaling axis, and thus to promote the elimination of malignant cells [310]. Advances in SCO sequencing technologies have enabled the comprehensive investigation of the dynamic properties of tumor-infiltrating immune cells during the course of cancer immunotherapies. Taking non-small-cell lung cancer as an example, temporal single-cell tracing analysis of 36 patients after anti-PD-1 therapy has found that precursor exhausted T (Texp) cells, which show low expression of coinhibitory genes and high expression of *GZMK*, accumulate after treatment in responsive patients. In contrast, nonresponsive patients did not exhibit the increased levels of Texp cells. In addition, paired TCR sequencing analysis has demonstrated that these Texp cells can accumulate not only through

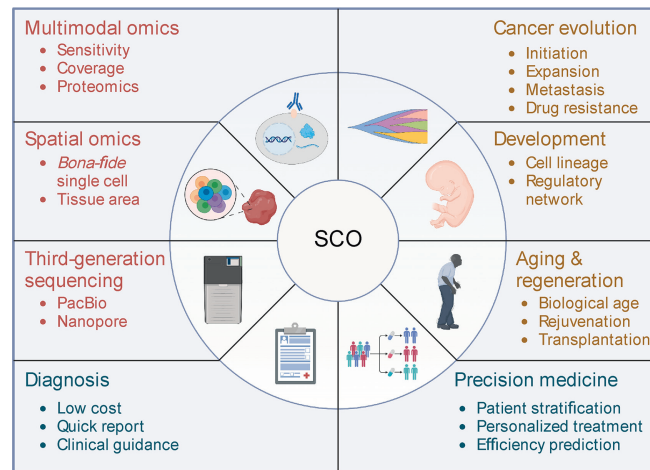


Figure 5 Perspectives of SCO

Perspectives of SCO consist of technological advances, human fundamental research, and clinical prospects.

replenishment from peripheral T cells, but also through their local expansion. This has been named the clonal revival phenomenon [300]. A single-cell meta-analysis of 225 samples from 102 ICB-treated patients across five cancer types has revealed that the *CXCL13* could be an effective marker of tumor-reactive CD8⁺ T cells within tumors, and that the high proportion of *CXCL13*⁺ CD8⁺ T cells is indicative of favorable responses to ICB therapy [311]. Another study found that CD4⁺ neoantigen-reactive T cells show significant *CXCL13* expression compared with bystander cells [312]. To evaluate the efficacy of a combination of chemotherapies and immunotherapies, a triple-negative breast cancer study was conducted to dissect the dynamics of immune cells in the TME and peripheral blood of patients treated with paclitaxel or paclitaxel plus atezolizumab using scRNA-seq and single-cell ATAC sequencing. These data have revealed that the expansion of tumor-reactive immune cells caused by atezolizumab can actually be hindered by paclitaxel, leading to ineffective combination therapy [276].

Despite the success of cancer immunotherapies, durable responses have only been achieved in a fraction of cancer patients. Clearly, the dual effect of many immune cells such as macrophages and neutrophils in cancer poses a challenge for cancer immunotherapies. Specifically, the influential role of neutrophils in cancer biology and their potential as therapeutic targets are now only marginally recognized. Current cancer immunotherapies primarily target a single immune cell type or a single subpopulation of cell types. In the future, SCO sequencing technologies will play a greater role in the prediction of treatment efficiency and the development of more personalized therapies.

Perspectives

There is little doubt that the increasing use of SCO technologies in human biology and disease research will continue for the foreseeable future. This increased trend has been enabled by improvements in technologies and usability, both of which will continue to develop (Figure 5).

Spatially resolved single-cell multimodal omics sequencing technologies

Although dozens of single-cell multimodal omics sequencing technologies have been developed (Table 1), many challenges still remain [7]. Compared to unimodal solutions, the detection sensitivity of individual modalities in single-cell multimodal sequencing methods is relatively low, which poses difficulties in distinguishing technical noise from biological signals. Additionally, these methods typically exhibit limited coverage, in the sense that not all genome-wide fragments can be captured uniformly, and thus the resulting data are too sparse to provide comprehensive characteristics of cells. Another weakness lies in the fact that few solutions have implemented single-cell high-throughput profiling of the proteome [313], which directly links the cellular output to its function. Specifically, assays for detecting genome and proteome simultaneously will enable the investigation of the extent to which genetic variants affect the physiological functions of cells and tissues. The combined analysis of transcriptome and proteome allows to delineate the dynamic relationship between RNA and protein abundance across different cell types and tissues.

Although SCO sequencing technologies have enabled the high-resolution investigations of biological systems, cell-cell spatial relationships and communication are lost during tissue dissociation. In fact, cellular functions not only depend on intracellular molecules and events, but also are tightly regulated by their well-organized niche cells [314]. Recent years have witnessed the rapid development of sequencing-based spatially resolved omics technologies [315–330]. However, these spatial strategies capture partial cellular content or a mixture of multiple cells, instead of profiling cells at *bona-fide* single-cell resolution. Although the imaging-based spatially resolved methods can achieve single-cell resolution, they are still limited in terms of overall throughput and gene coverage [331]. In the future, exact single-cell spatially resolved omics sequencing may be possible by encoding cellular molecules in a self-adaption rather than grid-like fashion for individual cells. In addition, combining single-cell CRISPR-based perturbations with spatial information may enable

in situ investigations of the complex connections between genotypes and phenotypes [325].

Finally, all current single-cell multimodal and spatially resolved omics sequencing technologies are based on next-generation sequencing platforms. To sequence long fragments directly and improve the sequencing accuracy, more high-throughput technologies based on the third-generation sequencing platforms such as Nanopore and PacBio are likely to emerge and grow.

Decoding temporal dynamics underlying human biology and disease

In fact, most biological processes in humans are involved in temporal dynamics, such as embryo development, individual aging, and disease progression. For example, the cancer evolution consists of multiple crucial transitions, including tumor initiation from precancerous lesions to malignancy, local expansion and distant metastasis, as well as progression to the drug-resistant state [332]. In these processes, SCO sequencing technologies can play roles in unraveling the diversity of cell states, the dynamics of cell fate, and the complexity of cell–cell interactions through the sequential sampling strategies. Additionally, integrated analysis of single-cell sequencing data and a mass of archived bulk sample sequencing data with pair clinical information in the International Cancer Genome Consortium (ICGC) [333], TCGA [334], and Genotype-Tissue Expression (GTEx) [335], will reveal potential biomarkers for detecting precancerous, metastatic, and drug-resistant transitions, improve the early detection and patient stratification, and facilitate drug screening and personalized treatment.

Another threat to human health is aging, in which cells lose their physiological integrity and organs gradually display dysfunctional states. Notably, aging directly contributes to many diseases, such as neurodegeneration, cancer, and cardiovascular diseases [336]. The application of SCO sequencing technologies makes it possible to measure biological ages at the molecular and the cellular levels [337]. Calculating aging scores at the cellular level will uncover heterogeneity between cells and asynchronism between biological and chronological ages. Specifically, immune aging characterized by systemic pathogen-free inflammation in aged individuals can lead to profound effect on immune processes, and comprehensive elucidation of aging immune cells may reveal key targets to rejuvenate the immune system [338].

In addition to aging, organic diseases can also lead to functional deficiencies in organs. From a long-term perspective, transplants are the most effective treatment for organ failure. For example, human pluripotent stem cell-derived islets have been a promising strategy for the therapy of insulin-deficient diabetes [339]. The combined analysis of SCO sequencing technologies and organoids will play an important role in the exploration of stem cell maintenance and differentiation.

Overcoming the bottleneck of translating SCO technologies to clinical applications

Significantly, SCO sequencing is beginning to profoundly impact the development of precision medicine, including more accurate patient stratification and more personalized treatment. Ultimately, translating these technological advances into

clinical practice will greatly improve the accuracy of disease diagnosis and treatment. For example, functional CRISPR screens with single-cell readout will facilitate the dissection of disease mechanisms and accelerate drug discovery.

The first of many challenges that need to be addressed is the high cost associated with SCO examinations, which hinders their widespread adoption in clinical settings. A substantial reduction in cost is necessary to make single-cell-based assays affordable for most patients in a variety of diagnostic settings. One way to achieve this is through automation of sampling, library construction, sequencing, as well as medical equipment operation. Another approach is to streamline the bioinformatic analysis process via user-friendly software tools that automate data processing and analysis. In addition, it is imperative to reduce the time taken for single-cell-based assays. In general, it should take no more than two days from clinical sampling to the generation of informative medical reports for effective diagnostic or therapeutic purposes. Reducing the time and expertise required for experiments and data analysis will make single-cell assays more accessible to a wider range of healthcare. Furthermore, the vast majority of the current SCO works are of a research nature, with the primary goal of generating new hypotheses or advancing new research directions. In clinical settings, however, SCO-derived results need to provide specific guidance to medical staff members for achieving more precise diagnosis and actionable plans. Thus, it is essential to establish more definitive links between SCO-derived results and specific clinical parameters, such as treatment responses and detailed pathological classifications.

In summary, clinical desirability calls for a concerted effort by all stakeholders involved in the healthcare industry. Advances in the technology and medical research will not only improve patient outcomes but also increase the efficiency of the healthcare system. Although the road ahead will be long, SCO sequencing technologies represent a hugely promising gateway to precision medicine.

Competing interests

Zemin Zhang is a founder of Analytical BioSciences and is a board member for InnoCare Pharma. Other authors have declared no competing interests.

CRedit authorship contribution statement

Qiang Shi: Conceptualization, Investigation, Writing – original draft, Visualization, Funding acquisition. **Xueyan Chen:** Investigation, Writing – original draft, Visualization. **Zemin Zhang:** Supervision, Writing – review & editing, Funding acquisition. All authors have read and approved the final manuscript.

Acknowledgments

We apologize for the exclusion of numerous important studies due to limited space. Part of the analysis was performed on the High Performance Computing Platform of the Center for Life Sciences in Peking University, China. Part of figures was created with the help of BioRender. This review was supported by the National Natural Science Foundation of China

(Grant Nos. 81988101, 31991171, 91959000, 62203019, 92159305, and 92259205), the Beijing Municipal Science and Technology Commission (Grant No. Z221100007022002), and the Changping Laboratory, China. Qiang Shi was supported in part by the China Postdoctoral Science Foundation (Grant Nos. 2021TQ0012 and 2022M720246), the Peking University Boya Postdoctoral Fellowship, and the Postdoctoral Fellowship of Peking-Tsinghua Center for Life Sciences, China.

ORCID

ORCID 0000-0003-1851-773X (Qiang Shi)

ORCID 0000-0001-8106-5313 (Xueyan Chen)

ORCID 0000-0003-3789-6536 (Zemin Zhang)

References

- [1] Ye Z, Sarkar CA. Towards a quantitative understanding of cell identity. *Trends Cell Biol* 2018;28:1030–48.
- [2] Liao J, Lu X, Shao X, Zhu L, Fan X. Uncovering an organ's molecular architecture at single-cell resolution by spatially resolved transcriptomics. *Trends Biotechnol* 2021;39:43–58.
- [3] Elmentaite R, Dominguez Conde C, Yang L, Teichmann SA. Single-cell atlases: shared and tissue-specific cell types across human organs. *Nat Rev Genet* 2022;23:395–410.
- [4] Buechler MB, Pradhan RN, Krishnamurthy AT, Cox C, Calviello AK, Wang AW, et al. Cross-tissue organization of the fibroblast lineage. *Nature* 2021;593:575–9.
- [5] Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. *Nat Rev Genet* 2019;20:631–56.
- [6] 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.
- [7] Zhu C, Preissl S, Ren B. Single-cell multimodal omics: the power of many. *Nat Methods* 2020;17:11–4.
- [8] Stubbington MJT, Rozenblatt-Rosen O, Regev A, Teichmann SA. Single-cell transcriptomics to explore the immune system in health and disease. *Science* 2017;358:58–63.
- [9] Papalexi E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol* 2018;18:35–45.
- [10] Camp JG, Platt R, Treutlein B. Mapping human cell phenotypes to genotypes with single-cell genomics. *Science* 2019;365:1401–5.
- [11] Method of the year 2013. *Nat Methods* 2014;11:1.
- [12] Method of the year 2019: single-cell multimodal omics. *Nat Methods* 2020;17:1.
- [13] Crick F. Central dogma of molecular biology. *Nature* 1970;227:561–3.
- [14] Wen L, Tang F. Recent advances in single-cell sequencing technologies. *Precis Clin Med* 2022;5:pbac002.
- [15] Zong C, Lu S, Chapman AR, Xie XS. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* 2012;338:1622–6.
- [16] Fu Y, Li C, Lu S, Zhou W, Tang F, Xie XS, et al. Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification. *Proc Natl Acad Sci U S A* 2015;112:11923–8.
- [17] Chen C, Xing D, Tan L, Li H, Zhou G, Huang L, et al. Single-cell whole-genome analyses by Linear Amplification via Transposon Insertion (LIANTI). *Science* 2017;356:189–94.
- [18] Chu WK, Edge P, Lee HS, Bansal V, Bafna V, Huang X, et al. Ultraaccurate genome sequencing and haplotyping of single human cells. *Proc Natl Acad Sci U S A* 2017;114:12512–7.
- [19] Xing D, Tan L, Chang CH, Li H, Xie XS. Accurate SNV detection in single cells by transposon-based whole-genome amplification of complementary strands. *Proc Natl Acad Sci U S A* 2021;118:e2013106118.
- [20] Blokzijl F, de Ligt J, Jager M, Sasselli V, Roerink S, Sasaki N, et al. Tissue-specific mutation accumulation in human adult stem cells during life. *Nature* 2016;538:260–4.
- [21] Lodato MA, Rodin RE, Bohrsen CL, Coulter ME, Barton AR, Kwon M, et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* 2018;359:555–9.
- [22] Kazazian Jr HH, Moran JV. Mobile DNA in health and disease. *N Engl J Med* 2017;377:361–70.
- [23] Ludwig LS, Lareau CA, Ulirsch JC, Christian E, Muus C, Li LH, et al. Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell* 2019;176:1325–39.e22.
- [24] Fan X, Yang C, Li W, Bai X, Zhou X, Xie H, et al. SMOOTH-seq: single-cell genome sequencing of human cells on a third-generation sequencing platform. *Genome Biol* 2021;22:195.
- [25] Kelsey G, Stegle O, Reik W. Single-cell epigenomics: recording the past and predicting the future. *Science* 2017;358:69–75.
- [26] Klemm SL, Shipony Z, Greenleaf WJ. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet* 2019;20:207–20.
- [27] Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, et al. The accessible chromatin landscape of the human genome. *Nature* 2012;489:75–82.
- [28] Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 2015;523:486–90.
- [29] Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* 2015;348:910–4.
- [30] Preissl S, Fang R, Huang H, Zhao Y, Raviram R, Gorkin DU, et al. Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nat Neurosci* 2018;21:432–9.
- [31] Lareau CA, Duarte FM, Chew JG, Kartha VK, Burkett ZD, Kohlway AS, et al. Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility. *Nat Biotechnol* 2019;37:916–24.
- [32] Satpathy AT, Granja JM, Yost KE, Qi Y, Meschi F, McDermott GP, et al. Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nat Biotechnol* 2019;37:925–36.
- [33] Xu W, Wen Y, Liang Y, Xu Q, Wang X, Jin W, et al. A plate-based single-cell ATAC-seq workflow for fast and robust profiling of chromatin accessibility. *Nat Protoc* 2021;16:4084–107.
- [34] Jin W, Tang Q, Wan M, Cui K, Zhang Y, Ren G, et al. Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. *Nature* 2015;528:142–6.
- [35] Lai B, Gao W, Cui K, Xie W, Tang Q, Jin W, et al. Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. *Nature* 2018;562:281–5.
- [36] Dekker J, Mirny L. The 3D genome as moderator of chromosomal communication. *Cell* 2016;164:1110–21.
- [37] Finn EH, Pegoraro G, Brandao HB, Valton AL, Oomen ME, Dekker J, et al. Extensive heterogeneity and intrinsic variation in spatial genome organization. *Cell* 2019;176:1502–15.e10.
- [38] Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, et al. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 2013;502:59–64.
- [39] Nagano T, Lubling Y, Yaffe E, Wingett SW, Dean W, Tanay A, et al. Single-cell Hi-C for genome-wide detection of chromatin interactions that occur simultaneously in a single cell. *Nat Protoc* 2015;10:1986–2003.
- [40] Flyamer IM, Gassler J, Imakaev M, Brandão HB, Ulianov SV, Abdennur N, et al. Single-nucleus Hi-C reveals unique chromatin

- reorganization at oocyte-to-zygote transition. *Nature* 2017;544:110–4.
- [41] Nagano T, Lubling Y, Várnai C, Dudley C, Leung W, Baran Y, et al. Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* 2017;547:61–7.
- [42] Ramani V, Deng X, Qiu R, Gunderson KL, Steemers FJ, Disteché CM, et al. Massively multiplex single-cell Hi-C. *Nat Methods* 2017;14:263–6.
- [43] Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* 2017;544:59–64.
- [44] Tan L, Xing D, Chang CH, Li H, Xie XS. Three-dimensional genome structures of single diploid human cells. *Science* 2018;361:924–8.
- [45] Tan L, Ma W, Wu H, Zheng Y, Xing D, Chen R, et al. Changes in genome architecture and transcriptional dynamics progress independently of sensory experience during post-natal brain development. *Cell* 2021;184:741–58.e17.
- [46] Arrastia MV, Jachowicz JW, Ollikainen N, Curtis MS, Lai C, Quinodoz SA, et al. Single-cell measurement of higher-order 3D genome organization with scSPRITE. *Nat Biotechnol* 2022;40:64–73.
- [47] Millan-Zambrano G, Burton A, Bannister AJ, Schneider R. Histone post-translational modifications — cause and consequence of genome function. *Nat Rev Genet* 2022;23:563–80.
- [48] Rotem A, Ram O, Shores N, Sperling RA, Goren A, Weitz DA, et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol* 2015;33:1165–72.
- [49] Ai S, Xiong H, Li CC, Luo Y, Shi Q, Liu Y, et al. Profiling chromatin states using single-cell HiChIP-seq. *Nat Cell Biol* 2019;21:1164–72.
- [50] Grosselin K, Durand A, Marsolier J, Poitou A, Marangoni E, Nemati F, et al. High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. *Nat Genet* 2019;51:1060–6.
- [51] Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* 2019;10:1930.
- [52] Wang Q, Xiong H, Ai S, Yu X, Liu Y, Zhang J, et al. CoBATCH for high-throughput single-cell epigenomic profiling. *Mol Cell* 2019;76:206–16.e7.
- [53] Carter B, Ku WL, Kang JY, Hu G, Perrie J, Tang Q, et al. Mapping histone modifications in low cell number and single cells using antibody-guided chromatin tagmentation (ACT-seq). *Nat Commun* 2019;10:3747.
- [54] Harada A, Maehara K, Handa T, Arimura Y, Nogami J, Hayashi-Takanaka Y, et al. A chromatin integration labelling method enables epigenomic profiling with lower input. *Nat Cell Biol* 2019;21:287–96.
- [55] Meers MP, Llagas G, Janssens DH, Codomo CA, Henikoff S. Multifactorial profiling of epigenetic landscapes at single-cell resolution using MulTI-Tag. *Nat Biotechnol* 2023;41:708–16.
- [56] Greenberg MVC, Bourchis D. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol* 2019;20:590–607.
- [57] Guo H, Zhu P, Wu X, Li X, Wen L, Tang F. Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res* 2013;23:2126–35.
- [58] Guo H, Zhu P, Guo F, Li X, Wu X, Fan X, et al. Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nat Protoc* 2015;10:645–59.
- [59] Luo C, Keown CL, Kurihara L, Zhou J, He Y, Li J, et al. Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. *Science* 2017;357:600–4.
- [60] Mulqueen RM, Pokholok D, Norberg SJ, Torkency KA, Fields AJ, Sun D, et al. Highly scalable generation of DNA methylation profiles in single cells. *Nat Biotechnol* 2018;36:428–31.
- [61] Smallwood SA, Lee HJ, Angermueller C, Krueger F, Saadeh H, Peat J, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat Methods* 2014;11:817–20.
- [62] Shareef SJ, Bevil SM, Raman AT, Aryee MJ, van Galen P, Hovestadt V, et al. Extended-representation bisulfite sequencing of gene regulatory elements in multiplexed samples and single cells. *Nat Biotechnol* 2021;39:1086–94.
- [63] Trapp A, Kerepesi C, Gladyshev VN. Profiling epigenetic age in single cells. *Nat Aging* 2021;1:1189–201.
- [64] Gabbutt C, Schenck RO, Weisenberger DJ, Kimberley C, Berner A, Househam J, et al. Fluctuating methylation clocks for cell lineage tracing at high temporal resolution in human tissues. *Nat Biotechnol* 2022;40:720–30.
- [65] Wang X, He Y, Zhang Q, Ren X, Zhang Z. Direct comparative analyses of 10X Genomics Chromium and Smart-seq2. *Genomics Proteomics Bioinformatics* 2021;19:253–66.
- [66] Picelli S, Bjorklund AK, 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.
- [67] Hagemann-Jensen M, Ziegenhain C, Chen P, Ramsköld D, Hendriks GJ, Larsson AJM, et al. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nat Biotechnol* 2020;38:708–14.
- [68] Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretsky I, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* 2014;343:776–9.
- [69] Fan X, Tang D, Liao Y, Li P, Zhang Y, Wang M, et al. Single-cell RNA-seq analysis of mouse preimplantation embryos by third-generation sequencing. *PLoS Biol* 2020;18:e3001017.
- [70] Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* 2015;161:1187–201.
- [71] Macosko EZ, Basu A, Satija R, Nemes J, Shekhar K, Goldman M, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 2015;161:1202–14.
- [72] Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun* 2017;8:14049.
- [73] Lahmemann D, Koster J, Szczurek E, McCarthy DJ, Hicks SC, Robinson MD, et al. Eleven grand challenges in single-cell data science. *Genome Biol* 2020;21:31.
- [74] Erhard F, Baptista MAP, Krammer T, Hennig T, Lange M, Arampatzi P, et al. scSLAM-seq reveals core features of transcription dynamics in single cells. *Nature* 2019;571:419–23.
- [75] Hendriks GJ, Jung LA, Larsson AJM, Lidschreiber M, Andersson Forsman O, Lidschreiber K, et al. NASC-seq monitors RNA synthesis in single cells. *Nat Commun* 2019;10:3138.
- [76] Cao J, Zhou W, Steemers F, Trapnell C, Shendure J. Sci-fate characterizes the dynamics of gene expression in single cells. *Nat Biotechnol* 2020;38:980–8.
- [77] Qiu Q, Hu P, Qiu X, Govek KW, Camara PG, Wu H. Massively parallel and time-resolved RNA sequencing in single cells with scNT-seq. *Nat Methods* 2020;17:991–1001.
- [78] Chen W, Guillaume-Gentil O, Rainer PY, Gäbelein CG, Saelens W, Gardeux V, et al. Live-seq enables temporal transcriptomic recording of single cells. *Nature* 2022;608:733–40.
- [79] Li X, Xiong X, Yi C. Epitranscriptome sequencing technologies: decoding RNA modifications. *Nat Methods* 2016;14:23–31.
- [80] Moshitch-Moshkovitz S, Dominissini D, Rechavi G. The epitranscriptome toolbox. *Cell* 2022;185:764–76.
- [81] Vistain LF, Tay S. Single-cell proteomics. *Trends Biochem Sci* 2021;46:661–72.

- [82] Zheng Y, Huang X, Kelleher NL. Epiproteomics: quantitative analysis of histone marks and codes by mass spectrometry. *Curr Opin Chem Biol* 2016;33:142–50.
- [83] Dey SS, Kester L, Spanjaard B, Bienko M, van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. *Nat Biotechnol* 2015;33:285–9.
- [84] Li W, Calder RB, Mar JC, Vиж J. Single-cell transcriptogenomics reveals transcriptional exclusion of ENU-mutated alleles. *Mutat Res* 2015;772:55–62.
- [85] Macaulay IC, Haerty W, Kumar P, Li YI, Hu TX, Teng MJ, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods* 2015;12:519–22.
- [86] Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods* 2016;13:229–32.
- [87] Chew LF, Courtis ET, Tan Y, Viswanathan R, Xing Q, Tan RZ, et al. Single-cell multimodal profiling reveals cellular epigenetic heterogeneity. *Nat Methods* 2016;13:833–6.
- [88] Dixit A, Parnas O, Li B, Chen J, Fulco CP, Jerby-Arnon L, et al. Perturb-seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* 2016;167:1853–66. e17.
- [89] Frei AP, Bava FA, Zunder ER, Hsieh EW, Chen SY, Nolan GP, et al. Highly multiplexed simultaneous detection of RNAs and proteins in single cells. *Nat Methods* 2016;13:269–75.
- [90] Genshaft AS, Li S, Gallant CJ, Darmanis S, Prakadan SM, Ziegler CG, et al. Multiplexed, targeted profiling of single-cell proteomes and transcriptomes in a single reaction. *Genome Biol* 2016;17:188.
- [91] Hou Y, Guo H, Cao C, Li X, Hu B, Zhu P, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res* 2016;26:304–19.
- [92] Hu Y, Huang K, An Q, Du G, Hu G, Xue J, et al. Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol* 2016;17:88.
- [93] Jaitin DA, Weiner A, Yofe I, Lara-Astiaso D, Keren-Shaul H, David E, et al. Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. *Cell* 2016;167:1883–96. e15.
- [94] Datlinger P, Rendeiro AF, Schmid C, Krausgruber T, Traxler P, Klughammer J, et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods* 2017;14:297–301.
- [95] Guo F, Li L, Li J, Wu X, Hu B, Zhu P, et al. Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. *Cell Res* 2017;27:967–88.
- [96] Peterson VM, Zhang KX, Kumar N, Wong J, Li L, Wilson DC, et al. Multiplexed quantification of proteins and transcripts in single cells. *Nat Biotechnol* 2017;35:936–9.
- [97] Pott S. Simultaneous measurement of chromatin accessibility, DNA methylation, and nucleosome phasing in single cells. *Elife* 2017;6:e23203.
- [98] Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* 2017;14:865–8.
- [99] Bian S, Hou Y, Zhou X, Li X, Yong J, Wang Y, et al. Single-cell multiomics sequencing and analyses of human colorectal cancer. *Science* 2018;362:1060–3.
- [100] Cao J, Cusanovich DA, Ramani V, Aghamirzaie D, Pliner HA, Hill AJ, et al. Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* 2018;361:1380–5.
- [101] Chen X, Litzenburger UM, Wei Y, Schep AN, LaGory EL, Choudhry H, et al. Joint single-cell DNA accessibility and protein epitope profiling reveals environmental regulation of epigenomic heterogeneity. *Nat Commun* 2018;9:4590.
- [102] Clark SJ, Argelaguet R, Kapourani CA, Stubbs TM, Lee HJ, Alda-Catalinas C, et al. scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat Commun* 2018;9:781.
- [103] Han KY, Kim KT, Joung JG, Son DS, Kim YJ, Jo A, et al. SIDR: simultaneous isolation and parallel sequencing of genomic DNA and total RNA from single cells. *Genome Res* 2018;28:75–87.
- [104] Lake BB, Chen S, Sos BC, Fan J, Kaeser GE, Yung YC, et al. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat Biotechnol* 2018;36:70–80.
- [105] Chen S, Lake BB, Zhang K. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nat Biotechnol* 2019;37:1452–7.
- [106] Gerlach JP, van Buggenum JAG, Tanis SEJ, Hogeweg M, Heuts BMH, Muraro MJ, et al. Combined quantification of intracellular (phospho-)proteins and transcriptomics from fixed single cells. *Sci Rep* 2019;9:1469.
- [107] Gu C, Liu S, Wu Q, Zhang L, Guo F. Integrative single-cell analysis of transcriptome, DNA methylome and chromatin accessibility in mouse oocytes. *Cell Res* 2019;29:110–23.
- [108] Lee DS, Luo C, Zhou J, Chandran S, Rivkin A, Bartlett A, et al. Simultaneous profiling of 3D genome structure and DNA methylation in single human cells. *Nat Methods* 2019;16:999–1006.
- [109] Li G, Liu Y, Zhang Y, Kubo N, Yu M, Fang R, et al. Joint profiling of DNA methylation and chromatin architecture in single cells. *Nat Methods* 2019;16:991–3.
- [110] Liu L, Liu C, Quintero A, Wu L, Yuan Y, Wang M, et al. Deconvolution of single-cell multi-omics layers reveals regulatory heterogeneity. *Nat Commun* 2019;10:470.
- [111] Mateo LJ, Murphy SE, Hafner A, Cinquini IS, Walker CA, Boettiger AN. Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* 2019;568:49–54.
- [112] Mimitou EP, Cheng A, Montalbano A, Hao S, Stoeckius M, Legut M, et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat Methods* 2019;16:409–12.
- [113] Rodriguez-Meira A, Buck G, Clark SA, Povinelli BJ, Alcolea V, Louka E, et al. Unravelling intratumoral heterogeneity through high-sensitivity single-cell mutational analysis and parallel RNA sequencing. *Mol Cell* 2019;73:1292–305.e8.
- [114] Rooijers K, Markodimitraki CM, Rang FJ, de Vries SS, Chialastri A, de Luca KL, et al. Simultaneous quantification of protein–DNA contacts and transcriptomes in single cells. *Nat Biotechnol* 2019;37:766–72.
- [115] Rubin AJ, Parker KR, Satpathy AT, Qi Y, Wu B, Ong AJ, et al. Coupled single-cell CRISPR screening and epigenomic profiling reveals causal gene regulatory networks. *Cell* 2019;176:361–76. e17.
- [116] Yin Y, Jiang Y, Lam KG, Berletch JB, Distchele CM, Noble WS, et al. High-throughput single-cell sequencing with linear amplification. *Mol Cell* 2019;76:676–90.e10.
- [117] Zhu C, Yu M, Huang H, Juric I, Abnoui A, Hu R, et al. An ultra high-throughput method for single-cell joint analysis of open chromatin and transcriptome. *Nat Struct Mol Biol* 2019;26:1063–70.
- [118] Ma S, Zhang B, LaFave LM, Earl AS, Chiang Z, Hu Y, et al. Chromatin potential identified by shared single-cell profiling of RNA and chromatin. *Cell* 2020;183:1103–16.e20.
- [119] Xing QR, Farran CAE, Zeng YY, Yi Y, Warriar T, Gautam P, et al. Parallel bimodal single-cell sequencing of transcriptome and chromatin accessibility. *Genome Res* 2020;30:1027–39.
- [120] Zachariadis V, Cheng H, Andrews N, Enge M. A highly scalable method for joint whole-genome sequencing and gene-expression profiling of single cells. *Mol Cell* 2020;80:541–53.e5.
- [121] Chung H, Parkhurst CN, Magee EM, Phillips D, Habibi E, Chen F, et al. Joint single-cell measurements of nuclear proteins and RNA *in vivo*. *Nat Methods* 2021;18:1204–12.

- [122] Gu H, Raman AT, Wang X, Gaiti F, Chaligne R, Mohammad AW, et al. Smart-RRBS for single-cell methylome and transcriptome analysis. *Nat Protoc* 2021;16:4004–30.
- [123] Hwang B, Lee DS, Tamaki W, Sun Y, Ogorodnikov A, Hartoularos GC, et al. SCITO-seq: single-cell combinatorial indexed cytometry sequencing. *Nat Methods* 2021;18:903–11.
- [124] Liscovitch-Brauer N, Montalbano A, Deng J, Méndez-Mancilla A, Wessels HH, Moss NG, et al. Profiling the genetic determinants of chromatin accessibility with scalable single-cell CRISPR screens. *Nat Biotechnol* 2021;39:1270–7.
- [125] Mimitou EP, Lareau CA, Chen KY, Zorzetto-Fernandes AL, Hao Y, Takeshima Y, et al. Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat Biotechnol* 2021;39:1246–58.
- [126] Pierce SE, Granja JM, Greenleaf WJ. High-throughput single-cell chromatin accessibility CRISPR screens enable unbiased identification of regulatory networks in cancer. *Nat Commun* 2021;12:2969.
- [127] Plongthongkum N, Diep D, Chen S, Lake BB, Zhang K. Scalable dual-omics profiling with single-nucleus chromatin accessibility and mRNA expression sequencing 2 (SNARE-seq2). *Nat Protoc* 2021;16:4992–5029.
- [128] Reimegård J, Tarbier M, Danielsson M, Schuster J, Baskaran S, Panagiotou S, et al. A combined approach for single-cell mRNA and intracellular protein expression analysis. *Commun Biol* 2021;4:624.
- [129] Sun Z, Tang Y, Zhang Y, Fang Y, Jia J, Zeng W, et al. Joint single-cell multiomic analysis in Wnt3a induced asymmetric stem cell division. *Nat Commun* 2021;12:5941.
- [130] Swanson E, Lord C, Reading J, Heubeck AT, Genge PC, Thomson Z, et al. Simultaneous trimodal single-cell measurement of transcripts, epitopes, and chromatin accessibility using TEA-seq. *Elife* 2021;10:e63632.
- [131] Wang Y, Yuan P, Yan Z, Yang M, Huo Y, Nie Y, et al. Single-cell multiomics sequencing reveals the functional regulatory landscape of early embryos. *Nat Commun* 2021;12:1247.
- [132] Xiong H, Luo Y, Wang Q, Yu X, He A. Single-cell joint detection of chromatin occupancy and transcriptome enables higher-dimensional epigenomic reconstructions. *Nat Methods* 2021;18:652–60.
- [133] Yan R, Gu C, You D, Huang Z, Qian J, Yang Q, et al. Decoding dynamic epigenetic landscapes in human oocytes using single-cell multi-omics sequencing. *Cell Stem Cell* 2021;28:1641–56.e7.
- [134] Zhu C, Zhang Y, Li YE, Lucero J, Behrens MM, Ren B. Joint profiling of histone modifications and transcriptome in single cells from mouse brain. *Nat Methods* 2021;18:283–92.
- [135] Bartosovic M, Castelo-Branco G. Multimodal chromatin profiling using nanobody-based single-cell CUT&Tag. *Nat Biotechnol* 2023;41:794–805.
- [136] Chen AF, Parks B, Kathiria AS, Ober-Reynolds B, Goronzy JJ, Greenleaf WJ. NEAT-seq: simultaneous profiling of intranuclear proteins, chromatin accessibility and gene expression in single cells. *Nat Methods* 2022;19:547–53.
- [137] Fiskin E, Lareau CA, Ludwig LS, Eraslan G, Liu F, Ring AM, et al. Single-cell profiling of proteins and chromatin accessibility using PHAGE-ATAC. *Nat Biotechnol* 2022;40:374–81.
- [138] Luo C, Liu H, Xie F, Armand EJ, Siletti K, Bakken TE, et al. Single nucleus multi-omics identifies human cortical cell regulatory genome diversity. *Cell Genom* 2022;2:100107.
- [139] Pan L, Ku WL, Tang Q, Cao Y, Zhao K. scPCOR-seq enables co-profiling of chromatin occupancy and RNAs in single cells. *Commun Biol* 2022;5:678.
- [140] Rang FJ, de Luca KL, de Vries SS, Valdes-Quezada C, Boele E, Nguyen PD, et al. Single-cell profiling of transcriptome and histone modifications with EpiDamID. *Mol Cell* 2022;82:1956–70.e14.
- [141] Stuart T, Hao S, Zhang B, Mekerishvili L, Landau DA, Maniatis S, et al. Nanobody-tethered transposition enables multifactorial chromatin profiling at single-cell resolution. *Nat Biotechnol* 2023;41:806–12.
- [142] Tedesco M, Giannese F, Lazarević D, Giansanti V, Rosano D, Monzani S, et al. Chromatin velocity reveals epigenetic dynamics by single-cell profiling of heterochromatin and euchromatin. *Nat Biotechnol* 2022;40:235–44.
- [143] Vistain L, Van Phan H, Keisham B, Jordi C, Chen M, Reddy ST, et al. Quantification of extracellular proteins, protein complexes and mRNAs in single cells by proximity sequencing. *Nat Methods* 2022;19:1578–89.
- [144] Xu W, Yang W, Zhang Y, Chen Y, Hong N, Zhang Q, et al. ISSAAC-seq enables sensitive and flexible multimodal profiling of chromatin accessibility and gene expression in single cells. *Nat Methods* 2022;19:1243–9.
- [145] Zhang B, Srivastava A, Mimitou E, Stuart T, Raimondi I, Hao Y, et al. Characterizing cellular heterogeneity in chromatin state with scCUT&Tag-pro. *Nat Biotechnol* 2022;40:1220–30.
- [146] Hu Y, Jiang Z, Chen K, Zhou Z, Zhou X, Wang Y, et al. scNanoATAC-seq: a long-read single-cell ATAC sequencing method to detect chromatin accessibility and genetic variants simultaneously within an individual cell. *Cell Res* 2023;33:83–6.
- [147] Yu L, Wang X, Mu Q, Tam SST, Loi DSC, Chan AKY, et al. scONE-seq: a single-cell multi-omics method enables simultaneous dissection of phenotype and genotype heterogeneity from frozen tumors. *Sci Adv* 2023;9:eabp8901.
- [148] Zappia L, Phipson B, Oshlack A. Exploring the single-cell RNA-seq analysis landscape with the scRNA-tools database. *PLoS Comput Biol* 2018;14:e1006245.
- [149] Heumos L, Schaar AC, Lance C, Litinetskaya A, Drost F, Zappia L, et al. Best practices for single-cell analysis across modalities. *Nat Rev Genet* 2023;24:550–72.
- [150] Melsted P, Boeshaghi AS, Liu L, Gao F, Lu L, Min KHJ, et al. Modular, efficient and constant-memory single-cell RNA-seq preprocessing. *Nat Biotechnol* 2021;39:813–8.
- [151] McGinnis CS, Murrow LM, Gartner ZJ. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst* 2019;8:329–37.
- [152] Wolock SL, Lopez R, Klein AM. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst* 2019;8:281–91.
- [153] DePasquale EAK, Schnell DJ, Van Camp PJ, Valiente-Alandi I, Blaxall BC, Grimes HL, et al. DoubletDecon: deconvoluting doublets from single-cell RNA-sequencing data. *Cell Rep* 2019;29:1718–27.
- [154] Xi NM, Li JJ. Benchmarking computational doublet-detection methods for single-cell RNA sequencing data. *Cell Syst* 2021;12:176–94.
- [155] Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol* 2019;15:e8746.
- [156] Hao Y, Hao S, Andersen-Nissen E, Mauck 3rd WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell* 2021;184:3573–87.e29.
- [157] Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* 2018;19:15.
- [158] Bacher R, Chu LF, Leng N, Gasch AP, Thomson JA, Stewart RM, et al. SCnorm: robust normalization of single-cell RNA-seq data. *Nat Methods* 2017;14:584–6.
- [159] Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* 2019;20:296.
- [160] 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.
- [161] Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. *Nat Methods* 2018;15:1053–8.

- [162] Kiselev VY, Kirschner K, Schaub MT, Andrews T, Yiu A, Chandra T, et al. SC3: consensus clustering of single-cell RNA-seq data. *Nat Methods* 2017;14:483–6.
- [163] Andrews TS, Hemberg M. M3Drop: dropout-based feature selection for scRNA-seq. *Bioinformatics* 2019;35:2865–7.
- [164] Jiang L, Chen H, Pinello L, Yuan GC. GiniClust: detecting rare cell types from single-cell gene expression data with Gini index. *Genome Biol* 2016;17:144.
- [165] Ding J, Adiconis X, Simmons SK, Kowalczyk MS, Hession CC, Marjanovic ND, et al. Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nat Biotechnol* 2020;38:737–46.
- [166] van Dijk D, Sharma R, Nainys J, Yim K, Kathail P, Carr AJ, et al. Recovering gene interactions from single-cell data using data diffusion. *Cell* 2018;174:716–29.e27.
- [167] Zhang Z, Cui F, Wang C, Zhao L, Zou Q. Goals and approaches for each processing step for single-cell RNA sequencing data. *Brief Bioinform* 2021;22:bbaa314.
- [168] Li WV, Li JJ. An accurate and robust imputation method scImpute for single-cell RNA-seq data. *Nat Commun* 2018;9:997.
- [169] Huang M, Wang J, Torre E, Dueck H, Shaffer S, Bonasio R, et al. SAVER: gene expression recovery for single-cell RNA sequencing. *Nat Methods* 2018;15:539–42.
- [170] Linderman GC, Zhao J, Roulis M, Bielecki P, Flavell RA, Nadler B, et al. Zero-preserving imputation of single-cell RNA-seq data. *Nat Commun* 2022;13:192.
- [171] Jiang R, Sun T, Song D, Li JJ. Statistics or biology: the zero-inflation controversy about scRNA-seq data. *Genome Biol* 2022;23:31.
- [172] Svensson V. Droplet scRNA-seq is not zero-inflated. *Nat Biotechnol* 2020;38:147–50.
- [173] Cao Y, Kitanovski S, Kupperts R, Hoffmann D. UMI or not UMI, that is the question for scRNA-seq zero-inflation. *Nat Biotechnol* 2021;39:158–9.
- [174] Baran Y, Bercovich A, Sebe-Pedros A, Lubling Y, Giladi A, Chomsky E, et al. MetaCell: analysis of single-cell RNA-seq data using *K*-nn graph partitions. *Genome Biol* 2019;20:206.
- [175] Persad S, Choo ZN, Dien C, Sohail N, Masilionis I, Chaligne R, et al. SEACells infers transcriptional and epigenomic cellular states from single-cell genomics data. *Nat Biotechnol* 2023;41:1746–57.
- [176] Luecken MD, Buttner M, Chaichoompu K, Danese A, Interlandi M, Mueller MF, et al. Benchmarking atlas-level data integration in single-cell genomics. *Nat Methods* 2022;19:41–50.
- [177] Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* 2019;16:1289–96.
- [178] Hie B, Bryson B, Berger B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. *Nat Biotechnol* 2019;37:685–91.
- [179] Xu C, Lopez R, Mehlman E, Regier J, Jordan MI, Yosef N. Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. *Mol Syst Biol* 2021;17:e9620.
- [180] Lotfollahi M, Wolf FA, Theis FJ. scGen predicts single-cell perturbation responses. *Nat Methods* 2019;16:715–21.
- [181] van der Maaten L, Hinton GE. Visualizing data using t-SNE. *J Mach Learn Res* 2008;9:2579–605.
- [182] McInnes L, Healy J, Saul N, Großberger L. UMAP: uniform manifold approximation and projection. *J Open Source Softw* 2018;3:861.
- [183] Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* 2019;37:38–44.
- [184] Liu B, Li C, Li Z, Wang D, Ren X, Zhang Z. An entropy-based metric for assessing the purity of single cell populations. *Nat Commun* 2020;11:3155.
- [185] Hu C, Li T, Xu Y, Zhang X, Li F, Bai J, et al. Cell Marker 2.0: an updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. *Nucleic Acids Res* 2023;51:D870–6.
- [186] Franzen O, Gan LM, Björkegren JLM. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database (Oxford)* 2019;2019:baz046.
- [187] Zhang AW, O’Flanagan C, Chavez EA, Lim JLP, Ceglia N, McPherson A, et al. Probabilistic cell-type assignment of single-cell RNA-seq for tumor microenvironment profiling. *Nat Methods* 2019;16:1007–15.
- [188] Guo H, Li J. scSorter: assigning cells to known cell types according to marker genes. *Genome Biol* 2021;22:69.
- [189] Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, et al. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* 2019;20:163–72.
- [190] Li C, Liu B, Kang B, Liu Z, Liu Y, Chen C, et al. SciBet as a portable and fast single cell type identifier. *Nat Commun* 2020;11:1818.
- [191] Domínguez Conde C, Xu C, Jarvis LB, Rainbow DB, Wells SB, Gomes T, et al. Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science* 2022;376:eab15197.
- [192] Lotfollahi M, Naghipourfar M, Luecken MD, Khajavi M, Buttner M, Wagenstetter M, et al. Mapping single-cell data to reference atlases by transfer learning. *Nat Biotechnol* 2022;40:121–30.
- [193] Chen J, Xu H, Tao W, Chen Z, Zhao Y, Han JJ. Transformer for one stop interpretable cell type annotation. *Nat Commun* 2023;14:223.
- [194] Yang F, Wang W, Wang F, Fang Y, Tang D, Huang J, et al. scBERT as a large-scale pretrained deep language model for cell type annotation of single-cell RNA-seq data. *Nat Mach Intell* 2022;4:852–66.
- [195] Petukhov V, Igolkina A, Rydbirk R, Mei S, Christoffersen L, Khodosevich K, et al. Case-control analysis of single-cell RNA-seq studies. *bioRxiv* 2022;484475.
- [196] Sonesson C, Robinson MD. Bias, robustness and scalability in single-cell differential expression analysis. *Nat Methods* 2018;15:255–61.
- [197] Squair JW, Gautier M, Kathe C, Anderson MA, James ND, Hutson TH, et al. Confronting false discoveries in single-cell differential expression. *Nat Commun* 2021;12:5692.
- [198] Buttner M, Ostner J, Müller CL, Theis FJ, Schubert B. scCODA is a Bayesian model for compositional single-cell data analysis. *Nat Commun* 2021;12:6876.
- [199] Saelens W, Cannoodt R, Todorov H, Saey Y. A comparison of single-cell trajectory inference methods. *Nat Biotechnol* 2019;37:547–54.
- [200] Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* 2017;14:979–82.
- [201] Street K, Rizzo D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* 2018;19:477.
- [202] Wolf FA, Hamey FK, Plass M, Solana J, Dahlin JS, Gottgens B, et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol* 2019;20:59.
- [203] Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* 2020;38:1408–14.
- [204] Lange M, Bergen V, Klein M, Setty M, Reuter B, Bakhti M, et al. CellRank for directed single-cell fate mapping. *Nat Methods* 2022;19:159–70.
- [205] Gao M, Qiao C, Huang Y. UniTVelo: temporally unified RNA velocity reinforces single-cell trajectory inference. *Nat Commun* 2022;13:6586.

- [206] Alpert A, Moore LS, Dubovik T, Shen-Orr SS. Alignment of single-cell trajectories to compare cellular expression dynamics. *Nat Methods* 2018;15:267–70.
- [207] Sugihara R, Kato Y, Mori T, Kawahara Y. Alignment of single-cell trajectory trees with CAPITAL. *Nat Commun* 2022;13:5972.
- [208] Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 2017;14:1083–6.
- [209] Pratapa A, Jalihal AP, Law JN, Bharadwaj A, Murali TM. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nat Methods* 2020;17:147–54.
- [210] Chan TE, Stumpf MPH, Babbie AC. Gene regulatory network inference from single-cell data using multivariate information measures. *Cell Syst* 2017;5:251–67.
- [211] Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. Inferring regulatory networks from expression data using tree-based methods. *PLoS One* 2010;5:e12776.
- [212] Chen S, Mar JC. Evaluating methods of inferring gene regulatory networks highlights their lack of performance for single cell gene expression data. *BMC Bioinformatics* 2018;19:232.
- [213] Raharinirina NA, Peppert F, von Kleist M, Schutte C, Sunkara V. Inferring gene regulatory networks from single-cell RNA-seq temporal snapshot data requires higher-order moments. *Patterns (N Y)* 2021;2:100332.
- [214] González-Blas CB, De Winter S, Hulselmans G, Hecker N, Matetovici I, Christiaens V, et al. SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nat Methods* 2023;20:1355–67.
- [215] Kamal A, Arnold C, Claringbould A, Moussa R, Servaas NH, Kholmatov M, et al. GRaNIE and GRaNPA: inference and evaluation of enhancer-mediated gene regulatory networks. *Mol Syst Biol* 2023;19:e11627.
- [216] Fleck JS, Jansen SMJ, Wollny D, Zenk F, Seimiya M, Jain A, et al. Inferring and perturbing cell fate regulomes in human brain organoids. *Nature* 2023;621:365–72.
- [217] Kartha VK, Duarte FM, Hu Y, Ma S, Chew JG, Lareau CA, et al. Functional inference of gene regulation using single-cell multi-omics. *Cell Genom* 2022;2:100166.
- [218] Lynch AW, Theodoris CV, Long HW, Brown M, Liu XS, Meyer CA. MIRA: joint regulatory modeling of multimodal expression and chromatin accessibility in single cells. *Nat Methods* 2022;19:1097–108.
- [219] Garcia-Alonso L, Handfield LF, Roberts K, Nikolakopoulou K, Fernando RC, Gardner L, et al. Mapping the temporal and spatial dynamics of the human endometrium *in vivo* and *in vitro*. *Nat Genet* 2021;53:1698–711.
- [220] Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell–cell communication using Cell Chat. *Nat Commun* 2021;12:1088.
- [221] Liu Z, Sun D, Wang C. Evaluation of cell–cell interaction methods by integrating single-cell RNA sequencing data with spatial information. *Genome Biol* 2022;23:218.
- [222] Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods* 2020;17:159–62.
- [223] Jiang P, Zhang Y, Ru B, Yang Y, Vu T, Paul R, et al. Systematic investigation of cytokine signaling activity at the tissue and single-cell levels. *Nat Methods* 2021;18:1181–91.
- [224] Ren X, Zhong G, Zhang Q, Zhang L, Sun Y, Zhang Z. Reconstruction of cell spatial organization from single-cell RNA sequencing data based on ligand–receptor mediated self-assembly. *Cell Res* 2020;30:763–78.
- [225] Nitzan M, Karaiskos N, Friedman N, Rajewsky N. Gene expression cartography. *Nature* 2019;576:132–7.
- [226] Jerby-Arnon L, Regev A. DIALOGUE maps multicellular programs in tissue from single-cell or spatial transcriptomics data. *Nat Biotechnol* 2022;40:1467–77.
- [227] Li B, Zhang W, Guo C, Xu H, Li L, Fang M, et al. Benchmarking spatial and single-cell transcriptomics integration methods for transcript distribution prediction and cell type deconvolution. *Nat Methods* 2022;19:662–70.
- [228] Lopez R, Nazaret A, Langevin M, Samaran J, Regier J, Jordan MI, et al. A joint model of unpaired data from scRNA-seq and spatial transcriptomics for imputing missing gene expression measurements. *arXiv* 2019;1905.02269.
- [229] Cable DM, Murray E, Zou LS, Goeva A, Macosko EZ, Chen F, et al. Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat Biotechnol* 2022;40:517–26.
- [230] Kleshchevnikov V, Shmatko A, Dann E, Aivazidis A, King HW, Li T, et al. Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat Biotechnol* 2022;40:661–71.
- [231] Stuart T, Srivastava A, Madad S, Lareau CA, Satija R. Single-cell chromatin state analysis with Signac. *Nat Methods* 2021;18:1333–41.
- [232] Welch JD, Kozareva V, Ferreira A, Vanderburg C, Martin C, Macosko EZ. Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* 2019;177:1873–87. e17.
- [233] Dou J, Liang S, Mohanty V, Miao Q, Huang Y, Liang Q, et al. Bi-order multimodal integration of single-cell data. *Genome Biol* 2022;23:112.
- [234] Cao ZJ, Gao G. Multi-omics single-cell data integration and regulatory inference with graph-linked embedding. *Nat Biotechnol* 2022;40:1458–66.
- [235] Argelaguet R, Arnol D, Bredikhin D, Deloro Y, Velten B, Marioni JC, et al. MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol* 2020;21:111.
- [236] Ashuach T, Gabitto MI, Koodli RV, Saldi GA, Jordan MI, Yosef N. MultiVI: deep generative model for the integration of multimodal data. *Nat Methods* 2023;20:1222–31.
- [237] Villani AC, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* 2017;356:eaah4573.
- [238] Ballesteros I, Rubio-Ponce A, Genua M, Lusito E, Kwok I, Fernandez-Calvo G, et al. Co-option of neutrophil fates by tissue environments. *Cell* 2020;183:1282–97. e18.
- [239] Xie X, Shi Q, Wu P, Zhang X, Kambara H, Su J, et al. Single-cell transcriptome profiling reveals neutrophil heterogeneity in homeostasis and infection. *Nat Immunol* 2020;21:1119–33.
- [240] Sas AR, Carbajal KS, Jerome AD, Menon R, Yoon C, Kalinski AL, et al. A new neutrophil subset promotes CNS neuron survival and axon regeneration. *Nat Immunol* 2020;21:1496–505.
- [241] Velten L, Haas SF, Raffel S, Blaszkiewicz S, Islam S, Hennig BP, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol* 2017;19:271–81.
- [242] Litvinukova M, Talavera-Lopez C, Maatz H, Reichart D, Worth CL, Lindberg EL, et al. Cells of the adult human heart. *Nature* 2020;588:466–72.
- [243] Stewart BJ, Ferdinand JR, Young MD, Mitchell TJ, Loudon KW, Riding AM, et al. Spatiotemporal immune zonation of the human kidney. *Science* 2019;365:1461–6.
- [244] Eraslan G, Drokhyansky E, Anand S, Fiskin E, Subramanian A, Slyper M, et al. Single-nucleus cross-tissue molecular reference maps toward understanding disease gene function. *Science* 2022;376:eab4290.
- [245] Suo C, Dann E, Goh I, Jardine L, Kleshchevnikov V, Park JE, et al. Mapping the developing human immune system across organs. *Science* 2022;376:eabo0510.

- [246] Tabula Sapiens Consortium, Jones RC, Karkanias J, Krasnow MA, Pisco AO, Quake SR, et al. The Tabula Sapiens: a multiple-organ, single-cell transcriptomic atlas of humans. *Science* 2022;376:eabl4896.
- [247] Liu Z, Zhang Z. Mapping cell types across human tissues. *Science* 2022;376:695–6.
- [248] Gur C, Wang SY, Sheban F, Zada M, Li B, Kharouf F, et al. LGR5 expressing skin fibroblasts define a major cellular hub perturbed in scleroderma. *Cell* 2022;185:1373–88.e20.
- [249] Huang B, Chen Z, Geng L, Wang J, Liang H, Cao Y, et al. Mucosal profiling of pediatric-onset colitis and IBD reveals common pathogenics and therapeutic pathways. *Cell* 2019;179:1160–76.e24.
- [250] Martin JC, Chang C, Boschetti G, Ungaro R, Giri M, Grout JA, et al. Single-cell analysis of Crohn's disease lesions identifies a pathogenic cellular module associated with resistance to anti-TNF therapy. *Cell* 2019;178:1493–508.e20.
- [251] Winkler EA, Kim CN, Ross JM, Garcia JH, Gil E, Oh I, et al. A single-cell atlas of the normal and malformed human brain vasculature. *Science* 2022;375:eabi7377.
- [252] Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* 2019;570:332–7.
- [253] Cuomo ASE, Nathan A, Raychaudhuri S, MacArthur DG, Powell JE. Single-cell genomics meets human genetics. *Nat Rev Genet* 2023;24:535–49.
- [254] Soskic B, Cano-Gamez E, Smyth DJ, Ambridge K, Ke Z, Matte JC, et al. Immune disease risk variants regulate gene expression dynamics during CD4⁺ T cell activation. *Nat Genet* 2022;54:817–26.
- [255] Nathan A, Asgari S, Ishigaki K, Valencia C, Amariuta T, Luo Y, et al. Single-cell eQTL models reveal dynamic T cell state dependence of disease loci. *Nature* 2022;606:120–8.
- [256] Perez RK, Gordon MG, Subramaniam M, Kim MC, Hartoularos GC, Targ S, et al. Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. *Science* 2022;376:eabf1970.
- [257] Yazar S, Alquicira-Hernandez J, Wing K, Senabouth A, Gordon MG, Andersen S, et al. Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. *Science* 2022;376:eabf3041.
- [258] Jagadeesh KA, Dey KK, Montoro DT, Mohan R, Gazal S, Engreitz JM, et al. Identifying disease-critical cell types and cellular processes by integrating single-cell RNA-sequencing and human genetics. *Nat Genet* 2022;54:1479–92.
- [259] Zhang MJ, Hou K, Dey KK, Sakaue S, Jagadeesh KA, Weinand K, et al. Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data. *Nat Genet* 2022;54:1572–80.
- [260] Rood JE, Maartens A, Hupalowska A, Teichmann SA, Regev A. Impact of the Human Cell Atlas on medicine. *Nat Med* 2022;28:2486–96.
- [261] Sungnak W, Huang N, Becavin C, Berg M, Queen R, Litvinukova M, et al. SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat Med* 2020;26:681–7.
- [262] Ren X, Wen W, Fan X, Hou W, Su B, Cai P, et al. COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. *Cell* 2021;184:1895–913.e19.
- [263] Cao Y, Su B, Guo X, Sun W, Deng Y, Bao L, et al. Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of convalescent patients' B cells. *Cell* 2020;182:73–84.e16.
- [264] Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, et al. The Human Cell Atlas. *Elife* 2017;6:e27041.
- [265] Rozenblatt-Rosen O, Stubbington MJT, Regev A, Teichmann SA. The Human Cell Atlas: from vision to reality. *Nature* 2017;550:451–3.
- [266] HuBMAP Consortium. The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. *Nature* 2019;574:187–92.
- [267] Borner K, Teichmann SA, Quardokus EM, Gee JC, Browne K, Osumi-Sutherland D, et al. Anatomical structures, cell types and biomarkers of the Human Reference Atlas. *Nat Cell Biol* 2021;23:1117–28.
- [268] Hanahan D. Hallmarks of cancer: new dimensions. *Cancer Discov* 2022;12:31–46.
- [269] Sade-Feldman M, Yizhak K, Bjorgaard SL, Ray JP, de Boer CG, Jenkins RW, et al. Defining T cell states associated with response to checkpoint immunotherapy in melanoma. *Cell* 2018;175:998–1013.e20.
- [270] Ho YJ, Anaparthi N, Molik D, Mathew G, Aicher T, Patel A, et al. Single-cell RNA-seq analysis identifies markers of resistance to targeted BRAF inhibitors in melanoma cell populations. *Genome Res* 2018;28:1353–63.
- [271] Durante MA, Rodriguez DA, Kurtenbach S, Kuznetsov JN, Sanchez MI, Decatur CL, et al. Single-cell analysis reveals new evolutionary complexity in uveal melanoma. *Nat Commun* 2020;11:496.
- [272] Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 2016;352:189–96.
- [273] Tirosh I, Venteicher AS, Hebert C, Escalante LE, Patel AP, Yizhak K, et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* 2016;539:309–13.
- [274] Filbin MG, Tirosh I, Hovestadt V, Shaw ML, Escalante LE, Mathewson ND, et al. Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. *Science* 2018;360:331–5.
- [275] Yuan J, Levitin HM, Frattini V, Bush EC, Boyett DM, Samanamud J, et al. Single-cell transcriptome analysis of lineage diversity in high-grade glioma. *Genome Med* 2018;10:57.
- [276] Zhang Y, Chen H, Mo H, Hu X, Gao R, Zhao Y, et al. Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. *Cancer Cell* 2021;39:1578–93.e8.
- [277] Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* 2018;174:1293–308.e36.
- [278] Nalio Ramos R, Missolo-Koussou Y, Gerber-Ferder Y, Bromley CP, Bugatti M, Nunez NG, et al. Tissue-resident FOLR2⁺ macrophages associate with CD8⁺ T cell infiltration in human breast cancer. *Cell* 2022;185:1189–207.e25.
- [279] Wagner J, Rapsomaniki MA, Chevrier S, Anzeneder T, Langwieder C, Dykgers A, et al. A single-cell atlas of the tumor and immune ecosystem of human breast cancer. *Cell* 2019;177:1330–45.e18.
- [280] Pal B, Chen Y, Vaillant F, Capaldo BD, Joyce R, Song X, et al. A single-cell RNA expression atlas of normal, preneoplastic and tumorigenic states in the human breast. *EMBO J* 2021;40:e107333.
- [281] Savas P, Virassamy B, Ye C, Salim A, Mintoff CP, Caramia F, et al. Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. *Nat Med* 2018;24:986–93.
- [282] Liu Y, Zhang Q, Xing B, Luo N, Gao R, Yu K, et al. Immune phenotypic linkage between colorectal cancer and liver metastasis. *Cancer Cell* 2022;40:424–37.e5.
- [283] Chen B, Scurrah CR, McKinley ET, Simmons AJ, Ramirez-Solano MA, Zhu X, et al. Differential pre-malignant programs and microenvironment chart distinct paths to malignancy in human colorectal polyps. *Cell* 2021;184:6262–80.e26.

- [284] Pelka K, Hofree M, Chen JH, Sarkizova S, Pirl JD, Jorgji V, et al. Spatially organized multicellular immune hubs in human colorectal cancer. *Cell* 2021;184:4734–52.e20.
- [285] Becker WR, Nevins SA, Chen DC, Chiu R, Horning AM, Guha TK, et al. Single-cell analyses define a continuum of cell state and composition changes in the malignant transformation of polyps to colorectal cancer. *Nat Genet* 2022;54:985–95.
- [286] Joanito I, Wirapati P, Zhao N, Nawaz Z, Yeo G, Lee F, et al. Single-cell and bulk transcriptome sequencing identifies two epithelial tumor cell states and refines the consensus molecular classification of colorectal cancer. *Nat Genet* 2022;54:963–75.
- [287] Zhang L, Yu X, Zheng L, Zhang Y, Li Y, Fang Q, et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature* 2018;564:268–72.
- [288] Zhang L, Li Z, Skrzypczynska KM, Fang Q, Zhang W, O'Brien SA, et al. Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. *Cell* 2020;181:442–59.e29.
- [289] Kumar V, Ramnarayanan K, Sundar R, Padmanabhan N, Srivastava S, Koiwa M, et al. Single-cell atlas of lineage states, tumor microenvironment, and subtype-specific expression programs in gastric cancer. *Cancer Discov* 2022;12:670–91.
- [290] Kang B, Camps J, Fan B, Jiang H, Ibrahim MM, Hu X, et al. Parallel single-cell and bulk transcriptome analyses reveal key features of the gastric tumor microenvironment. *Genome Biol* 2022;23:265.
- [291] Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X, et al. Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. *Cell* 2017;169:1342–56.e16.
- [292] Xue R, Zhang Q, Cao Q, Kong R, Xiang X, Liu H, et al. Liver tumour immune microenvironment subtypes and neutrophil heterogeneity. *Nature* 2022;612:141–7.
- [293] Zhang Q, He Y, Luo N, Patel SJ, Han Y, Gao R, et al. Landscape and dynamics of single immune cells in hepatocellular carcinoma. *Cell* 2019;179:829–45.e20.
- [294] Young MD, Mitchell TJ, Vieira Braga FA, Tran MGB, Stewart BJ, Ferdinand JR, et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* 2018;361:594–9.
- [295] Enge M, Arda HE, Mignardi M, Beausang J, Bottino R, Kim SK, et al. Single-cell analysis of human pancreas reveals transcriptional signatures of aging and somatic mutation patterns. *Cell* 2017;171:321–30.e14.
- [296] Peng J, Sun BF, Chen CY, Zhou JY, Chen YS, Chen H, et al. Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res* 2019;29:725–38.
- [297] Leader AM, Grout JA, Maier BB, Nabet BY, Park MD, Tabachnikova A, et al. Single-cell analysis of human non-small cell lung cancer lesions refines tumor classification and patient stratification. *Cancer Cell* 2021;39:1594–609.e12.
- [298] Salcher S, Sturm G, Horvath L, Untergasser G, Kuempers C, Fotakis G, et al. High-resolution single-cell atlas reveals diversity and plasticity of tissue-resident neutrophils in non-small cell lung cancer. *Cancer Cell* 2022;40:1503–20.e8.
- [299] Zilionis R, Engblom C, Pfirschke C, Savova V, Zemmour D, Saaticioglu HD, et al. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* 2019;50:1317–34.
- [300] Liu B, Hu X, Feng K, Gao R, Xue Z, Zhang S, et al. Temporal single-cell tracing reveals clonal revival and expansion of precursor exhausted T cells during anti-PD-1 therapy in lung cancer. *Nat Cancer* 2022;3:108–21.
- [301] Wu F, Fan J, He Y, Xiong A, Yu J, Li Y, et al. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat Commun* 2021;12:2540.
- [302] Guo X, Zhang Y, Zheng L, Zheng C, Song J, Zhang Q, et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med* 2018;24:978–85.
- [303] Lambrechts D, Wauters E, Boeckx B, Aibar S, Nittner D, Burton O, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med* 2018;24:1277–89.
- [304] Maier B, Leader AM, Chen ST, Tung N, Chang C, LeBerichel J, et al. A conserved dendritic-cell regulatory program limits antitumour immunity. *Nature* 2020;580:257–62.
- [305] Blank CU, Haining WN, Held W, Hogan PG, Kallies A, Lugli E, et al. Defining “T cell exhaustion”. *Nat Rev Immunol* 2019;19:665–74.
- [306] 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.
- [307] Cheng S, Li Z, Gao R, Xing B, Gao Y, Yang Y, et al. A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell* 2021;184:792–809.e23.
- [308] Zheng L, Qin S, Si W, Wang A, Xing B, Gao R, et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* 2021;374:abe6474.
- [309] Krishnamurthy AT, Shyer JA, Thai M, Gandham V, Buechler MB, Yang YA, et al. LRRRC15⁺ myofibroblasts dictate the stromal setpoint to suppress tumour immunity. *Nature* 2022;611:148–54.
- [310] Topalian SL, Drake CG, Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell* 2015;27:450–61.
- [311] Liu B, Zhang Y, Wang D, Hu X, Zhang Z. Single-cell meta-analyses reveal responses of tumor-reactive CXCL13⁺ T cells to immune-checkpoint blockade. *Nat Cancer* 2022;3:1123–36.
- [312] Zheng C, Fass JN, Shih YP, Gunderson AJ, Sanjuan Silva N, Huang H, et al. Transcriptomic profiles of neoantigen-reactive T cells in human gastrointestinal cancers. *Cancer Cell* 2022;40:410–23.e7.
- [313] Bennett HM, Stephenson W, Rose CM, Darmanis S. Single-cell proteomics enabled by next-generation sequencing or mass spectrometry. *Nat Methods* 2023;20:363–74.
- [314] Gao S, Shi Q, Zhang Y, Liang G, Kang Z, Huang B, et al. Identification of HSC/MPP expansion units in fetal liver by single-cell spatiotemporal transcriptomics. *Cell Res* 2022;32:38–53.
- [315] Stahl PL, Salmen F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 2016;353:78–82.
- [316] Rodrigues SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* 2019;363:1463–7.
- [317] Vickovic S, Eraslan G, Salmen F, Klughammer J, Stenbeck L, Schapiro D, et al. High-definition spatial transcriptomics for *in situ* tissue profiling. *Nat Methods* 2019;16:987–90.
- [318] Liu Y, Yang M, Deng Y, Su G, Enniful A, Guo CC, et al. High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. *Cell* 2020;183:1665–81.e18.
- [319] Merritt CR, Ong GT, Church SE, Barker K, Danaher P, Geiss G, et al. Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nat Biotechnol* 2020;38:586–99.
- [320] Cho CS, Xi J, Si Y, Park SR, Hsu JE, Kim M, et al. Microscopic examination of spatial transcriptome using Seq-Scope. *Cell* 2021;184:3559–72.e22.
- [321] Stickels RR, Murray E, Kumar P, Li J, Marshall JL, Di Bella DJ, et al. Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat Biotechnol* 2021;39:313–9.
- [322] Chen A, Liao S, Cheng M, Ma K, Wu L, Lai Y, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis

- using DNA nanoball-patterned arrays. *Cell* 2022;185:1777–92.e21.
- [323] Deng Y, Bartosovic M, Kukanja P, Zhang D, Liu Y, Su G, et al. Spatial-CUT&Tag: spatially resolved chromatin modification profiling at the cellular level. *Science* 2022;375:681–6.
- [324] Deng Y, Bartosovic M, Ma S, Zhang D, Kukanja P, Xiao Y, et al. Spatial profiling of chromatin accessibility in mouse and human tissues. *Nature* 2022;609:375–83.
- [325] Dhainaut M, Rose SA, Akturk G, Wroblewska A, Nielsen SR, Park ES, et al. Spatial CRISPR genomics identifies regulators of the tumor microenvironment. *Cell* 2022;185:1223–39.e20.
- [326] Fu X, Sun L, Dong R, Chen JY, Silakit R, Condon LF, et al. Polony gels enable amplifiable DNA stamping and spatial transcriptomics of chronic pain. *Cell* 2022;185:4621–33.e17.
- [327] He S, Bhatt R, Brown C, Brown EA, Buhr DL, Chantranuvatana K, et al. High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat Biotechnol* 2022;40:1794–806.
- [328] Vickovic S, Lötstedt B, Klughammer J, Mages S, Segerstolpe Å, Rozenblatt-Rosen O, et al. SM-Omics is an automated platform for high-throughput spatial multi-omics. *Nat Commun* 2022;13:795.
- [329] Liu Y, DiStasio M, Su G, Asashima H, Enniful A, Qin X, et al. High-plex protein and whole transcriptome co-mapping at cellular resolution with spatial CITE-seq. *Nat Biotechnol* 2023;41:1405–9.
- [330] Zhang D, Deng Y, Kukanja P, Agirre E, Bartosovic M, Dong M, et al. Spatial epigenome-transcriptome co-profiling of mammalian tissues. *Nature* 2023;616:113–22.
- [331] Zhuang X. Spatially resolved single-cell genomics and transcriptomics by imaging. *Nat Methods* 2021;18:18–22.
- [332] Rozenblatt-Rosen O, Regev A, Oberdoerffer P, Nawy T, Hupalowska A, Rood JE, et al. The Human Tumor Atlas Network: charting tumor transitions across space and time at single-cell resolution. *Cell* 2020;181:236–49.
- [333] International Cancer Genome Consortium, Hudson TJ, Anderson W, Artez A, Barker AD, Bell C, et al. International network of cancer genome projects. *Nature* 2010;464:993–8.
- [334] Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* 2013;45:1113–20.
- [335] GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580–5.
- [336] Cai Y, Song W, Li J, Jing Y, Liang C, Zhang L, et al. The landscape of aging. *Sci China Life Sci* 2022;65:2354–454.
- [337] Rutledge J, Oh H, Wyss-Coray T. Measuring biological age using omics data. *Nat Rev Genet* 2022;23:715–27.
- [338] Mogilenko DA, Shchukina I, Artyomov MN. Immune ageing at single-cell resolution. *Nat Rev Immunol* 2022;22:484–98.
- [339] Nair GG, Tzanakakis ES, Hebrok M. Emerging routes to the generation of functional beta-cells for diabetes mellitus cell therapy. *Nat Rev Endocrinol* 2020;16:506–18.