SCIENCE CHINA PRESS

SCIENCE CHINA Life Sciences
life.scichina.com link.springer.com
https://doi.org/10.1007/s11427-024-2770-x

REVIEW

Advances and applications in single-cell and spatial genomics

Jingjing Wang[1†], Fang Ye[1†], Haoxi Chai[5†], Yujia Jiang[9,10†], Teng Wang[7,12†], Xia Ran[1,4†], Qimin Xia[7†], Ziye Xu[6†], Yuting Fu[1,2†], Guodong Zhang[1,2†], Hanyu Wu[1,2†], Guoji Guo[1,2,3,4#*], Hongshan Guo[1,4*], Yijun Ruan[5*], Yongcheng Wang[6*], Dong Xing[7,8*], Xun Xu[9,10,11*] & Zemin Zhang[7*]

[1]Bone Marrow Transplantation Center of the First Affiliated Hospital & Liangzhu Laboratory, Zhejiang University School of Medicine, Hangzhou 310058, China
[2]Center for Stem Cell and Regenerative Medicine, Zhejiang University School of Medicine, Hangzhou 310058, China
[3]Zhejiang Provincial Key Lab for Tissue Engineering and Regenerative Medicine, Dr. Li Dak Sum & Yip Yio Chin Center for Stem Cell and Regenerative Medicine, Hangzhou 310058, China
[4]Institute of Hematology, Zhejiang University, Hangzhou 310000, China
[5]Life Sciences Institute and The Second Affiliated Hospital, Zhejiang University, Hangzhou 310058, China
[6]Department of Laboratory Medicine of The First Affiliated Hospital & Liangzhu Laboratory, Zhejiang University School of Medicine, Hangzhou 310058, China
[7]Biomedical Pioneering Innovation Center (BIOPIC) and School of Life Sciences, Peking University, Beijing 100871, China
[8]Beijing Advanced Innovation Center for Genomics (ICG), Peking University, Beijing 100871, China
[9]BGI Research, Shenzhen 518083, China
[10]BGI Research, Hangzhou 310030, China
[11]Guangdong Provincial Key Laboratory of Genome Read and Write, BGI Research, Shenzhen 518083, China
[12]Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

†Contributed equally to this work
#Lead contact author
*Co-corresponding authors listed in an alphabetical order of surnames (Guoji Guo, email: ggj@zju.edu.cn; Hongshan Guo, email: guohs@zju.edu.cn; Yijun Ruan, email: yjruan@zju.edu.cn; Yongcheng Wang, email: yongcheng@zju.edu.cn; Dong Xing, email: dxing@pku.edu.cn; Xun Xu, email: xuxun@genomics.cn; Zemin Zhang, email: zemin@pku.edu.cn)

Received 20 August 2024; Accepted 10 October 2024; Published online 20 December 2024

The applications of single-cell and spatial technologies in recent times have revolutionized the present understanding of cellular states and the cellular heterogeneity inherent in complex biological systems. These advancements offer unprecedented resolution in the examination of the functional genomics of individual cells and their spatial context within tissues. In this review, we have comprehensively discussed the historical development and recent progress in the field of single-cell and spatial genomics. We have reviewed the breakthroughs in single-cell multi-omics technologies, spatial genomics methods, and the computational strategies employed toward the analyses of single-cell atlas data. Furthermore, we have highlighted the advances made in constructing cellular atlases and their clinical applications, particularly in the context of disease. Finally, we have discussed the emerging trends, challenges, and opportunities in this rapidly evolving field.

single-cell | technologies | cell atlas | multi-omics | spatial | clinical application

Introduction

The human body is a highly organized system of approximately 37 trillion cells encompassing hundreds of cell types (Han et al., 2020; Jovic et al., 2022; Wen and Tang, 2022). Despite originating from a single zygote, cells continue to accumulate genetic and epigenetic mutations as they divide and grow (Takeshima and Ushijima, 2019; Zong et al., 2012). In fact, cells within the same tissue, organ, or cell type may contribute differently to physiological or pathological processes. Understanding this heterogeneity at the single-cell level is crucial for uncovering insights into developmental biology, disease mechanisms, and therapeutic strategies. Analyzing the human body at the level of individual cells can facilitate data integration and modeling across diverse tissues, thereby advancing our comprehension of human systems.

Chapter 1. Overview of single-cell and spatial genomics

Conventionally, the genetic makeup of cells has been analyzed at the molecular or cellular level with limited markers. Conventional bulk-cell-based next-generation sequencing (NGS), also known as bulk sequencing, analyzes cells from bulk tissues as a whole and provides averaged signals from heterogeneous cell populations. Through bulk sequencing, our understanding of biology has revolutionized from relevant pathways to key genes, but it is stuck at the molecular level as the differences between individual cells are not inquired. Low-throughput techniques such as single-cell polymerase chain reaction (PCR) and immunofluorescence enable single-cell analyses, albeit only a limited number of markers can be quantified this way (Kaur et al., 2019). Thus, there is an urgent need to develop single-cell sequencing technologies that can systematically profile indivi-

Citation: Wang, J., Ye, F., Chai, H., Jiang, Y., Wang, T., Ran, X., Xia, Q., Xu, Z., Fu, Y., Zhang, G., et al. (2025). Advances and applications in single-cell and spatial genomics. Sci China Life Sci 68, 1226–1282. https://doi.org/10.1007/s11427-024-2770-x

dual cells.

To sequence single cells, it is a prerequisite to isolating single cells from heterogeneous cell populations (Walker and Parkhill, 2008). Major challenges encountered in single-cell isolation are the scalability and sensitivity of the isolation methods while ensuring the quality and purity of the individual cells (Gross et al., 2015). To tackle these challenges, various approaches have been adopted, which are mainly based on the following three principles (Hu et al., 2016a; Kaur et al., 2019): (i) physical features, such as the size, density, electric charges, and deformability; (ii) biological properties, such as the type of surface proteins; (iii) molecular barcoding by using a split-pool ligation-based strategy. Considering the various aspects of cell input/output, automation level, scalability, and cell viability after isolation, four major approaches have been adopted and widely applied across laboratories, including micromanipulation, laser-capture microdissection (LCM), fluorescence-activated cell sorting (FACS), and microfluidics-assisted droplet encapsulation. Of these approaches, microfluidics-assisted droplet encapsulation exhibits better throughput performance and automation (Zhou et al., 2021).

The second prerequisite of single-cell sequencing is to effectively amplify trace amounts of genetic materials from single cells (Walker and Parkhill, 2008). PCR with random or semi-random primers has laid the groundwork for single-cell amplification. Although it is bias-prone, PCR is widely used and later optimized across diverse hybridization-based or isothermal amplification methods (Kaur et al., 2019). The hybridization-based methods encompass methods that have been modified from the conventional PCR, including methods such as linker-adaptor PCR, primer extension preamplification PCR, and degenerate oligonucleotide-primed PCR. To overcome the limitations of hybridization methods (e.g., low coverage, allele dropouts, and amplification bias), the isothermal methods and *in vitro* transcription (IVT) were subsequently developed (Kaur et al., 2019; Walker and Parkhill, 2008).

Developed in parallel with single-cell genomic technologies, spatial genomic methods preserve the position information and map the spatial organization of gene expression and other molecular features within tissues. Exciting progress has been recorded for spatial transcriptomics (STs), which combines spatial information with transcriptomics, thereby allowing gene expression profiling while preserving tissue architecture. In addition, multiplexed *in situ* hybridization detects multiple RNA or DNA targets within a single tissue section and provides high-resolution spatial maps of gene expression.

With technological advances in single-cell barcoding and amplification, single-cell and spatial genomics has emerged, wherein cells can be profiled with both molecular and position indexes in a high-throughput manner. Recently, these technical forces have contributed to the construction of comprehensive cellular atlases involving mapping of the molecular characteristics of all cell types within an organism. Further mapping of cellular changes in various diseases can provide unprecedented insights into disease mechanisms and potential therapeutic targets. However, to fully realize the potential of these atlases, we need to overcome several hurdles. Here, we have discussed some of the key areas in the field of single-cell and spatial genomics identified in the past decade and with potential in the future.

Chapter 2. Single-cell sequencing technologies

So far, single-cell sequencing technologies have evolved from transcriptome to genome, epigenome, and proteome, from singular modality to multi-omics, from single-cell to subcellular, and scaled up from dozens of cells to millions of cells (Figure 1).

Single-cell RNA-seq

scRNA-seq estimates the RNA molecules of individual cells and provides a snapshot of the entire transcription landscape within the assayed cell to reveal the genes that are being actively transcribed into individual cells. Based on this information, the cell identity, state, and type can be defined at a more granular level (Cuomo et al., 2023). In a diploid cell, many expressed genes have more than a dozen copies of the mRNAs, which makes it relatively feasible to quantify them in individual cells (Wen and Tang, 2022). For this reason, perhaps, scRNA-seq stands out as the earliest, most commonly-used, and most cutting-edge (especially in terms of throughput (Tanay and Regev, 2017) technology that has been developed for single-cell sequencing.

The inaugural demonstration of scRNA-seq on a NGS platform dates back to 2009 (Tang et al., 2009). This first model analyzed the whole transcriptomic characterization of manually isolated single cells. Since then, scRNA-seq has scaled up rapidly (Svensson et al., 2020)—from only 8 cells in the first study to more than 11 million cells in a recent publication (Qiu et al., 2024). Over the ensuing nearly 15 years, the landscape of scRNA-seq sequencing technologies has witnessed an unprecedented surge in terms of development and diversification (Table S1). The conventional scRNA-seq workflow encompasses crucial stages, including single-cell separation, library construction, high-throughput sequencing, and subsequent data analysis. The common steps involved in transcriptome library generation include reverse transcription (RT) into first-strand cDNA, second-strand synthesis, and cDNA amplification. Notably, diverse scRNA-seq technologies exhibit distinct advantages and limitations throughout the process. Although certain preliminary technologies have now turned obsolete, some others have undergone continuous refinements, such as the markedly enhancing pace, throughput, sensitivity, coverage, capacity for long reads, and overall consistency. Pioneering technologies such as Smart-seq (Ramsköld et al., 2012), introduced in 2012, have now evolved into more advanced iterations such as the recently updated Smart-seq3xpress (Hagemann-Jensen et al., 2022). In plate-based scRNA-seq techniques such as Smart-seq and CEL-Seq, the experimental cells are physically separated into micro-well plates (such as a 96-well or 384-well plate) in a one-cell one-well fashion, with only a few dozen or hundred cells analyzed at a time. However, in the droplet-based strategy, such as Drop-seq and inDrop, single cells are encapsulated into microdroplets containing beads that are prebarcoded by unique oligonucleotides. After cell lysis, mRNA molecules are hybridized and reverse transcribed with the barcodes and then eventually sequenced using labels that correspond to their cell origin (Cuomo et al., 2023). Notably, only the 5′ or 3′ end of transcripts is sequenced and thousands of cells are analyzed simultaneously. The pivotal moment in scRNA-seq's commercialization transpired with the launch of 10x Genomics' Chromium in 2015, which provides a benchtop solution accessible to scientists and catalyzing advancements in other commercial scRNA-seq

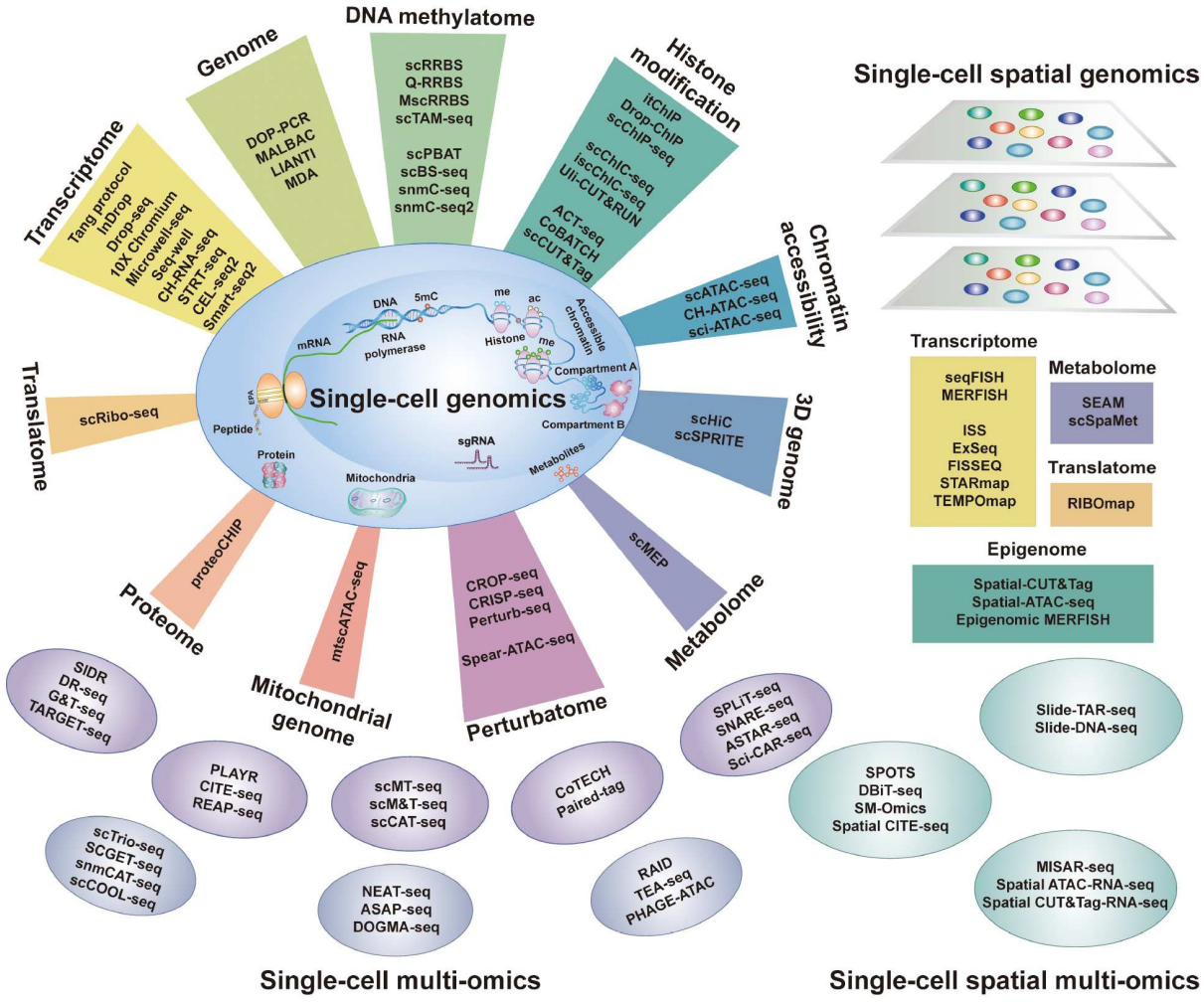

Figure 1. Overview of single-cell technologies in different omics.

technologies (Zheng et al., 2017b). Together, these events have helped to shape the dynamics and rapidly evolving landscape of single-cell transcriptomics.

Single-cell separation

As the primary distinction from conventional sequencing methods, the separation of single cells from samples is a crucial step in gathering transcriptome information from individual cells. Conventional techniques for single-cell separation have been thoroughly scrutinized in other studies (Hwang et al., 2018; Pensold and Zimmer-Bensch, 2020; Tan et al., 2010), including limiting dilution, micromanipulation, LCM, and FACS. In low-throughput scRNA-seq technologies (e.g., Smart-seq2 (Picelli et al., 2013), single-cell universal poly(A)-independent RNA sequencing (SUPeR-seq) (Fan et al., 2015b), molecular crowding single-cell RNA barcoding and sequencing (mcSCRB-seq) (Bagnoli et al., 2018), and massively parallel single-cell RNA-sequencing (MARS-Seq) (Jaitin et al., 2014)), the single cells obtained via these techniques can be individually segregated into tubes, which subsequently undergo distinct reactions. Certain conventional low-throughput methods for single-cell separation have now evolved to facilitate detailed examination of intricate solid tissues, rare samples, and even subcellular fractions. For instance, LCM utilizes a computer-aided laser system for the precise dissection of single cells from tissues with spatial accuracy (Chen et al., 2017b; Nichterwitz et al., 2016). A microfluidic device based on micropillar arrays offers highly efficient isolation, detection, and collection of low-abundance cells (e.g., circulating tumor cells (CTCs)) (Chen and Wang, 2023). Minimally invasive nanotweezers allow spatially controlled extraction of samples from living cells with single-molecule/organelle precision (Nadappuram et al., 2019).

An additional avenue within the design of single-cell separation methods seeks to optimize both throughput and efficiency. The release of the first single-cell automated prep system, Fluidigm C1, which can be implemented with CEL-Seq2 (Hashimshony et al., 2016) and Smart-seq/C1 (Shalek et al., 2014), signified a significant advancement. Microwell-based scRNA-seq methods (e.g., CytoSeq (Fan et al., 2015a), Seq-Well (Gierahn et al., 2017), and Microwell-Seq (Han et al., 2018b)), wherein cells are deposited into wells by gravity at random, offer the advantages of low cost and high throughput. The introduction of droplet microfluidic technology has revolutionized cell throughput (Wen et al., 2016). Basically, it enables the encapsulation of single cells and reagents in independent aqueous microdroplets (typically tens to hundreds of μm in diameter) in a high-throughput manner (e.g., Drop-seq (Macosko et al., 2015), inDrop (Klein et al., 2015) and 10x Genomics'

Chromium). Nevertheless, both microwell- and droplet-based scRNA-seq methods rely on stochastic pairing of cells and barcoded beads, which is constrained by the mathematical principle of Poisson statistics, resulting in low utilization of both the cells and beads. To resolve these limitations, innovative approaches such as dTNT-seq (Bai et al., 2020), which integrates dielectrophoresis (DEP)-trapping-nano well-transfer, and Well-paired-seq (Yin et al., 2022), which consists of thousands of size exclusion and quasi-static hydrodynamic dual wells, have been developed to enhance microreactor and cell/bead capture efficiency. Single-cell combinatorial indexing (sci), incorporating *in situ* nucleic acids indexing and a "split-pool" strategy for unique labeling, serves as a low-cost methodological framework for high-throughput scRNA-seq (e.g., sci-RNA-seq (Cao et al., 2017) and SPLiT-seq (Rosenberg et al., 2018)). Datlinger et al. (2021) introduced a "single-cell combinatorial fluidic indexing" (scifi) strategy and developed scifi-RNA-seq—a method that combines one-step combinatorial preindexing of entire transcriptomes with subsequent single-cell RNA-seq using microfluidics—which eliminated the need for absolute single-cell separation and overcame the limitations of Poisson distribution in encapsulation. These limitations of single-cell separation methods originate from the essential use of microwells or microfluidic devices necessary for sample processing. The microfluidics-free particle-templated instant partition sequencing (PIP-seq) (Clark et al., 2023) allows single-cell encapsulation and barcoding of cDNA in uniform droplet emulsions by using only a vortex. Recently, an instrument-free scRNA-seq method RevGel-seq (Komatsu et al., 2023), which utilizes a reversible hydrogel, has presented a novel cell-bead pairing strategy that does not require nanowells or droplets, thereby introducing a new direction in the development of single-cell separation in scRNA-seq technologies.

Droplet microfluidic technology also offers the benefits of adaptability and transportability in high-throughput scRNA-seq. Various cell characterization technologies can be incorporated with microfluidic chips to further select specific cells during single-cell separation, which can be achieved by regulating the micron-diameter channels' liquid flow. Computational Sorting and Mapping of Single Cells (COSMOS) is a platform based on artificial intelligence (AI) and microfluidics that can characterize and sort single cells based on high-dimensional embedding vectors of morphology without needing biomarker labels and stains/dyes (Salek et al., 2023). In addition, an automated Raman-based sorting method involves the use of optical tweezers and microfluidics to sort individual cells of interest (Lee et al., 2021). Digital microfluidics (DMF) is also an emerging technique that employs dielectric wetting to precisely manipulate discrete fluids, allowing efficient and non-destructive isolation of single cells (Zeng et al., 2024).

Cell barcoding

A significant advancement that substantially elevates the enumeration of profiled single cells involves the introduction of single-cell tagged reverse transcription (STRT-seq) (Islam et al., 2011), which uses cell-specific barcodes allowed for multiplexed single-cell RNA-seq on the Illumina platform. Barring these low-throughput scRNA-seq technologies (e.g., Smart-seq2, Smart-seq/C1, SUPeR-seq, multiple annealing, and dC-tailing-based quantitative single-cell RNA-seq (MATQ-seq) (Sheng et al., 2017)), virtually all scRNA-seq technologies utilize cell-specific barcodes, each employing distinct application methodologies. The cellular origin of RNA is labeled with specific-barcoded RT primers, e.g., MARS-Seq, Drop-seq, CytoSeq, and vast transcriptome analysis of single cells by dA-tailing (VASA-seq) (Salmen et al., 2022). To further facilitate scalable profiling of single cells, cost-effective and high-throughput scRNA-seq methods (e.g., sci-RNA-seq, SPLiT-seq, and sci-RNA-seq3 (Cao et al., 2019)) have been developed, wherein cells are combinatorial indexed by RT, cDNA ligation, and PCR. The majority of droplet- and microwell-based scRNA-seq technologies incorporate microparticles or hydrogel microspheres (beads), each with distinct barcoded primers. Each barcoded bead is coencapsulated with an individual cell and reaction reagents utilizing microfluidic chips. However, the diversity of such barcoded beads presents a considerable challenge for these bead-based high-throughput scRNA-seq methods. Drop-seq introduced a 12-round "split-and-pool" strategy to directly synthesize oligonucleotide primers on beads. inDrop and Microwell-Seq synthesized the library of barcoded hydrogel microspheres using two or three split-pool rounds of hybridization and primer extension. Generally, among all cell barcoding methods, the microwell- and nanowell-based methods demonstrate comparable sensitivity, throughput, and lower cost when compared with the droplet platform method. The stability of commercialized equipment in both droplet and microwell platforms has provided diverse choices for users based on the sample size, cell size, and budget. The droplet platform may not be appropriate for large cells that can lead plugs. On the other hand, pool-split indexing methods are convenient and require no special equipment, making them suitable for almost all laboratories (does not need an automatic pipetting system). Nonetheless, there is a drawback to the need for manual pipetting proficiency to ensure reproducible outcomes.

cDNA library generation strategies

Together with the enhancements in throughput, the cDNA library generation of scRNA-seq technologies has incessantly evolved to amplify the sensitivity and accuracy of RNA in single cells. This aspect remains the most salient distinction among various scRNA-seq methods. In the following section, we have reviewed these diverse strategies.

Before the library generation reactions, some of these scRNA-seq technologies (e.g., droplet-based Drop-seq and microwell-based CytoSeq) lyse single fresh cells in separated chambers and capture the released RNA molecules through hybridization using primers. Some scRNA-seq technologies (e.g., plate-based SPLiT-seq, microwell-based Microwell-seq 2.0 (Chen et al., 2021a), Microwell-seq3 (Ye et al., 2024b), and droplet-based snRandom-seq (Xu et al., 2023f)) are compatible with fixed cells and nuclei, allowing the capturing of RNA molecules and the generation of cDNA through an *in situ* RT reaction.

scRNA-seq methodologies can be categorized into three major types—oligo-dT, random/not-so-random primer (NSRs), and probe-based—depending on the primers employed for the initial cDNA synthesis. Ever since the first scRNA-seq of Tang et al. (2009) a significant proportion of classical scRNA-seq technologies use oligo-dT primers for the first-strand cDNA synthesis. Oligo-dT priming is an efficient approach to capturing mRNAs with minimal uninformative reads (e.g., those derived from rRNAs). The state-of-the-art oligo-dT-based scRNA-seq methods (e.g., plate-based Smart-seq2, droplet-based 10x Genomics' Chromium, and microwell-based CytoSeq) are adequately sensitive to quantify and determine cell states with high

accuracy. Nevertheless, this strategy solely captures and amplifies polyadenylated RNAs to consequently yield an incomplete depiction of the transcriptome. Moreover, the oligo-dT-primed cDNA synthesis requires high-quality intact RNAs, thereby precluding the application of oligo-dT-based scRNA-seq technologies with samples that are deteriorated or degraded (e.g., formalin-fixed paraffin-embedded (FFPE) samples). Random priming is another approach for the first-strand cDNA synthesis, which can cover full-length of any size of transcripts. SUPeR-seq introduced random primers to enable the simultaneous detection of both poly(A)+ and poly(A)− RNA species (e.g., circular RNAs). MATQ-seq is more sensitive with the use of RT primers that consist of the MALBAC primer, three consecutive G or T, and 20 consecutive T (dT20). RamDA-seq combined a novel RT technology: RT with random displacement amplification (RT-RamDA) and NSRs (Hayashi et al., 2018). The high-throughput MATQ-drop (Niu et al., 2023), snRandom-seq (Xu et al., 2023f), and snHH-seq (Chen et al., 2024a) also used random primers for total RNA capture and average gene body coverage. 10x Genomics' Chromium Single Cell Gene Expression Flex applies probe-based chemistry for RNA capture in relatively low-quality fixed samples. In the FFPE samples with worse RNA quality, the random primer-based snRandom-seq displayed much higher sensitivity when compared with other probe-based (snPATHO-Seq (Vallejo et al., 2022)) or poly(A)-based (snFFPE-seq (Chung et al., 2022)) scRNA-seq methods of FFPE samples.

Once the RNA molecule undergoes RT into cDNA, the comprehensive amplification of the cDNA library is mandated before its in-depth sequencing. Presently, scRNA-seq primarily employs three strategies for cDNA library amplification, namely PCR-based amplification, linear isothermal amplification via T7-based IVT, and amplification utilizing Phi29 DNA polymerase. As the Phi29 DNA polymerase-based amplification has not been updated recently, we have reviewed scRNA technologies using PCR- and IVT-based amplification.

Based on the synthesis pattern of second-strand cDNA with a second anchor sequence for PCR amplification, PCR-based amplification can be categorized into homopolymer tailing, template-switching, and random priming. In the first scRNA-seq research by Tang et al. (2009), ~30-nt polyA tails are attached to the first-strand cDNA at the 3′-end through terminal deoxynucleotidyl transferase. Subsequently, a second poly(dT) primer, anchored differently, is employed to procure double-stranded DNA, which is subsequently amplified via PCR. Subsequent Quartz-Seq (Sasagawa et al., 2013), MATQ-seq, snRandom-seq, and snHH-seq also adopted a similar homopolymer tailing strategy. Considering that the polyA or polyC tail can be added to any excrescent RT primers, the cDNA library will contain byproducts from the surviving primers. In Quartz-Seq, all the steps of the whole-transcript amplification are executed in a single PCR tube. The preponderance of superfluous RT primer is digested by exonuclease I, while the amplification of byproducts is inhibited through the application of suppression PCR primer. Conversely, in MATQ-seq, the excrescent RT primers are digested by adding T4 DNA polymerase into the original reaction tube to avoid the production of by-products. The droplet-based MATQ-drop, snRandom-seq, and snHH-seq are performed on fixed single cells or nuclei, which allow indirect washing with buffer after the RT reactions. STRT-seq was the first to introduce a template-switching strategy for second-strand cDNA synthesis. In STRT-seq, mRNA is reverse-transcribed into cDNA by a tailed oligo-dT primer with a barcode and an upstream primer-binding sequence, and the barcode is introduced into the 5′-end of the transcripts by Moloney Murine Leukemia Virus reverse transcriptase (MMLV RT) (Zhu et al., 2001). Smart-seq epitomizes a robust and precise methodology that enhances read coverage across transcripts, with marked sensitivity and accuracy in quantification. It also leans heavily on the terminal transferase and template-switching activities intrinsic to MMLV RT. It has been hypothesized that template switching transpires sequentially, wherein once the RT reaction reaches the 5′-end of an RNA molecule, there is an untempled appendage of three protruding nucleotides (+CCC) to the 3′-end of the nascent cDNA, which then hybridizes with the complementary rGrGrG 3′-end of the template-switching oligo (TSO), thereby facilitating RT to swap its template and resume transcription. The SMART strategy, which significantly enriches transcripts with intact 5′-ends and eliminates the second-strand synthesis, is the most feasible strategy to integrate with plate-, microwell-, or droplet-based full-length scRNA-seq methods. Thus, the switching mechanism at the 5′-end of the RNA template (SMART) has become the most common strategy used by Smart-seq2, CytoSeq, Drop-seq, and 10x Genomics' Chromium. An IVT strategy based on T7 RNA polymerase was developed for the synthesis of the second strand without any need for an anchor sequence. The IVT strategy has been adopted by several scRNA-seq technologies, such as CEL-Seq (Hashimshony et al., 2012), CEL-Seq2, MARS-seq, and inDrop. CEL-seq is the first method to employ IVT in scRNA-seq, which overcomes the limitation of the small initial amounts of RNA. In CEL-seq, the amplified RNA via IVT strategy is fragmented to a size distribution appropriate for sequencing and then converted into a library compatible with Illumina sequencing through ligation with the 3′ adaptor. However, the ligation step is not efficient and introduces primer dimers that interfere with sequencing. CEL-Seq2, a refined iteration of CEL-Seq executed on Fluidigm's C1 system, dispenses with the ligation phase by incorporating the Illumina adaptor directly during the second RT step, serving as a 5′-tail affixed to a random hexamer. MARS-Seq provided an optional addition of plate barcode by ligating the 3′ adaptor on the amplified and fragmented RNA. In inDrop, the amplification of RNA according to CEL-Seq was conducted in nanoliter droplets. The recently reported high-throughput VASA-seq also embraced this IVT tactic. However, it involves the engagement of a picoinjection apparatus for two successive operations in the droplets of VASA-seq, which poses a challenge to its broad application due to the increased complexity. Although Quartz-Seq uses a homopolymer tailing strategy for the cDNA library amplification, IVT has been used in the Quartz-Chip method to synthesize labeled cRNA from the amplified cDNA for the microarray analysis.

Transcript coverage in cDNA library generation

Predicated on the transcript coverage, these scRNA-seq technologies present a clear demarcation into 3′-end, 5′-end, and full-length modalities. Traditional NGS is restricted to abbreviated transcript fragments, introducing a noticeable 3′-end bias in read distribution for high-throughput scRNA methodologies that use oligo-dT priming, such as Drop-seq, inDrop, and 10x Genomics' Chromium (Zhang et al., 2019d). Conversely, 5′-tag counting scRNA-seq methods like STRT-seq and STRT-seq-2i (Hochgerner et al., 2017) exhibited a bias toward the 5′-end of transcripts. Although both 5′-end and 3′-end scRNA-seq demonstrated

comparable cell typing capabilities, 5′-end approaches offer specific advantages in certain applications. For instance, C1 Cap analysis gene expression (C1 CAGE) (Kouno et al., 2019) employed an original sample multiplexing strategy in the C1 microfluidic system for the detection of transcription start sites and enhancer activity at single-cell resolution. In addition, scBCR-seq (Goldstein et al., 2019), which combines 5′-end scRNA-seq with single-cell V(D)J sequencing, profiled B cell receptor repertoires. The ultra-high throughput 5′-end scRNA-Seq method FIve PRime End Single-cell Combinatorial Indexing RNA sequencing (FIPRESCI) (Li et al., 2023d) combines *in situ* preindexing using Tn5 transposomes with droplet template switch oligo (TSO) barcoding, thereby allowing sample multiplexing and multimodal profiling of the entire transcriptome and immune receptor repertoire. However, the 3′- or 5′-end bias limits the mutation and splicing analysis (Chen et al., 2024a). Smart-seq2 has been the gold standard for full-length scRNA-seq owing to its sensitivity and robustness. Each specimen procured via the low-throughput oligo-dT-based Smart-seq2 undergoes individual processing. The cDNA libraries of Smart-seq2, once fragmented, are primed for thorough sequencing and deeper analysis. This step unlocks opportunities for characterizing splice isoforms, allelic variants, and single-nucleotide polymorphisms at a single-cell resolution. High-throughput VASA-seq employs oligo-dT primers as well but staggers and polyadenylates every RNA molecule in a lysed single cell, thereby facilitating the apprehension of total RNAs via oligo-dT primers. Even SUPeR-seq, which uses random (AnchorX-T15N6) primers, shows mild 3′-end bias. In fact, some scRNA-seq technologies like MATQ-seq, snRandom-seq, and snHH-seq also incorporate random primers, thereby integrating a multiple annealing tactic during the RT process for the capturing of full-length transcripts.

Sample multiplexing in cDNA library generation

Pre-indexing based on separated tagmentation or RT provides an effective way for sample multiplexing. Multiplexing for distinct samples in one experiment can be achieved through antibody labeling, genetic barcoding, and click chemistry reactions. Cell Hashing (Stoeckius et al., 2018) used oligo-tagged antibodies against ubiquitously expressed surface proteins to mark cells obtained from different samples. The mimic poly-A tails in the oligos were captured and analyzed in the sequencing data to distinguish these samples. This method was then further developed to simultaneously analyze target proteins and RNA at a single-cell level. In contrast, CellTag Indexing (Guo et al., 2019) used heritable predefined genetic barcodes to label different cell types sourced from different samples, which allows long-term cell tracking both *in vitro* and *in vivo*. The genetic barcoding strategies were adapted to gene expression perturbation and lineage tracing. For sample multiplexing based on click chemistry reaction, the "ClickTags" method (Gehring et al., 2020) involved rapid cross-linking reaction along with the identification of methyltetrazine-modified DNA oligonucleotides by using cross-linker NHS-trans-cyclooctene (NHS-TCO). Another method, MULTI-seq, used lipid- and cholesterol-modified oligonucleotides as an anchor with plasma membranes for cell labeling (McGinnis et al., 2019). In general, click chemistry reactions can provide flexible and effective labeling approaches for cells, nuclei, and molecules. Thus, the modifications of DNA and proteins can thereby be distinguished and sequenced by other multi-omics methods.

RNA dynamics

scRNA-seq of gene expression dynamics can provide unique insights into transcription perturbation during development, drug treatment, and disease progression. Time-resolved dimension of RNA synthesis in the system can achieve temporal resolution, which is important in an organism. In 2019, single-cell, thiol-(SH)-linked alkylation of RNA for metabolic-labeling sequencing (scSLAM-seq) (Erhard et al., 2019) and new transcriptome alkylation-dependent single-cell RNA sequencing (NASC-seq) (Hendriks et al., 2019) utilizes nucleoside analog 4-thiouridine (4sU) and iodoacetamide-based chemical conversion reactions to label new RNA during transcription. The abundant T-C conversions in new RNA distinguish them from unlabeled "old" RNA. In 2020, sci-fate (Cao et al., 2020b) and scNT-seq (Qiu et al., 2020) reported similar metabolic labeling methods by using split-pool indexing and droplet microfluidics platform to achieve high-throughput sequencing of RNA dynamics in single cells, respectively. In addition to these metabolic labeling methods, Liver-seq allows direct extraction of RNA in single cells through fluidic force microscopy and preserves cell viability (Chen et al., 2022b). However, the equipment and throughput restrict its further application. Recently, single-cell nascent RNA sequencing (scGRO-seq) employed click chemistry rather than metabolic labeling to directly evaluate new RNA and to transcribe RNA polymerases in single cells (Mahat et al., 2024). The dynamic landscape of global transcription unveiled co-transcription between pluripotency genes and their enhancers.

Single-microbe RNA-seq

Microbial traits exhibit both heterogeneity within populations and precise transcriptional control, which emphasizes the need for conducting single-microbe transcriptional analyses. However, the existing scRNA-seq technologies face substantial incompatibility with single-microbe RNA-seq due to several formidable technical challenges. For instance, a prime constraint arises from the lack of 3′-end poly(A) tails in bacterial mRNAs, which are significant capture regions for poly(T) primers. Moreover, the RNA content in an ordinary bacterium is nearly two orders of magnitude less compared with that in a standard mammalian cell (Westermann and Vogel, 2021), with ribosomal RNA (rRNA) constituting over 80% of total bacterial RNAs (Giannoukos et al., 2012). The rigid cell wall of bacteria presents additional challenges toward cell lysis and the subsequent extraction of RNA in droplets. Nevertheless, the emergence of recent technical enhancements has expedited the realization of bacterial single-cell transcriptomics. The primary procedure for single-microbe RNA-seq significantly parallels the protocols devised previously for eukaryotic cells. For single microbe separation, in the early single-bacterium RNA-seq method, BaSiC RNA-seq (Wang et al., 2015), individual bacterial cells underwent isolation, with subsequent purification and amplification of total RNA from each cell separately. Some relatively low-throughput methods for single-bacterium RNA-seq (such as Imdahl et al.'s method (Imdahl et al., 2020) and Homberger's method (Homberger et al., 2023)) have been devised by employing FACS. Two high-throughput bacterial scRNA-seq methods (prokaryotic expression profiling by tagging RNA *in situ* and sequencing (PETRI-seq) (Blattman et al., 2020) and microSPLiT (Kuchina et al., 2021)) adopted a split-pool barcoding strategy, which was initially

established for eukaryotes to allow researchers to simultaneously analyze thousands of bacteria. Recently, the commercially available 10x Genomics droplet platform—a popular choice for high-throughput single-cell RNA-seq—was embraced by probe-based bacterial sequencing (ProBac-seq) (McNulty et al., 2023), M3-Seq (Wang et al., 2023a), and BacDrop (Ma et al., 2023b). M3-Seq combined plate-based *in situ* indexing by RT with droplet-based indexing by ligation. Moreover, smRandom-seq (Xu et al., 2023e) employed an enhanced barcoding platform, which was originally devised for single-cell barcoding. This platform is characterized by smaller barcoded beads and droplets, as complemented by a refined primer release and barcode synthesis method. For cDNA library generation, Imdahl et al. (2020) and Homberger et al. (2023) developed methods based on the poly (A)-independent MATQ-seq strategy, which was previously established for eukaryotes. In Homberger's refined bacterial single-cell RNA-seq method through automated MATQ-seq, the enhancements in gene coverage and gene detection thresholds improved, and small RNAs were successfully identified at the single-cell level. PETRI-seq tagged RNA molecules *in situ* within a fixed single microbe using random primers. microSPLiT poly-adenylated the mRNA in-cell with *E. coli* poly(A) polymerase I (PAP). The subsequent protocols for these two technologies were both built on SPLiT-seq. ProBac-seq captured individual transcripts with pre-designed DNA probes complementary to protein-coding sequences, which potentially limited unbiased discovery. Recently, three high-throughput single-microbe RNA-seq methods (BacDrop, smRandom-seq, and M3-seq) were developed using random primers for *in situ* cDNA generation. The previous single-microbe RNA-seq technologies, such as PETRI-seq, exhibited an abundance of signals from rRNA, which could compromise mRNA detection and result in high sequencing costs. For the mRNA enrichment process, microSPLiT used mRNA-preferred PAP to polyadenylate RNA. BacDrop performed rRNA depletion *in situ* through probe-hybridization and followed RNase H digestion. These *in situ* mRNA enrichment methods may pose a risk of losing unamplified non-rRNA transcripts. M3-seq also adopted this probe-hybridization-based strategy but performed rRNA depletion after library amplification. Another rRNA depletion strategy (CRISPR-based DASH (Gu et al., 2016)) was employed by Homberger's method and smRandom-seq at the sequencing library level. Although these singular microbe RNA-seq methodologies have proven effective with both gram-negative and gram-positive laboratory bacterial strains, the intricate particulars of these protocols presumptively necessitate further refinement for single-microbe RNA-seq experiments involving microorganisms derived from complex natural communities.

Single-cell genome sequencing (or scDNA-seq)

Along with the development of scRNA-seq, several conventional bulk DNA sequencing methods have been reconfigured to enable the analysis of individual cells. Different from scRNA-seq, scDNA-seq typically amplifies the 2 copies of the genome in mammalian cells, which requires sensitive amplification techniques with single molecular detection power and limited biases. scDNA-seq measures the DNA molecules within each cell of a given sample to allow the detection of genetic variants of a certain frequency and the DNA copy number changes in the assayed cells. In scDNA-seq, DNA extracted from isolated cells is most frequently amplified using isothermal PCR with random or semi-random primers, such as multiple displacement amplification (MDA) and multiple annealing and looping-based amplification cycle (MALBAC). MDA provides a better but non-uniform genome coverage, making it suitable for the detection of single-nucleotide variations (SNVs). In comparison, MALBAC provides a reduced but more even coverage and is thus more suitable for detecting copy number variations (CNVs). To further minimize priming and amplification biases, linear amplification via transposon insertion (LIANTI) was later developed, which demonstrated the highest genome amplification uniformity (Chen et al., 2017a). In addition to the conventional scDNA-seq method, some novel technologies have been developed based on scDNA-seq. For example, Leighton et al. (2023) developed a multipatient-targeted scDNA-seq that enables the construction of mutational lineage in breast cancer.

Single-cell epigenome sequencing

The epigenome comprises a set of heritable modifications to DNA, RNA, or histones within a cell, which can alter gene expression. Considering that epigenetic information is generally dispersed throughout the genome, single-cell epigenome sequencing poses particular challenges (Wen and Tang, 2022).

DNA methylation

In mammals, chemical modifications to the DNA mainly involve the addition or removal of methyl groups from the C-5 position of cytosine (5mC) or the N^6 position of adenine (m^6A). As the dominant type of DNA modification, 5mC occurs mainly in the form of 5′-3′ CpG dinucleotides (Liu et al., 2023b). Remarkably, DNA methylation has been demonstrated to play a critical role in the development of human diseases; analysis of DNA methylation at the single-cell level is thus essential to reveal cellular heterogeneity in DNA methylation. The optimization of bulk DNA methylation sequencing technologies has accelerated the development of a variety of single-cell DNA methylation sequencing (scDNAm-seq) technologies, mainly based on restriction digestion or post-bisulfite adapter tagging (PBAT) (Liu et al., 2023b). The first scDNAm-seq technology was developed in 2013 based on reduced representation bisulfite sequencing (scRRBS), which utilizes the strategy of one-tube enzymatic reaction to reduce unnecessary DNA loss (Guo et al., 2013). Subsequently, on the basis of restriction digestion, methods such as Q-RRBS (Yang et al., 2015), MscRRBS (Charlton et al., 2018), and scTAM-seq (Bianchi et al., 2022) were developed sequentially; these methods have increased the throughput by up to 10,000 cells. Meanwhile, scBS-seq is the first PBAT-based scDNAm-seq technology developed in 2014 that utilizes the strategy of two rounds of random priming and analyses of 96 cells (Smallwood et al., 2014). Subsequently, methods such as snmC-seq (Luo et al., 2017) and snmC-seq2 (Luo et al., 2018) were developed with improved multiplexing strategies and increased throughput to 384 cells. Notably, a challenge in the development of scDNAm-seq technologies is that most methods are based on bisulfite treatment, which causes significant DNA degradation and limited library complexity (Liu et al., 2023b). Therefore, new scDNAm-seq methods are expected to be developed through enzyme-based bisulfite-free approaches, such as Cabernet, a recently published method to measure 5mC and its oxidative form (5hmC) at a single-base resolution with high

genomic coverage (Cao et al., 2023).

Chromatin accessibility

Chromatin accessibility is a window into gene regulation, which denotes the activity of regulatory regions and their accessibility to transcription factors (TFs) (Tsompana and Buck, 2014). The profiling of chromatin accessibility at the single-cell level is, therefore, of vital importance to determine cellular heterogeneity in gene regulatory programs. Currently, chromatin accessibility is mainly assessed based on the accessibility of an enzyme such as Tn5 transposase, DNase I, MNase, or GpC methyltransferase (Wen and Tang, 2022). Probably due to the complexity of the protocol and the intrinsic bias of the enzyme, single-cell DNase-seq (Gao et al., 2021) has not been broadly adopted despite being developed. In contrast, the Tn5-based ATAC-seq methods are readily adaptable for single-cell assay, and various methodologies based on microwell (Chen et al., 2018; Xu et al., 2021b), nanowell plate (Mezger et al., 2018), microfluidic (Buenrostro et al., 2015), or droplet (Lareau et al., 2019) have been developed for scATAC-seq. The first reported scATAC-seq method by Buenrostro et al. paved the way for subsequent scATAC-seq approaches based on droplet, microwell, and pool-split indexing. A plausible reason for the development of scATAC-seq is that it slices the genome and adds adaptor tags to the accessible regions simultaneously in a cell, which makes it exemplary for high-throughput analysis (Wen and Tang, 2022). Different from ATAC-seq, the GpC methyltransferase-based method utilizes exogenous GpC methyltransferase to artificially methylate cytosines within GpC dinucleotides. Considering that GpC dinucleotide occurs at approximately every 25 bp in the mammalian genome, this method provides a higher resolution and enables the analysis of chromatin states and DNA methylation simultaneously within the same single cell (Wen and Tang, 2022).

Histone modifications

Histone modifications are basically posttranslational modifications (e.g., acetylation, phosphorylation, and methylation) that involve adding or removing from histone tails so as to modulate a specific gene expression (Bannister and Kouzarides, 2011). There are multiple histone activation marks (such as H3K4me1, H3K4me3, H3K36me3, H3K79me2, H3K9Ac, H3K27Ac, and H4K16Ac) and repressive marks (such as H3K27me2/3 and H3K9me1/2/3) that display differences in the genomic distribution patterns. To profile histone modifications in individual cells, several methods have been developed, which mainly fall into two categories. The first group of methods is based on ChIP-seq, such as Drop-ChIP (Grosselin et al., 2019), scChIP-seq (Rotem et al., 2015), and itChIP-seq (Ai et al., 2019). The main obstacle in ChIP-seq-based single-cell sequencing is the low specificity and sensitivity of antibody capture and the need for adding cell-specific barcodes before aggregating the cells for immunoprecipitation (Wen and Tang, 2022). For this reason, itChIP-seq adds cell barcodes via Tn5 transposase tagmentation, while scChIP-seq and Drop-ChIP add cell barcodes via MNase digestion and ligation (Wen and Tang, 2022). The second set of methods is based on chromatin immunocleavage (e.g., CUT&RUN and CUT&Tag), such as scChIC-seq (Ku et al., 2019), uli-CUT&RUN (Hainer et al., 2019), iscChIC-seq (Ku et al., 2021), ACT-seq (Carter et al., 2019), combinatorial barcoding, targeted chromatin release (CoBATCH) (Wang et al., 2019a), and scCUT&Tag (Bartosovic et al., 2021; Wu et al., 2021), wherein the enzyme used for chromatin fragmentation is first fused to protein A/G (pA/G) and subsequently conjugated to primary antibody. When compared with ChIP-seq-based methods, these methods omit chromatin fragmentation before antibody incubation, often adapting a combinatory indexing strategy to increase the throughput and automation level (Preissl et al., 2023). In these methods, no multi-step library preparation is required, which minimizes cell loss significantly. Nevertheless, for minimizing pA-Tn5 bias under high-salt conditions, a significant challenge is encountered when profiling TF binding (Preissl et al., 2023).

3D genome organization

Chromatin architecture refers to the 3D organization of DNA and DNA-associated proteins, including topologically associated domains (TADs) and A/B compartments formed therefrom (Preissl et al., 2023). In 3D genome architecture, distal enhancers are placed spatially close to their targeted gene promoters. Single-cell profiling of chromatin architecture is thus needed to understand how different genomic regions interact with each other in individual cells. High-throughput/resolution chromosome conformation capture (Hi-C) is a conventional approach that involves capturing chromatin architecture in bulk samples. Based on this approach, scHi-C was developed to capture chromatin architecture at a single-cell level, which is also the most widely applied method. In the original study, only 10 T helper cells were assayed, with 0.01–0.03 million contacts detected in each cell (Nagano et al., 2013). By omitting the biotin-related step, Flyamer et al. (2017) detected a median of 0.34 million contacts. Subsequently, Tan et al. (2018) increased the number of contacts to a median of 1 million per cell by employing a Tn5-based whole-genome amplification protocol. In this method, Tan et al. (2021) used unindexed Tn5 and profiled more than 3,000 cells, with approximately 0.4 million contacts per cell. Meanwhile, combinatorial barcoding methods were employed in single-cell profiling of chromatin architecture. Although thousands of cells can be profiled otherwise, the resulting data are sparser than those yielded by the scHi-C method. To address this point, Mulqueen et al. (2021) employed a Tn5-based library strategy and increased the number of contacts to approximately 0.1 million per cell. In addition, several rounds of split-pool barcoding were used by Arrastia et al. in their recent introduction of single-cell split-pool recognition of interactions by tag extension (scSPRITE) to capture TADs, compartments, and inter-chromosomal interactions (Arrastia et al., 2022).

Single-cell sequencing of other modalities

In addition to the profiling of genome, transcriptome, and epigenome in individual cells, single-cell genomics has branched out into other modalities like proteins, metabolites, and ribosomes. First, to comprehend the proteomic constitution of individual cells, several distinct methods have been exploited, including: (i) mass spectrometry (MS)-based; (ii) antibody-based; (iii) imaging-based; (iv) single-molecule sequencing (Redit et al., 2023). Among these strategies, MS-based methods (e.g., proteoCHIP) have achieved the most success. Second, to catalog the chemical contents of individual cells, an array of novel technologies has been developed. For example, scMEP uses high-dimensional antibody-based methods to quantify proteins

regulating metabolic pathways for the characterization of the metabolic regulome of individual cells (Hartmann et al., 2021). Third, to understand mitochondrial DNA in individual cells, methods such as mtscATAC-seq (Lareau et al., 2023; Lareau et al., 2021a) have been developed to detect chromatin accessibility and, thereby, mitochondrial genotypes.

Chapter 3. Single-cell multi-omics technologies

Introduction to single-cell multi-omics

Recent advances in single-omics technologies have ushered in a new era of single-cell analysis that has facilitated detailed examination of cellular diversity across multiple biological layers—from genomics and epigenomics to transcriptomics, translatomics, and proteomics (Figure 1). This evolving field has substantially deepened our current understanding of cellular heterogeneity and developmental trajectories in health and disease. To elucidate the complex interplay between these layers, concurrently profiling multiple molecular modalities within the same cell is crucial. Recent years have witnessed the emergence and refinement of single-cell multi-omics approaches that integrate genomics and transcriptomics, which have now expanded to include epigenomics, capturing DNA methylation, chromatin accessibility, histone modifications, and chromatin organization. These approaches also extend to proteomics, epigenomics, and metabolomics, which provide a holistic view of cellular functions and the central dogma.

Single-cell multi-omics profiles have revealed new insights into the identification of sub-clusters, cellular heterogeneity, and the impact of genomic and epigenomic variation on biological functions. Principally, multi-omics studies aim to detect co-localization or co-occurrence of omic features within individual cells or across cell types and lineages to provide a basis for inferring causal relationships between chromosomal information and molecular manifestations, including mRNA, ncRNA, and proteins. Integrating bioinformatics analysis to construct trajectories introduces a temporal aspect to these data, which can enhance our understanding of cellular dynamics. Moreover, the advent of spatial omics has augmented our knowledge of cellular distribution within tissues. Single-cell multi-omics technologies are particularly advantageous for measuring limited samples, enabling comprehensive multi-dimensional analysis without needing sample division for separate assays, which potentially reduces the overall costs.

Single-omics methods have laid the groundwork for multi-omics approaches. The transcriptome plays a central role in linking genotype to phenotype as a crucial intermediary between the genome and proteome. As such, transcriptomics, being one of the most mature and earliest-developed omics, offers multiple strategies and has been established as a cornerstone in the evolution of multi-omics methodologies. Multi-omics techniques, which are rooted in single-omics methods, retain the major distinct procedure, such as the initial cell dissociation and collection via cell capture platforms, including microfluidics, mouth-pipetting, flow cytometry, or sophisticated labeling strategies like split-and-pool barcoding (Figure 2). Subsequently, the information from various modalities is transformed into features that allow sequencing through diverse strategies and deciphering through data processing and integration. Crucially, multi-omics approaches underscore the necessity of separating modalities to preserve data integrity and analytical efficiency—a principle that can be considered as an essential one of multi-omics.

Multi-omics methods have enhanced our understanding of cell state transitions, gene expression profiles, and regulatory networks in normal tissues and for diseases across species. Further applications of single-cell multi-omics have been elaborated in some previous reviews (Baysoy et al., 2023; Vandereyken et al., 2023). In this section, we have first addressed the various strategies known for modality separation, which is a key distinguishing feature between multi-omics and single-omics methods. We have then highlighted representative methodologies to illustrate the progression and state-of-the-art advancements in multi-omics. In the discussion, we have presented a comprehensive technical timeline on the combined detection and separation of modalities developed over the past decade.

Modality separation in multi-omics

Over the past decade, multi-omics methodologies have advanced into systems that incorporate multiple strategies to distinguish among modalities. An effective multi-omics approach ensures the maximum efficiency of each modality while minimizing signal loss and contamination or interference. Other beneficial factors include compatibility with standard laboratory setups, cost, and hands-on time. Broadly speaking, the evolution of multi-omics techniques has been guided through the advancements made with single-omics methods. For example, the recent development of the CUT&Tag method represents an advancement over the conventional single-cell ChIP-seq by efficiently labeling modification sites using Tn5 barcoding (Bartosovic et al., 2021; Wu et al., 2021). This advanced method leverages the simplicity of segregating barcoded tags from other modalities, thereby forming the basis for many chromatin-associated protein bi-omics methods that employ a Tn5 conjugated antibody recognition framework. However, multi-omics is more than a mere aggregation of single-omics techniques. Here, each step is meticulously optimized; for instance, (i) it involves a robust cell carrier platform to enhance efficiency and throughput; (ii) each modality is distinctly isolated to prevent cross-contamination and signal loss; (iii) a precise and convenient protocol has been established for amplification and library construction; (iv) there is an unbiased data acquisition process and the availability of appropriate downstream analysis tools. Multi-omics technologies leverage the platforms, protocols, and analytical techniques of singleomics. However, it is the challenge and the necessity of modality separation that fundamentally distinguishes multi-omics. We have categorized the modality separation strategies into three: physical separation, *in silico* separation, and enzymatic conversion (Figure 3). Moreover, we have delineated amplification strategies into separated and co-amplification approaches.

Physical separation

Physical separation techniques partition target molecules into distinct pools for individual library construction. A fundamental method involves extracting a fraction of a cell lysate for personalized profiling. This strategy serves as a cornerstone for current transcriptome and translatome bi-omics methods, which rely on direct lysate distribution (Hu et al., 2016b; Zou et al., 2022). Despite potentially losing half the signal, physical

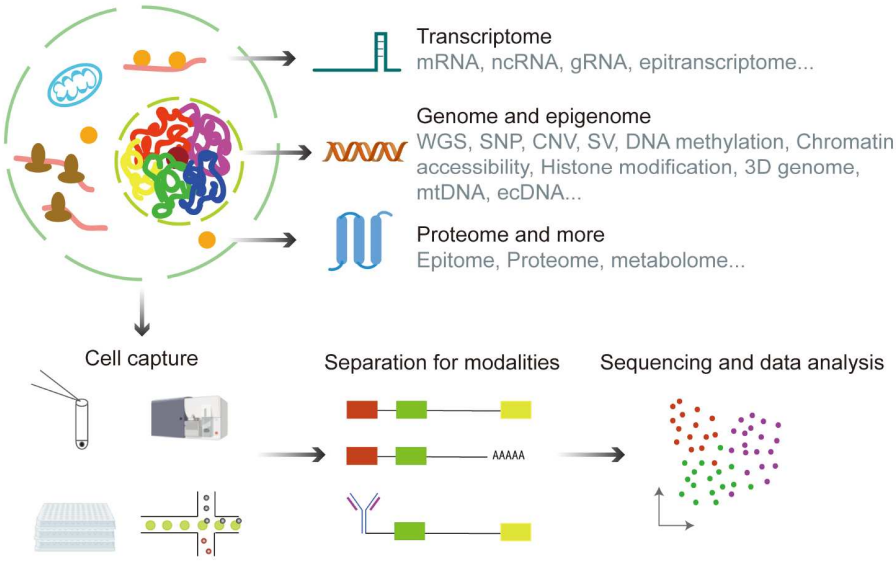

Figure 2. Workflow for single-cell multi-omics detection. Cells containing multiple omics information are collected through different platforms similar to that by single-omics methods. After the modalities are distinguished through different strategies, libraries containing sequence information and barcodes are sequenced, and the information is processed with pipelines to separate and integrate multiple omics. Different colors indicate distinguished molecules, sequences, or features in individual graphs.

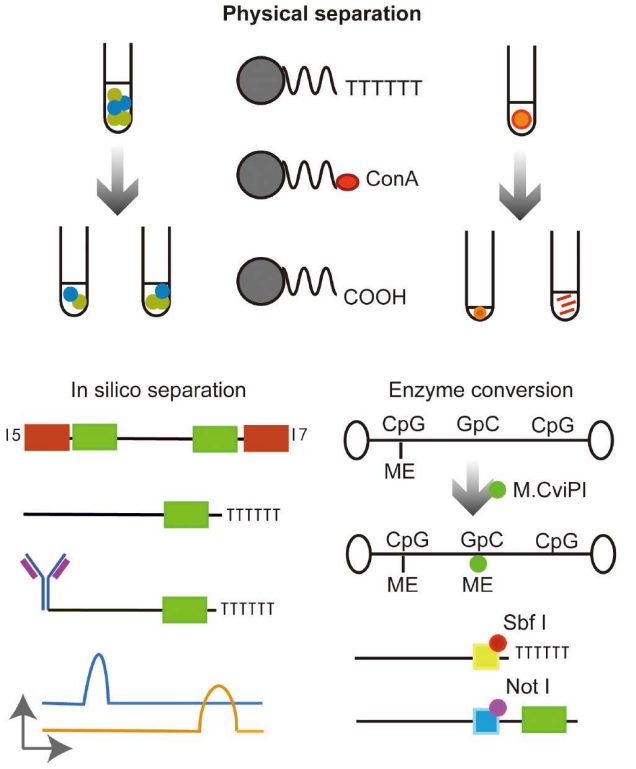

Figure 3. Different classes of modality-separation strategies. Strategies are classified into three groups: physical separation, *in silico* separation, and enzyme conversion. Each group includes multiple methods, as described in the main text and selectively drawn. Different colors and shapes indicate different molecules or elements in each method. Specifically, solid balls indicate enzymes or molecules; the green box indicates barcode position (in different sequences); the red box indicates Tn5 insertion i5/i7 sequence; the yellow or blue box indicates enzyme-cutting sites.

separation can efficiently isolate molecules and maintain the integrity of labile entities, such as RNA and proteins, under various lysis conditions. This separation process can occur post-profiling as well; for example, sci-CAR amplifies cDNA and fragmented DNA from divided lysate portions (Cao et al., 2018).

For enhanced separation, magnetic beads are used to isolate specific molecules. For example, oligo(dT) beads are used to extract poly-A RNA from lysates for use in downstream single-cell RNA-seq protocols. In SCTG, genomic DNA is purified with AMPure XP beads or ethanol, which may incur additional losses (Li et al., 2015). Simultaneous high-throughput ATAC and RNA expression with sequencing (SHARE-seq) combines combinatorial indexing with bead-based isolation of cDNA to improve throughput (Ma et al., 2020b). Bead separation has been integrated into advanced tri-omics methods such as single-cell chromatin accessibility, RNA barcoding, DNA methylation sequencing (scChaRM-seq), single-cell nucleosome occupancy, methylome and RNA expression sequencing (scNOMeRe-seq), and single-nucleus methylcytosine, chromatin accessibility, and transcriptome sequencing (snmCAT-seq) for mRNA isolation, thereby highlighting the technique's utility and popularity (Luo et al., 2022; Wang et al., 2021d; Yan et al., 2021).

Nuclei separation, which involves a two-step lysis, segregates intact nuclei from cytosolic components. After gentle cell lysis, nuclei are extracted by following different protocols. One pioneering multi-omics method employs microfluidics to partition nuclei for DNA retrieval (Han et al., 2014). High-speed centrifugation is also an effective means of isolating nuclei, as demonstrated in scTrio-seq (Bian et al., 2018; Hou et al., 2016). Plate-based centrifugation has increased throughput for isolation, as observed with scCAT-seq or direct nuclear tagmentation and RNA sequencing (DNTR-seq) (Qu et al., 2019; Zachariadis et al., 2020). In addition, beads can facilitate nucleus trapping to form visible clumps for easy manipulation. In the simultaneous isolation of genomic DNA and total RNA sequencing (SIDR-seq), antibody-conjugated beads have been used for nuclear precipitation (Han et al., 2018a), while scSET-seq employed Concanavalin A conjugated paramagnetic beads to bind specific membrane glycoproteins (Sun et al., 2021a). A recent comprehensive study conducted on colorectal cancer employed carboxylic acid beads

for profiling genomic DNA and mRNA bi-omics (Zhou et al., 2020). Such a bead-assisted nuclear separation method has been incorporated into LiMCA—single-cell transcriptomics plus 3D genome methodology (Wu et al., 2024a). Nuclei separation effectively minimizes cross-contamination between nuclear and cytosolic analytes but may result in the loss of molecular complexity, which affects components such as nuclear mRNA and mitochondrial DNA.

The aforementioned strategies aim to separate modalities before amplification or, in the initial stages, to ensure complete differentiation in subsequent analyses. However, these approaches often entail lower throughput, require more time and resources, and may cause some signal loss. An enhancement to this approach is implementing physical separation post-amplification. For this reason, DR-seq is devised to co-amplify cDNA and gDNA, followed by partitioning the amplicons into distinct DNA and RNA libraries (Dey et al., 2015). Although cDNA can be tagged with specific T7 adaptors, there remains a risk of cDNA contamination in the gDNA library. SNARE-seq employs a splint oligo complementary to the Tn5 adaptor sequence post-tagmentation, which enables the simultaneous capture of chromatin accessibility and transcriptome. This method allows unique cell barcoding within each droplet, which helps streamline the process (Chen et al., 2019).

Enzyme conversion

Enzyme conversion leverages enzyme-driven chemical modifications to distinguish or label different modalities. Commonly, this process involves the use of a DNA methylase, such as M.CviPI, to modify GpC sites in open chromatin regions that are not bound to nucleosomes, resulting in the conversion to methylated cytosines (Kelly et al., 2012). Subsequently, unmethylated cytosines are converted into uracil through bisulfite conversion. This method allows for the simultaneous profiling of naturally methylated CpG sites across the genome and enzyme-converted GpC sites within accessible chromatin. Techniques such as ScTrio-seq2, scCOOL-seq, NMT-seq, and scNOMe-seq have adopted and refined this approach to create a unified set of methodologies (Bian et al., 2018; Clark et al., 2018; Guo et al., 2017; Pott, 2017). Although this strategy tends to favor cytosine-rich loci and provides a higher coverage around transcription start sites, it can be integrated with RNA-seq. For example, cnmCAT-seq incorporates 5′-methyl-dCTP during cDNA synthesis in RT, thereby protecting cDNA from bisulfite damage and eliminating the need for separate mRNA isolation (Luo et al., 2022).

Another enzyme conversion strategy, as exemplified by Paired-seq and Paired-tag, employs restriction enzymes to differentiate libraries post-amplification (Zhu et al., 2019; Zhu et al., 2021). In this process, following split-and-pool barcoding of single cells, cDNA and DNA are tagged with distinct adaptor sequences. The pre-amplified mixture is then partitioned and selectively cleaved with NotI or SbfI. Strategies based on enzyme conversion heavily depend on the specific performance of enzymes and represent areas primed for further exploration and development.

Post-sequencing separation

Another significant category of the separation method involves either co-amplification or copool of different modalities, which distinguishes them after sequencing. In some instances, modalities can be directly identified from the sequence. More commonly, modalities requiring indirect identification utilize barcodes. Primers or adaptors are often employed to append barcodes for targeting molecules during RT or PCR amplification. Barcode separation offers several advantages, as follows: (i) it obviates the need for additional separation steps, as barcodes are integrated during profiling; (ii) it facilitates high throughput by combining modality and cell barcodes; (iii) it minimizes signal loss; (iv) it distinguishes between physically similar omics, such as the various histone modifications. Tn5 transposases, which cleave and tag double-stranded DNA or RNA-cDNA hybrids with adaptors, are useful for barcode addition (Penkov et al., 2023).

In summary, each modality separation strategy has distinct applications. Physical separation typically coincides with separate amplification, whereas enzyme conversion and *in silico* separation are generally employed before co-amplification. Some methods generate individual libraries for each modality to optimize the sequencing depth, while others construct a mixed library to streamline procedures and reduce the processing time. A single method may integrate multiple strategies, especially in tri-omics technologies. Although transcriptome profiling has utilized all of these three strategies, specific combinations such as DNA methylation and chromatin accessibility are predominantly associated with a single strategy, emphasizing its unique advantage. Thus, the choice of separation strategy and timing is dictated by the modalities involved, and these strategies continue to evolve to better accommodate diverse modalities.

Development of multi-omics

Over the past decade, single-cell multi-omics methods have undergone significant evolution, progressing from their initial, low-throughput versions that covered a limited range of modalities to sophisticated, comprehensive systems that incorporate a variety of strategies and encompass nearly all modalities. We have compiled here a selection of these methods, ranging from 2014 to 2023, in Table S2. Notably, there has been a marked acceleration in the development of new methodologies, particularly since 2021. The integration of modalities, including the genome, epigenome, transcriptome, and other cytoplasmic analytes, into the multi-omics toolkit has progressively expanded during this period.

Genomics and epigenomics

The Central Dogma underscores the flow of genetic information from DNA to RNA, highlighting a key objective of multi-omics, i.e., exploring the connection between genomic or epigenomic variation and transcriptional output. G&T-seq and DR-seq exemplify two primary methodologies that have distinct approaches (Dey et al., 2015; Macaulay et al., 2015). For instance, G&T-seq employs oligo-dT-covered beads to capture full-length mRNA for profiling via Smart-like protocols, while gDNA is processed using methods such as MDA or PicoPlex, whichever is applicable. Conversely, DR-seq co-amplifies cDNA and gDNA to minimize sample loss, although it possibly compromises full-length mRNA detection and introduces potential cDNA interference in DNA signals. TARGET-seq, like DR-seq, employs co-amplification for targeted mutation detection alongside cDNA and gDNA primers (Rodriguez-Meira et al., 2019). Sci-L3 increases throughput by co-amplifying DNA and RNA with distinct barcodes and combinatorial indexing (Yin et al., 2019). ScONE-seq facilitates a one-tube reaction for genomics and transcriptomics with Tn5 insertion post-RT (Yu et al., 2023a),

enabling separate sequencing of DNA and RNA libraries at optimal depths.

Following the widespread utilization of genomics, studies integrating epigenomics have emerged. While genomics concentrates on individual variations or mutations, epigenetics highlights cell differentiation and gene expression regulation. The first epigenetic multi-omics methods primarily focused on DNA methylation. In 2016, MT-seq, scTrio-seq, and M&T-seq, based on RRBS or genome-wide BS-seq, followed a separation strategy similar to that of G&T-seq, thereby providing a foundational framework for DNA methylation multi-omics (Angermueller et al., 2016; Hou et al., 2016; Hu et al., 2016b). Smart-RRBS, a recent method, involves the application of tens to thousands of cells without explicit nucleic acid extraction (Gu et al., 2021). Newer methods using enzyme-based methylation site conversion or third-generation sequencing reduce DNA loss to achieve higher coverage, implying potential multi-omics applications (Cao et al., 2023; Sun et al., 2021b).

The earliest chromatin accessibility multi-omics did not stem from existing single-omics methods such as DNase-seq, MNase-seq, or ATAC-seq (Buenrostro et al., 2015; Lai et al., 2018; Song et al., 2011). Enzyme conversion using GpC methylase was first employed in scCOOL-seq and scNOMe-seq for profiling open chromatin coupled with DNA methylation (Guo et al., 2017; Pott, 2017). ScNMT-seq physically isolated mRNA from nuclei, subsequently profiled via scRNA-seq, single-cell DNA methylation, and chromatin accessibility, ultimately becoming one of the earliest tri-omics methods, followed by methods such as scNOMeRe-seq, scChaRM-seq, scCAT-seq, and snmCAT-seq (Clark et al., 2018; Luo et al., 2022; Wang et al., 2021d; Yan et al., 2021) (Liu et al., 2019). Sci-CAR, the initial ATAC-based method, combined direct splitting of cDNA and DNA fragments with combinatorial indexing of pooled cells (Cao et al., 2018) to profile thousands of single cells into a single experiment. Since 2018, ATAC-related Tn5 fragmentation has been widely adopted in chromatin accessibility-associated methods owing to its convenience and reliability, such as SHARE-seq (Ma et al., 2020b), SNARE-seq (Chen et al., 2019), Paired-seq, and *in situ* sequencing hetero RNA-DNA-hybrid after assay for transposase-accessible chromatin-sequencing (ISSAAC-seq) (Xu et al., 2022b). Chromatin accessibility has become a frequently involved omics, second only to transcriptome (Cheng et al., 2021a; Qu et al., 2019; Xing et al., 2020).

Histone modifications and TFs, targeted by Tn5-conjugated antibody fragmentation, were initially reported in bulk methods and then quickly adapted to single-cell analysis (Kaya-Okur et al., 2019). Methods such as CoTECH, Paired-Tag, and scSET-seq, which employ insertion strategies similar to ATAC-related techniques and incorporate transcriptomics, have also been developed (Sun et al., 2021a; Xiong et al., 2021; Zhu et al., 2021). Alternatively, scPCOR-seq uniquely uses MNase conjugated to protein A for histone site localization (Pan et al., 2022). Since 2020, the simultaneous profiling of multiple chromatin-associated proteins within the same cells has become feasible (Gopalan et al., 2021; Janssens et al., 2022a; Lochs et al., 2024; Meers et al., 2023; Stuart et al., 2022; Tedesco et al., 2022; Yeung et al., 2023). This strategy has also been combined with epitope detection, as in scCUT&Tag-pro, wherein epitopes were profiled using oligo-conjugated antibodies alongside scCUT&Tag data (Zhang et al., 2022a). Another strategy involves the use of DNA adenine methyltransferase (Dam) to methylate adenines at protein-associated GATC sites, marking DNA sites' interaction with labeled proteins. ScDam&T-seq and EpiDamID are designed to cleave these methylated GATC sites in single cells, coupled with IVT-based amplification (Markodimitraki et al., 2020; Rang et al., 2022). DamID relies on *in vivo* transfection and reaction of Dam fusion proteins; hence, the results often reflect cumulative interactions or residence time rather than immediate protein binding states.

Finally, the relatively underdeveloped yet crucial aspect of epigenomics lies in understanding the 3D genome. Hi-C, a widely utilized sequencing method, is based on spatial proximity ligation of DNA fragments for profiling chromatin architecture (Lieberman-Aiden et al., 2009). The initial 3D genome-related multi-omics approaches were developed in combination with DNA methylation. ScMethyl-HiC and snm3C-seq detect bisulfite-converted methylated sites along with ligated fragments from different genomic loci (Lee et al., 2019; Li et al., 2019). Despite potential contact reductions due to bisulfite conversion, these methods have displayed correlations between methylated CpG loci and higher-order chromatin structures, particularly compartments. The first transcriptome and 3D genome method, named HiRES (Hi-C and RNA-seq employed simultaneously), was reported recently (Liu et al., 2023f). HiRES employs barcode separation of DNA and cDNA, followed by co-amplification of both ligated genomic DNA and cDNA in a single-tube reaction. Several methods quickly followed this report, combining various strategies such as nuclei separation, barcode separation, and combinatorial indexing (Luo et al., 2023; Wu et al., 2024a; Zhou et al., 2023). Benchmarks for these methods are essential for further evaluation. Considering the integral role of the 3D genome in reflecting chromatin dynamics, it is likely interconnected with other omics aspects. Therefore, new methods that integrate the 3D genome with other omics are eagerly anticipated.

Our discussion primarily focuses on nuclear DNA, which differs from cytoplasmic DNA, including extrachromosomal DNA (ecDNA) and organelle DNA, such as mitochondrial DNA (mtDNA), which are often excluded from nuclei separation. Nonetheless, some studies have incorporated the detection of these DNA types. In PHAGE-ATAC, phage-mediated tagmentation enabled the detection of mtDNA mutations (Fiskin et al., 2022); mtscATAC-seq adapted a modified 10x ATAC chromium system for mtDNA detection, which traces thousands of cancer cells into clones and sub-clones (Lareau et al., 2023). ScEC&T-seq physically separated DNA, allowing ecDNA to be identified by its circular feature (Chamorro González et al., 2023); scGTP-seq similarly captured ecDNA by using a single-molecule real-time sequencing of long fragments amplified through transposon insertion (SMOOTH-seq) like processes (Chang et al., 2023a). Both methods confirmed a high correlation between oncogene expression and ecDNA formation.

CRISPR-mediated gene editing, when combined with single-cell approaches, significantly enhances our understanding of genomic element functions. CRISPR-based methods such as CRISP-seq, CRISPR droplet sequencing (CROP-seq), and Perturb-seq co-profile transcriptomics, are often implemented on plates or microfluidics, which enable genome-wide screening (Datlinger et al., 2017; Dixit et al., 2016; Jaitin et al., 2016). Recent advancements involve integrating gRNAs with ATAC-seq to explore the effects of gene perturbation on chromatin accessi-

bility. Perturb-ATAC and Spear-ATAC allow single-cell ATAC-seq on microfluidics platforms, which enriches reverse-transcribed gRNA with barcoded primers (Pierce et al., 2021; Rubin et al., 2019). CRISPR-sciATAC increases throughput with combinatorial indexed ATAC-seq and two-round PCR enrichment of gRNA (Liscovitch-Brauer et al., 2021). Expanded CRISPR-compatible cellular indexing of transcriptomes and epitopes by sequencing (ECCITE-seq) incorporates protein information, facilitating precise sub-clustering of immune cells based on the receptor type (Mimitou et al., 2019).

Post-transcription omics

Multi-omics technologies are mainly centered around transcriptomics, not only because of the extensive development and sophistication of transcriptomic methodologies but also because of its central role in the flow of genetic information. However, RNA profiling encompasses more than just mRNA quantification. For example, full-length RNA sequencing provides additional insights into alternative splicing and allelic expression; these aspects may be lost in methods that focus solely on 3′- or 5′-enriched cDNA. Another critical aspect is the posttranscriptional modifications of RNA, which play crucial roles in mRNA regulation, RNA metabolism, and translation activity. One notable modification in mammalian mRNA and lncRNA is m^6A (Sun et al., 2019). Recently, sn-m^6A-CT has been developed based on a 10x Chromium platform to simultaneously profile the full transcriptome and m^6A methylomes (Hamashima et al., 2023). These m^6A methylomes were found to effectively differentiate cell types and exhibit variations in several pluripotent marker genes.

The coupling of transcription and translation has raised increasing attention recently. Ribo-seq enables the profiling of ribosome-bound mRNA both at the population and single-cell levels (VanInsberghe et al., 2021; Xiong et al., 2022). However, detecting single-cell ribosome-bound mRNA poses challenges owing to its scarcity and potential loss. T&T-seq and R2-lite have successfully separated transcriptome and ribosome-bound translatome in mammalian oocytes or embryos (Hu et al., 2022; Zou et al., 2022). T&T-seq captures ribosome-mRNA complexes with A-site beads, whereas R2-lite directly splits the lysate for individual profiling. Notably, R2-lite identified pre-produced mRNA in oocytes functional in early embryos, thereby potentially initiating zygotic gene activation. Hence, single-cell translatome profiling can help uncover post-transcriptional regulation across cell types by following a method that applies to cells of standard size.

Post-translation omics

Cytosolic analytes, which encompass a diverse array of molecules in the cytoplasm, present unique challenges and opportunities for single-cell analysis. Unlike nucleic acids, proteins cannot be directly amplified or sequenced, which poses significant hurdles in single-cell proteomics. Nevertheless, proteins can be indirectly profiled by using oligo-conjugated antibodies. Techniques such as cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq), as well as RNA expression and protein sequencing assay (REAP-seq), have successfully helped profile surface epitopes with dozens of oligo-conjugated antibodies, thereby facilitating transcriptome-like analysis. The extensions of this approach to intracellular proteins have been achieved using methods such as Single-Cell Protein and RNA Co-profiling, single-cell RNA and immunodetection (RAID-seq), and inCITE-seq (Chung et al., 2021; Gerlach et al., 2019; Reimegård et al., 2021). However, this strategy is limited to proteins for which efficient antibodies are available, thereby restricting its scope. A recent method, single-cell simultaneous transcriptome and proteome (scSTAP), tackles this issue by dividing oocyte lysate (like R2-lite) and profiling RNA and proteins separately through MATQ-seq and MS-based techniques (Jiang et al., 2023c).

Metabolomics focuses on small molecules, such as lipids or sugars, that play critical roles in signaling pathways and energy processes. Metabolic profiling typically relies on MS, which may cause irreversible damage to the molecules. Research in the field of single-cell metabolomics and its combination with proteomics is emerging (Li et al., 2021b; Shrestha, 2020). One of these approaches involves the use of oligo-EG conjugated microwells to extract metabolites from adsorbed proteins for separate MS analysis. However, metabolomics methods remain on the fringe of multi-omics, and there is a need to develop more innovative combinations.

New era in single-cell multi-omics: tech choices and further applications

Overall, currently, known single-cell multi-omics can be characterized based on several key features: (i) the majority of current methods focus on a combination of transcriptomes with other genomic modalities such as chromatin accessibility, DNA methylation, chromatin-associated proteins, and 3D genome; (ii) a range of previously unexplored modalities have recently been incorporated into multimodal assays; (iii) there is a rapid increase in the number of new methods; (iv) established methods are progressively being applied to address biological problems and then contribute to the construction of tissue atlases. Consequently, the advancement in the development of new monomodal profiling workflows and strategies for modality separation are propelling multi-omics research forward.

Ideally, a mature multimodal method should offer comparable efficiency across each single-ome, feature a user-friendly protocol, and be cost-effective with the availability of accessible downstream analysis tools. Certain combinations, such as integrating 3D genomes with chromatin accessibility and histone modifications or tri-omics involving transcriptome, translatome, and proteome, are currently unachievable. For combinations that involve multiple methodological variations, systematic comparison of these methods and their standardization is essential for widespread, large-scale applications. Here, we compared several strategies with throughput, time, costs, availability, and core instruments (Table 1). For laboratories with limited single-cell experience, fewer cell diversity, and simple biological models, some easy and cheap methods can be sufficient, such as those including physical separation methods or mouth-pipetting platforms. Microfluidics, combinatorial barcoding, or *in silico* optimization can benefit a throughput design that serves huge atlas projects or complex systems. As mentioned earlier, multi-omics methods greatly depend on modality separation and data quality. Modality contamination or exogenous contamination should be particularly avoided because multiple modality reactions extend over the entire method period, which introduces additional potential risks. Some methods share similar limitations, such as the excessive loss in direct separation or RNA-to-DNA contamination in *in silico*

https://doi.org/10.1007/s11427-024-2770-x

Table 1. Comparison of different strategies, including modality separation and single-cell platforms

Strategy	Throughput	Time	Cost	Availability	Core instruments
Direct lysate distribution	Medium	Fast	Low	Easy	Pipette
Bead separation	Low	Medium	Medium	Medium	Molecules binding beads
Nuclei separation	Low	Medium	Medium	Medium	Nuclei binding beads
Enzyme conversion	Usually High	Fast	Medium	Easy	Converting or cutting enzymes
In silico separation	Usually High	Medium	Low	Required well-designed algorithms	Good PC
Mouthpipette	Low	Slow	Low	Required skillful person	Mouthpipette tube
Plate sorting	Medium	Medium	High for instrument	Required FACS in clean room	FACS for single-cell plating
Microfluidics	High	Medium	High for instrument	High requirement	Microfluidics
Combinatorial barcoding	High	Fast	Medium	Required primers	Plates and barcode primers

methods such as DR-seq. As there is no perfect method established yet, some trade-offs should be made when selecting a method that better serves the project.

It has been shown in the past that when compared with single dimensional methods, multi-omics methods mainly study an internal correlation among diverse cell types of different modalities. In the long past, scientists applied DNA/RNA bi-omics to study mutation and clonal expansion with crucial gene switching in leukemia (Zachariadis et al., 2020). These correlations were further verified as causality by methods, including gene editing (Bhattarai-Kline et al., 2022). As multi-omics are centralized with transcriptomes, RNA information can help cluster basic cell types in complex tissues or provide an output for gene regulation of the genome or epigenome. RNA dynamics reflect a cell lineage within the development process, which produces a temporal perspective to study other omics. The potential interactions of different epigenomes can be seen in bi-epiomics methods; for instance, DNA methylation can mold chromatin architecture to present a diversity of neural cells (Lee et al., 2019; Tian et al., 2023b). A widespread application of post-transcription methods, including translatome, proteome, or metabolome, remains short. These methods can be gradually unlocked with the development of new methods or improvement in efficiency. Alternative strategies developed in the future can facilitate the establishment of a comprehensive system and setting up industry benchmarks. Moreover, computational methods have evolved to precisely distinguish different modalities and even integrate different single-ome data into "multi-omics integrated" data. With the rapid evolution of single-cell multi-omics technologies, we can expect more insightful discoveries associated with diverse fundamental biological processes such as development, aging, and diseases.

Chapter 4. Spatial genomics technologies

Introduction of spatial genomics

"The Human Genome Project" stands as a pivotal moment in contemporary biology, which initiated humanity's systematic capability to decipher the genetic code of living organisms (International Human Genome Sequencing Consortium, 2004). Positioned as a transitional milestone in the evolution of genetic molecular detection techniques, it has progressed from scrutinizing individual molecules to high-throughput molecular synchronization detection (Amaral et al., 2023). This progression also signifies the shift from a single-genome era to the current age and phase of multi-omics integration (Vandereyken et al., 2023). Spatial omics technology, which is emerging in the post-Human Genome Project era, aligns with the historical trajectory of genetic molecular detection methods and introduces a novel framework for unraveling the biological genetic code (Marx, 2021).

The method employed in this study to identify the relative spatial arrangements and expressions of tissues, cells, and biomolecules for deciphering the mechanisms of biological events is termed "spatial omics technology" (Larsson et al., 2021). Although various multi-omics approaches, such as spatial genomics and spatial proteomics, have emerged, the current focus of spatial omics technology predominantly centers on measuring and analyzing transcripts (Tang, 2021). This spatial detection technique, at the transcriptional level, has a long-standing history. In its early stages, direct detection of the localization of nucleic acid and molecules within tissues and cells posed a significant challenge. A clever solution involved the use of artificially introduced fluorescent probes for *in situ* hybridization of nucleic acids within tissues (Femino et al., 1998). By detecting the fluorescent probe, it acts as a surrogate for directly identifying nucleic acid molecules. In essence, fluorescent imaging can be employed to observe the spatial positions and abundance of nucleic acid expression within cell or tissue sections. This methodology has been recognized as spatial *in situ* hybridization (SISH) technology (Figure 4). Diverging from the principles of hybridization imaging technology, which entail chemically labeling base fragments for identification, the advancement and maturation of techniques such as PCR, ligation-based methods, and synthesis-based sequencing have facilitated the identification of individual bases in nucleic acid sequences with robust signals (Koboldt et al., 2013; Markham, 1993). For instance, in the case of *in situ* sequencing (ISS) (Ke et al., 2013), following the binding of target RNA with padlock probes, *in situ* amplification is performed, followed by ligation-based sequencing to ultimately determine the spatial characteristics of nucleic acid information. This technique, performing *in situ* base sequencing to glean genetic molecular information, is recognized as spatial ISS (SISS) technology.

With the accomplishment of "The Human Genome Project," the introduction of Roche's first 454 next-generation sequencers in 2005, and Illumina's release of the Genetic Analyzer 2 in 2007, the landscape of biological sequencing transformed low-throughput to high-throughput NGS (Metzker, 2010). In this high-throughput sequencing paradigm, all nucleic acids are extracted by subjecting them to biochemical reactions on the entire tissue. Subsequently, they are amalgamated in a single tube for sequencing, alignment, and assembly, which marks a

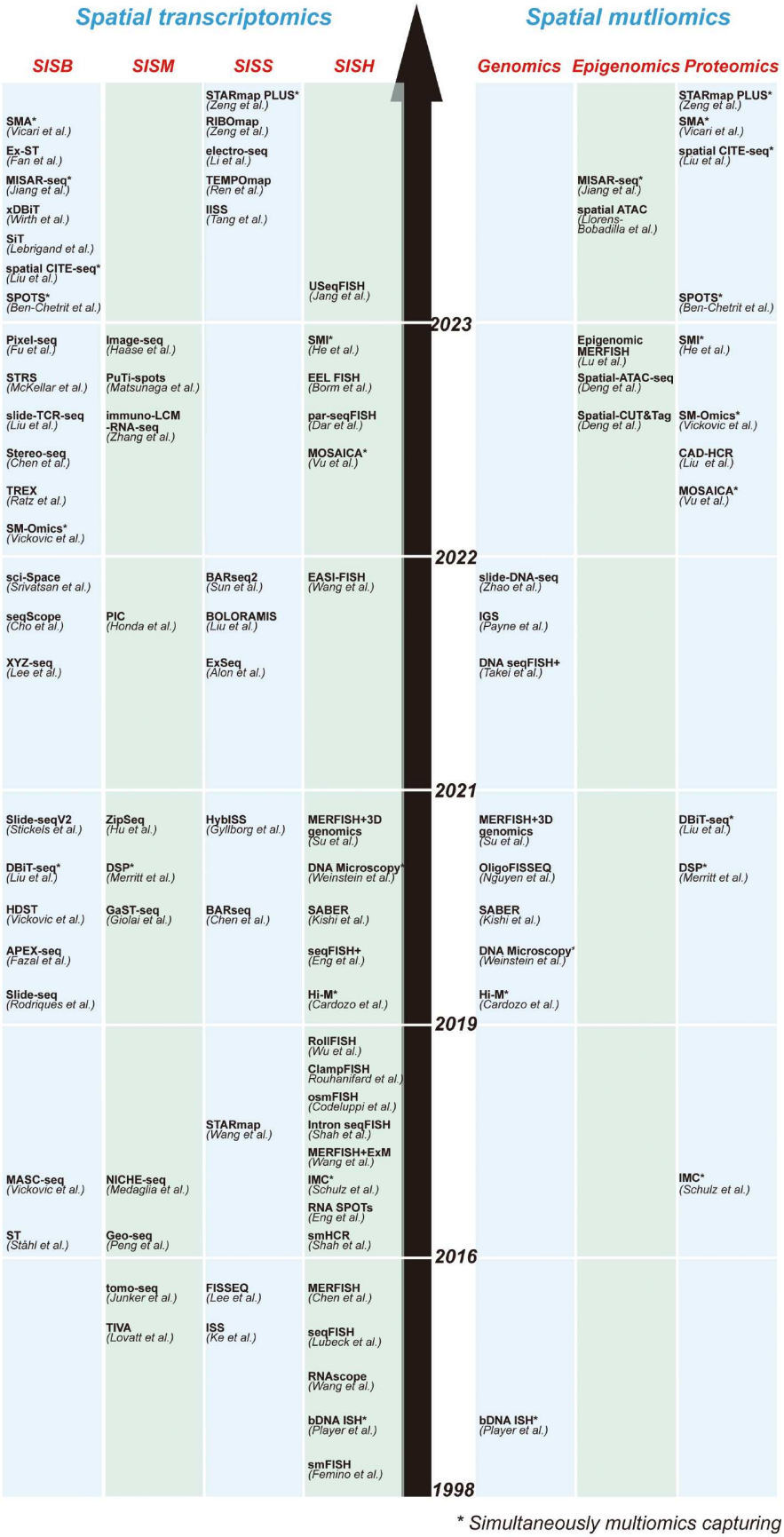

Figure 4. Timeline of spatial transcriptomics and spatial multi-omics methods.

https://doi.org/10.1007/s11427-024-2770-x

significant shift from the conventional workflow that involves sequential hybridizing and sequencing of individual nucleic acids. This shift also invited novel perspectives for spatial omics technology, giving rise to spatial *in situ* microdissection (SISM) technology. Through high-precision microdissection techniques, specific regions of tissue sections are physically separated, and the dissected tissues undergo high-throughput sequencing of mixed nucleic acid libraries (Emmert-Buck et al., 1996). This approach helps associate each sequencing result with a unique spatial position during the reconstruction of spatial locations. The early implementations of spatial SISM encompass Tomo-seq (Junker et al., 2014) and commercially available Digital Spatial Profiling (DSP) (Merritt et al., 2020) techniques. SISM involves the initial physical cutting of selected regions, followed by the assignment of spatial location markers. However, this process limits spatial resolution because of the constraints introduced by the physical cutting often performed on cell populations (Bressan et al., 2023). Although single-cell cutting is achievable, it remains exceptionally challenging. Consequently, the enhancement of spatial resolution has posed a limitation in the evolution of spatial omics technology. To overcome the constraints associated with physical operations, it is imperative to challenge the conventional logic of slicing first and then assigning spatial information. A breakthrough emerged in this direction by pre-labeling tissue cells and subjecting them to subsequent biochemical analyses, thereby mitigating the limitations of physical operations. Furthermore, the potential for independent labeling of individual cells or even nucleic acids within cells promised an escalation in the resolution from multicellular to single-cell or subcellular levels. The foundation of conventional sequencing technology can be referred for this innovative strategy. In single-cell sequencing approaches, barcode probes can pre-label each nucleic acid within individual cells, followed by biochemical reactions and sequencing for individual cells (Kolodziejczyk et al., 2015; Li and Wang, 2021; Sankaran et al., 2022). The advantages of barcode technology include the following: (i) advanced labeling of cells and nucleic acids, which circumvents the limitations of cutting first and then assigning spatial information. (ii) A vast array of barcode probes that theoretically exceed the number of nucleic acids in the tissues, ensuring a unique barcode for each nucleic acid. (iii) Integration of barcodes with polyT to form capture probes that are specific for the 3′ polyA tails of mRNA, thereby enabling unbiased transcriptome to capture instead of targeting. (iv) Implantation of barcode probes on diverse carriers, including common magnetic beads and modified glass slide surfaces. By leveraging the maturity of barcode technology and the progress in spatial omics, Ståhl et al. (2016) pioneered the implantation of barcodes on glass surfaces in 2016. This technique allowed unbiased capture of nucleic acids within covered tissues, which, when followed by *in situ* RT, can generate an abundance of amplified sequences for high-throughput NGS. This groundbreaking method, known as spatial *in situ* barcode (SISB) technology, achieved high-throughput transcription detection within tissue sections based on NGS in the spatial dimension. Thereafter, diverging from the implantation of barcode probes on flat surfaces, subsequent innovative techniques emerged. For example, Slide-seq (Rodriques et al., 2019) combines barcode sequences on magnetic beads, allows deterministic barcoding in tissue for spatial omics sequencing (DBiT-seq) (Zhang et al., 2020b), and introduces barcodes through fluidic channels, while Stereo-seq (Chen et al., 2022a) and Seq-Scope (Cho et al., 2021), provides single-cell precision through the development of sequencing chips.

In summary, spatial omics technology can be classified into two primary categories based on the distinct inherent principles for achieving spatial resolution and then subdivided into four subcategories: (i) Spatial Imaging: SISH and SISS; (ii) Spatial Sequencing: SISM and SISB. The rapid advancement of these spatial omics technologies has significantly expedited scientific research in the fields of developmental biology, neuroscience, pathology, botany, and other cutting-edge research domains (Moses and Pachter, 2022). These innovations offer unique biological insights into a range of previously unexplored life phenomena (Table S3).

Spatial technologies

SISH technology

smFISH (Femino et al., 1998) is regarded as the classical SISH; it involves the combination of individual RNA molecules in tissues with multiple fluorescently labeled probes. This technique allows the capture of images, facilitating the detection of RNA spatial localization and expressions. The fundamental logic underlying SISH can be summarized in two key steps: (i) Oligonucleotide Probes Binding to Targeted RNA Fragments by using specific oligonucleotide probes that bind to target RNA fragments and (ii) Capturing Images of Fluorescent Groups on Probes to detect signals and positions of RNA molecules. However, SISH encounters challenges with respect to signal-to-noise ratio and detection sensitivity, which necessitates optimization of these core steps with various solutions. However, enhancing the signal value and sensitivity of RNA molecule detection can be achieved by increasing the number of probes binding to each mRNA (Raj et al., 2008). Moreover, branch amplification techniques, such as branch DNA "double Z" probes (Player et al., 2001), hybridization chain reaction (HCR) (Shah et al., 2016), and rolling circle amplification (RCA) reaction can help enhance the detection signal following probe-target hybridization (Wu et al., 2018). ClampFISH (Rouhanifard et al., 2019) employs click chemistry for probe end hybridization and signal amplification is achieved through successive rounds of hybridization and click reactions in this method. The detection signal can also be amplified by applying the overall cycling smFISH strategy (Codeluppi et al., 2018) and through the incorporation of primer-exchange reactions (Kishi et al., 2019) (Figure 5).

The incorporation of spectral combinatorial labeling strategies and high-resolution, super-resolution microscopy in imaging can significantly augment the detection capabilities and sensitivity, thereby offering a solution for optical signal congestion. In seqFISH (Lubeck and Cai, 2012), each gene is designated a unique color post-hybridization, and upon probe stripping, the subsequent rounds of hybridization and colorimetric detection ensue. Theoretically, by employing 4 colors and 8 rounds of amplification, the entire genome can be encoded, yielding 4^8=65,536 potential codes. RNA SPOT expands the color palette to 12 "pseudo-color panels," and seqFISH+ (Eng et al., 2019) further extends it to 60 colors. Through the amalgamation of images and error-correcting hybridization, precise imaging of 10,000 genes can be achieved using this approach.

Another conventional method that employs spectral combinatorial barcoding is multiplexed error-robust FISH (MERFISH) (Chen et al., 2015). It characterizes genes as "1" and "0" by

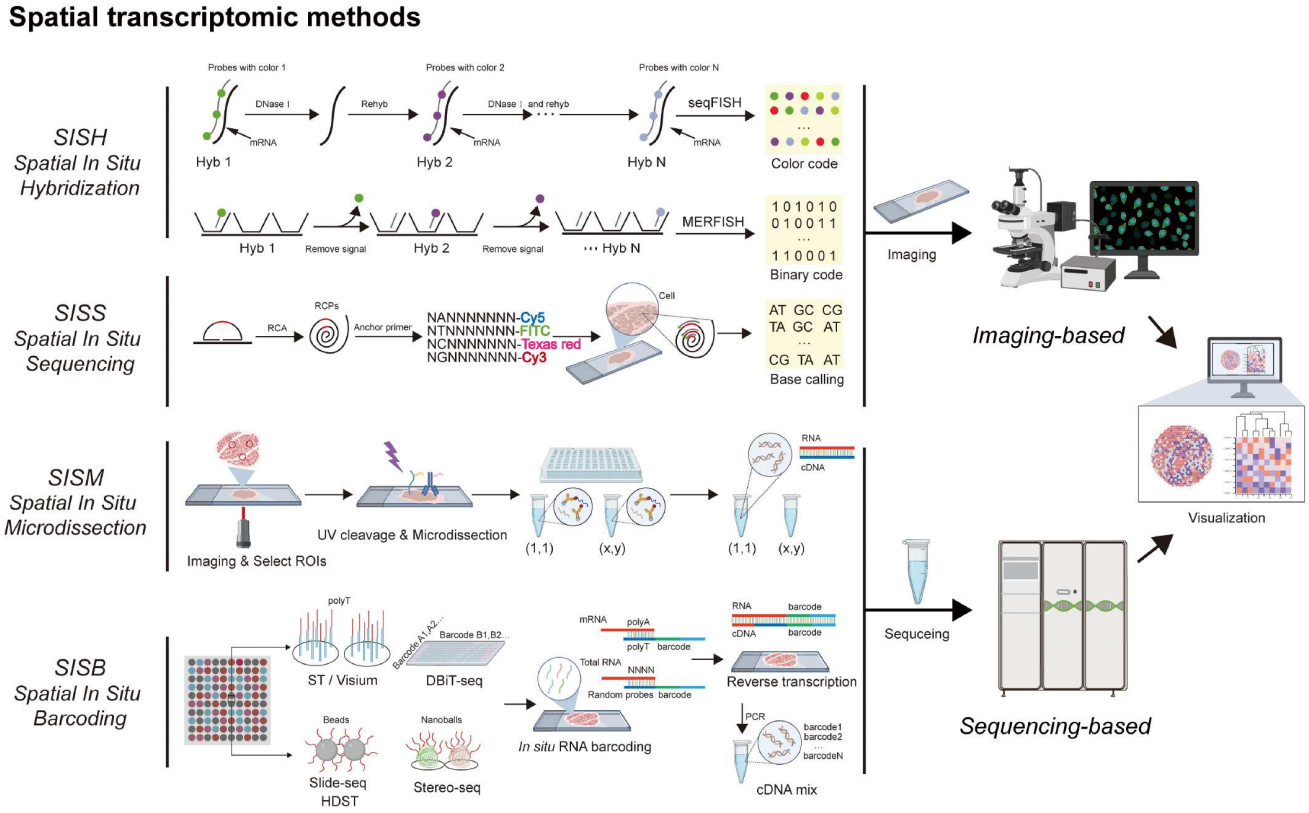

Figure 5. Overview of the four main spatial transcriptomics methods.

employing hybridization and colorimetric detection. Subsequently, the genes are encoded into a binary form through multiple rounds of hybridization. Error-correcting codes are implemented using a minimum Hamming distance, facilitating the detection of 10,000 genes. Integration with the expansion microscopy further enhances the overall density of detectable RNA, as with MERFISH (Wang et al., 2018a). Similarly, EASI-FISH achieves 3D detection of brain slices with a thickness of 300 μm by combining the use of hydrogel and amplifying the detection signal via the HCR (Wang et al., 2021c).

In addition to the conventional MERFISH and seqFISH techniques, spatial *in situ* detection methods inspired by the principles of smFISH have emerged, each offering unique technical characteristics and application scenarios. For instance, EEL FISH (Borm et al., 2023) utilizes electrophoresis to transfer RNA from tissue sections onto a capture surface for data acquisition, offering distinct advantages with respect to signal-to-noise ratio and speed. The par-seqFISH (Dar et al., 2021) method allows the recording of the gene expression and spatial background in micro assemblies at a single-cell level with molecular resolution. SISH has experienced rapid enhancements in signal-to-noise ratio, detection capabilities, sensitivity, and application scenarios with further development. Nonetheless, the intricacies of multiple hybridizations and limitations in detection capabilities due to optical congestion continue to impede significant technological breakthroughs.

SISS technology

The distinction between SISH and SISS lies in the necessity for a more robust nucleic acid signaling in sequencing so as to accurately detect and categorize each base, thereby completing the sequencing process. Consequently, the crux of SISS lies in the *in-situ* bulk amplification of nucleic acid fragments. PCR technology forms the fundamental logic for nucleic acid amplification, and the initial proposed solution for cloning and amplifying DNA on glass microscope slides using PCR reactions is Polony technology (Mitra and Church, 1999). In 2003, building upon Polony technology, Church's group (Mitra et al., 2003) introduced the first-generation Fluorescent ISS (FISSEQ) technology, enabling ISS of amplified DNA sequences from *in situ* PCR and demonstrating its application potential in tissues like bacteria and mice. In 2014, the second-generation FISSEQ technology emerged (Lee et al., 2014). Differing from its predecessor, this iteration directly captures non-targeted RNA molecules within tissue sections, reverse-transcribes them into cDNA, amplifies them through RCA, and employs phi29 polymerase. This process generates numerous continuous amplified copies fixed in the three-dimensional gel space, allowing for subcellular resolution ISS and showcasing potent application capabilities in tissues such as mouse brains, embryos, and *Drosophila* embryos (Lee et al., 2015). Another notable SISS is ISS, introduced by Ke et al. (2013) in 2013. ISS utilizes padlock DNA probes to bind to target nucleic acid sequences, followed by gap filling and ligation to form a complete circular structure. Amplification is then performed, and connection-based sequencing is employed to achieve spatial ISS. FISSEQ and ISS technologies have established the foundational framework for SISS, encompassing tissue and nucleic acid fixation, nucleic acid amplification, and ISS. These innovations have paved the way for subsequent SISS techniques with enhanced efficiency, resolution,

https://doi.org/10.1007/s11427-024-2770-x

and specificity.

Built upon ISS technology, the HybISS (Gyllborg et al., 2020) method integrates barcoded padlock probes (PLPs) and RCA approaches, utilizing bridged probes for sequence-by-hybridization chemistry and highly multiplexed detection, in contrast to the sequence-by-ligation scheme employed in ISS. This modification facilitates higher-throughput spatial targeted sequencing of larger tissue areas. In contrast, IISS (Tang et al., 2023c) alters the probe binding scheme through combinatorial probe anchoring chemistry (icPAL) and adopts a two-base barcode strategy, enhancing sequencing quality, signal intensity, and specificity. Spatially Resolved and signal-diluted Next-generation Targeted sequencing (SPRINTseq) (Chang et al., 2023b) employs synthetic sequencing by synthesis to expedite the reaction alongside a microfluidic flow cell and a novel padlock probe hybridization design. This significantly reduces the time needed for ISS, boosting detection throughput. Remarkably, SPRINTseq enables the detection of 142 million transcripts in 450,000 cells in <2 days, which showcases a substantial advancement in technology. The enduring principles of ISS underscore the remarkable progress achieved in the development of these innovative technologies.

The integration of SISS with hydrogel technology is emerging as a notable trend. The incorporation of hydrogel in tissue clearing serves to diminish optical interferences from the tissue background. STARmap (Wang et al., 2018b) exemplified this trend, excelling in achieving high-resolution ISS of three-dimensional tissues. In this methodology, STARmap/STARmap PLUS initially uses pairs of DNA probes to bind the target mRNA within the tissues, resulting in amplified DNA nanoballs (DNBs) anchored within the three-dimensional hydrogel space. Subsequently, tissue clearing is executed by eliminating proteins and lipids, paving the way for *in situ* sequencing, and enabling the detection of over 1,000 genes within the cleared tissues (Shi et al., 2023; Zeng et al., 2023b). ExSeq (Alon et al., 2021) uses the expandability of hydrogel to secure tissue nucleic acids within the hydrogel. Similar to FISSEQ, ExSeq is capable of nontargeted sequencing through the amplification of RNA into cDNA nanoballs. This method showcases the prospect of achieving subcellular resolution in diverse tissues, including mouse hippocampal neurons, visual cortex, and human breast cancer.

Building upon the imaging methodologies of SISH and SISS, enhancements have been implemented in hybridization principles, combinatorial labeling, error correction schemes, and various other aspects. When coupled with advanced microscopy and hydrogel technology, these techniques have progressively elevated throughput and sensitivity for spatial *in situ* detection across the transcriptome. The authors even demonstrated a significant potential in true 3D spatial omics. However, challenges continue to persist, such as in the experimental intricacies of multi-hybridization imaging schemes, optical congestion arising from high-density nucleic acids, and the inherent low-throughput nature of current sequencing schemes. These factors collectively hinder fundamental parameter improvements, making it challenging to harness the full potential of second-generation sequencing technologies.

SISM technology

In the period following the initiation of the human genome project, spatial omics technologies have evolved in tandem with second-generation sequencing techniques. The primary detection method involves conducting high-throughput sequencing on the entire composite library, as opposed to capturing precise *in situ* images of individual nucleic acid sequences. This strategy mitigates the constraints associated with multiple hybridizations and optical congestion to provide the benefits of batch processing and high-throughput detection. However, it is apparent that delineating distinct spatial dimensions within the mixed sequencing library has emerged as a new bottleneck and, hence, a constraint in the process.

In SISM, the imparting and restoration of spatial information for tissues and cells before and after sequencing constitute fundamental elements of the technique. This is achieved through physical partitioning or chemically labeling selected regions of interest (ROIs). Subsequently, the isolated libraries undergo either bulk sequencing or single-cell sequencing, which allows the straightforward detection of spatial heterogeneity between ROIs. LCM (Emmert-Buck et al., 1996), a widely employed microdissection technology, employs laser irradiation on the selected regions by fusing the artificial membrane with the underlying cells. This process involves region segmentation, recovery into a container, and subsequent sequencing of different libraries. Tomo-seq (Junker et al., 2014) further enhances spatial resolution by slicing the entire zebrafish embryo into 50–100 sections, extracting RNA from each section for individual sequencing and thereby achieving a relatively coarse-resolution 3D ST detection. Building on the principles of LCM and Tomo-seq, higher-resolution spatial omics technologies have been developed, such as Geo-seq (Peng et al., 2016), immuno-LCM-RNAseq (Zhang et al., 2022b)—which incorporates immunofluorescence-guided laser cutting—STRP-seq (Schede et al., 2021)—which involves the cutting of adjacent tissue sections at different angles—and GaST-seq (Giolai et al., 2019)—which is specifically designed for plant research.

Another application of ROI involves labeling with photoactive groups within the tissues, followed by biochemical reactions to assign spatial identifications (IDs) for sequencing detection. Notable techniques in this category include TIVA-tag (Lovatt et al., 2014), PIC (Honda et al., 2021), and ZipSeq (Hu et al., 2020). Among these, DSP (Merritt et al., 2020) technology has achieved commercial success.

In the context of SISM, spatial positioning information is applied to tissues, and high-throughput measurements are conducted on the entire library. However, the constraints encountered in physical sectioning operations considerably limit enhancement in spatial resolution. This limitation poses a challenge when conducting large-scale studies at a single-cell or subcellular level.

SISB technology

Microdissection technology involves the segmenting of various regions of tissues. The potential to differentiate each mRNA within a cell and perform high-throughput sequencing presents an opportunity to improve spatial resolution from the cellular level to the single-cell or even the subcellular level. Spatial tagging technologies have played a crucial role in supporting this concept. In 2016, Lundeberg's group (Ståhl et al., 2016) introduced ST technology, a groundbreaking approach that incorporated high-density probes with spatial barcodes and polyT sequences on a glass surface for preparing a capture chip. Tissue slices were placed on the chip, and treatments were applied to enable penetration and perforation of cell membrane structures,

facilitating the targeted capture of the 3′ polyA structures of mRNA within cells by using the chip's probes. *In situ* RT was then performed to generate cDNA, which was subsequently dissociated from the glass slide, collected into a container, amplified by PCR, and subjected to high-throughput sequencing. By aligning barcode sequences uniquely associated with mRNA in the sequencing results with the known barcode positions during implantation, the spatial locations of mRNA can be calculated, thereby enabling spatially resolved transcriptome sequencing. The spot spacing of ST technology, which refers to the distance between the centers of spots, is 200 μm, which gives a spatial resolution of 200 μm. Following its acquisition by 10x Genomics and subsequent technical optimization, ST technology was rebranded as "Visium," offering an improved resolution of 100 μm. Although Visium does not achieve an absolute single-cell resolution, considering the presence of multiple cells within each spot, it still offers significant technical advantages when compared with microdissection techniques. In addition, ST technology is the first to successfully assign spatial IDs to mRNA using barcodes, thereby introducing an original concept for subsequent spatial omics technologies and laying the foundation for the current popular spatial omics methodologies (Cheng et al., 2023a).

The essence of barcode-based spatial omics technologies lies in the method of assigning distinct barcodes to individual mRNA molecules. In ST technology, different barcode probe clusters are physically implanted at specific distances from each other on a glass slide to ensure minimal interference. An alternative approach involves carrying different barcode probes on separable carrier units, which are then assembled to form a chip with the carrier as the minimum spatial resolution. For example, Slide-seq (Rodriques et al., 2019) and Slide-seqV2 (Stickels et al., 2021) have implemented this concept by using magnetic beads, which are widely being used presently, whose spatial resolution is determined by the diameter of the magnetic beads, which is typically 10 μm. In contrast, high-definition STs (HDST) (Vickovic et al., 2019) use smaller magnetic beads that are placed in microwells to discriminate spatial barcodes so as to achieve a finer spatial resolution of 3 μm. However, due to limited detection signals, only the regions formed by multiple magnetic beads can be used as the minimum analysis unit without attaining single-cell resolution. Microfluidic devices have demonstrated notable advantages in the distribution of spatial barcodes. In the case of DBiT-seq (Zhang et al., 2020b), a microfluidic device is employed, permitting separate liquid addition in the x and y directions. This capability facilitates the addition of x-axis and y-axis coordinates to mRNA within tissue slices to accomplish spatial ID transposition. Despite these advancements, the spatial resolution is constrained by the limitations of the microfluidic channel size, and it has not yet attained the level of single-cell resolution.

The evolving trend in barcode technology is focused on reading and editing nucleic acid sequences with higher spatial resolution accuracy. This process involves equipping them with unique sequence spatial barcodes for planar capture and RNA sequencing. Unlike the development of barcode assignment schemes from the ground up in previous technologies, high-throughput sequencing technologies have been naturally progressing in this direction over several decades. High-throughput sequencing schemes entail biochemically attaching numerous cDNA amplification sequences with specific physical spacing onto the surface of a sequencing chip (Reuter et al., 2015). Subsequently, individual high-resolution base sequencing is performed on each amplification sequence. Over continuous iterations, the minimum physical spacing of nucleic acid sequences on Illumina and BGI sequencing chips reaches the nanometer level, with the number of nucleic acid clusters on a single chip reaching hundreds of millions. The parameters of these chips, whether physical or optical, have pushed the boundaries of the field to their maximum. Considering this point, the idea of replacing the nucleic acid fragments on the sequencing chip with capture probes consisting of different combinations of barcode and polyT sequences opens the possibility of capturing and sequencing tissue mRNA. Theoretically, this approach can achieve nanometer-level/subcellular-level spatial resolution omics technologies. The imminent challenge lies in reaching the technological limits of the field in the short term. In the realm of technological progress, Stereo-seq based on BGI sequencing chips and seqScope (Cho et al., 2021) based on Illumina sequencing chips have gradually emerged around 2022, which aligns with the trajectory of technological development and reshaping of spatial omics technologies. The distinction lies in the composition of the Stereo-seq chip's surface, which comprises matrix-arranged spot points that leverage BGI sequencing's patented DNBs to carry spatial barcode probes, which achieve a consistent resolution of 500 nm. Registering and segmenting single-stranded DNA (ssDNA)-stained tissue images has accomplished genuine single-cell resolution in spatial sequencing. In addition, owing to the expensive dimensions of BGI sequencing chips' reaction chip, the Stereo-seq spatial chip can span larger centimeter-level sizes, reaching up to 13 cm×13 cm, which facilitates spatial omics detection across entire sections of monkey brains. Both two single-cell resolution SISB methods involve comparisons of diverse barcodes within mixed libraries to ascertain mRNA positions, necessitating two rounds of sequencing for each chip and subsequently escalating the experimental costs. Piexl-seq (Fu et al., 2022) introduces an innovative solution wherein barcode transfer on the capture chip is accomplished by following a stamp-like method. This novel approach enables the acquisition of several dozen original spatial omics chips in a single sequencing run, thereby ultimately reducing the overall operations cost.

Sequencing-based spatial omics technologies such as SISM and SISB offer notable technical merits amid the evolution of high-throughput sequencing technology. For instance, they have excelled in achieving unbiased whole transcriptome detection, and their enhanced throughput and user-friendly operation have propelled their widespread adoption in the spatial omics technology market. However, as previously mentioned, sequencing chip-based spatial omics technologies have, to an extent, encountered constraints in the field, which have posed challenges for significant short-term enhancements to key technical parameters. Naturally, the development of technology hinges on the imperatives of scientific research. The existing strengths of STs suffice to cater to the demands of a majority of research endeavors, thereby exhibiting robust application capabilities across domains such as neuroscience, developmental biology, pathology, and plant biology (Tian et al., 2023a). In the immediate future, the discernible technological trajectory involves transitioning from single STs to multi-omics, encompassing genomics, epigenomics, proteomics, and much more. The pivotal focus is thus on achieving simultaneous detection of multiple omics within the same tissue section.

Spatial multi-omics

Spatial multi-omics research spans genomics, epigenomics, transcriptomics, proteomics, and metabolomics, as well as investigations into transcriptional and translational states (Vandereyken et al., 2023). Presently, spatial omics technologies predominantly center on the transcriptome dimension for deducing the expression patterns and regulatory mechanisms in biological phenomena through mRNA abundance analysis. In the context of central dogma, RNA expression is not only the outcome of upstream gene transcription and epigenetic regulation but also the catalyst for downstream protein translation. The central dogma has evolved into the "spatiotemporal central dogma," which allows simultaneous characterization of the biological phenomena at multiple omics levels, encompassing temporal and spatial dimensions, allowing for a more comprehensive understanding of cellular fate regulation mechanisms. This evolution significantly enhances our microscopic understanding of macroscopic life phenomena. Pioneering spatial omics technologies are forging a technological ecosystem geared toward concurrent multi-omics detection, with a primary emphasis on the transcriptome dimension. These technologies exhibit promising prospects, particularly in domains such as mouse brain research and disease studies (Figure 6).

Spatial genomics

Despite the fact that genome research boasts a lengthy history, the technology available for delineating the spatial dimension of chromatin states or DNA information within tissues remains nascent. Hi-M (Cardozo Gizzi et al., 2019), which is grounded in multiplexed imaging, represents a method that can concurrently discern chromosomal organization and transcriptional information within individual nuclei. However, due to constraints in imaging technology and hybridization capabilities, Hi-M encounters limitations in detecting tissues and DNA loci. Through the optimization of multiple hybridization schemes, SISH has markedly enhanced the detection of cell numbers and genomic loci. DNA-MERFISH technology (Sibai et al., 2020), which employs sequential hybridization, provides direct visualization of chromatin structural domains and compartments, unveils interchromosomal interactions, and facilitates simultaneous imaging of over 1,000 gene loci and nascent transcripts. DNA seqFISH+ (Takei et al., 2021) allows imaging of 3,660 chromosome loci in mouse embryonic stem cells. Similar to transcriptomics technology, *in situ* imaging-based techniques, such as SISH and SISS, primarily characterize targeted loci, while sequencing-based methods capture unbiased whole-tissue DNA information. Leveraging the foundation of Slide-seq, Chen's group (Zhao et al., 2022) pioneered Slide-DNA-seq by using barcode and capture sequence modifications on densely packed beads to achieve an unbiased capture of DNA within the tissue surface. Following sequencing, spatial genomic information has become accessible and applicable in the study of human cancer samples, uncovering clonally specific genetic alterations across diverse genomes. Although slide-DNA-seq can be integrated with adjacent sections for multi-omics characterization, it presently falls short of achieving simultaneous multi-omics sequencing on the same tissue section. In summary, spatial genomics technology offers distinctive insights into developmental biology, human diseases, and various research domains. However, the availability of widely applicable spatial genomics technologies remains limited. Therefore, future technological advancements should focus on developing methods that enable the simultaneous detection of the genome and transcriptome.

Spatial epigenomics

Spatial epigenomics introduces a novel perspective for analyzing the spatial dimension of cellular fate regulation mechanisms, and its technological development has rapidly progressed in tandem with transcriptomics technology. Zhuang's group (Lu et al., 2022) employed the MERFISH technique based on multiplexed SISS to capture specific epigenetic modifications on chromatin *in situ*. The authors labeled the DNA near the modification site with a T7 promoter, transcribed the labeled DNA fragments *in situ* to generate RNA, and detected and characterized the target RNA using MERFISH' multiplexed *in situ* fluorescence hybridization technology, thereby enabling spatial epigenomic detection, called epigenomic MERFISH. Sequencing-based spatial omics technologies have exhibited enhanced compatibility in omics studies. In spatialATAC, Ståhl's group (Llorens-Bobadilla et al., 2023) integrated Tn5 transposase into the tissue clearing step of ST-based RNA capture technology, which has allowed the labeling of open chromatin and the restoration of spatial position through spatial barcoding, thereby achieving spatial ATAC detection. Similarly, by leveraging microfluidic channels, Fan's group proposed spatial-ATAC-seq (Deng et al., 2022b) and Spatial CUT&Tag (Deng et al., 2022a) technology. The Spatial CUT&Tag involves the introduction of antibodies targeting modified proteins into tissue sections, followed by enhancement of pA-Tn5 transposase binding with secondary antibodies. Next, adapters are inserted into the antibody recognition sites of DNA with histone marks, and the sequences with spatial barcodes are *in situ* connected to the adapters through bilateral channels. Subsequent sequencing then generates a genome-wide map for the spatial analysis of histone modifications.

Microfluidic indexing-based spatial assay for transposase-accessible chromatin and RNA-sequencing (MISAR-seq) adopted similar strategies for spatially resolved joint profiling of chromatin accessibility and gene expression (Jiang et al., 2023a). In conclusion, epigenetic regulation stands as a crucial facet for comprehending the mechanisms underlying microscale cellular fate regulation and macroscopic biological events such as development and disease. Spatial epigenetic regulation technology has gradually emerged and garnered attention. However, the simultaneous detection of the epigenome and transcriptome encounters technical challenges, hindering the rapid development of simultaneous multi-omics epigenetic technologies. This area thus remains a significant avenue for future advancements in multi-omics technology.

Spatial proteomics

In the realm of spatial omics technology, the essential prerequisite for accurately pinpointing biomolecules in space lies in their specific capture *in situ* or an unbiased capture, followed by reconstruction using the carried positional IDs. Guided by these principles, proteins can be detected *in situ* through the specific binding of antigen-antibody interactions, which can then facilitate targeted spatial detection. Alternatively, proteins can be quantified and sequenced by linking oligonucleotide barcodes with antibodies, thus leveraging the CITE-seq principle. Owing to the stability of proteins and the compatibility of experimental

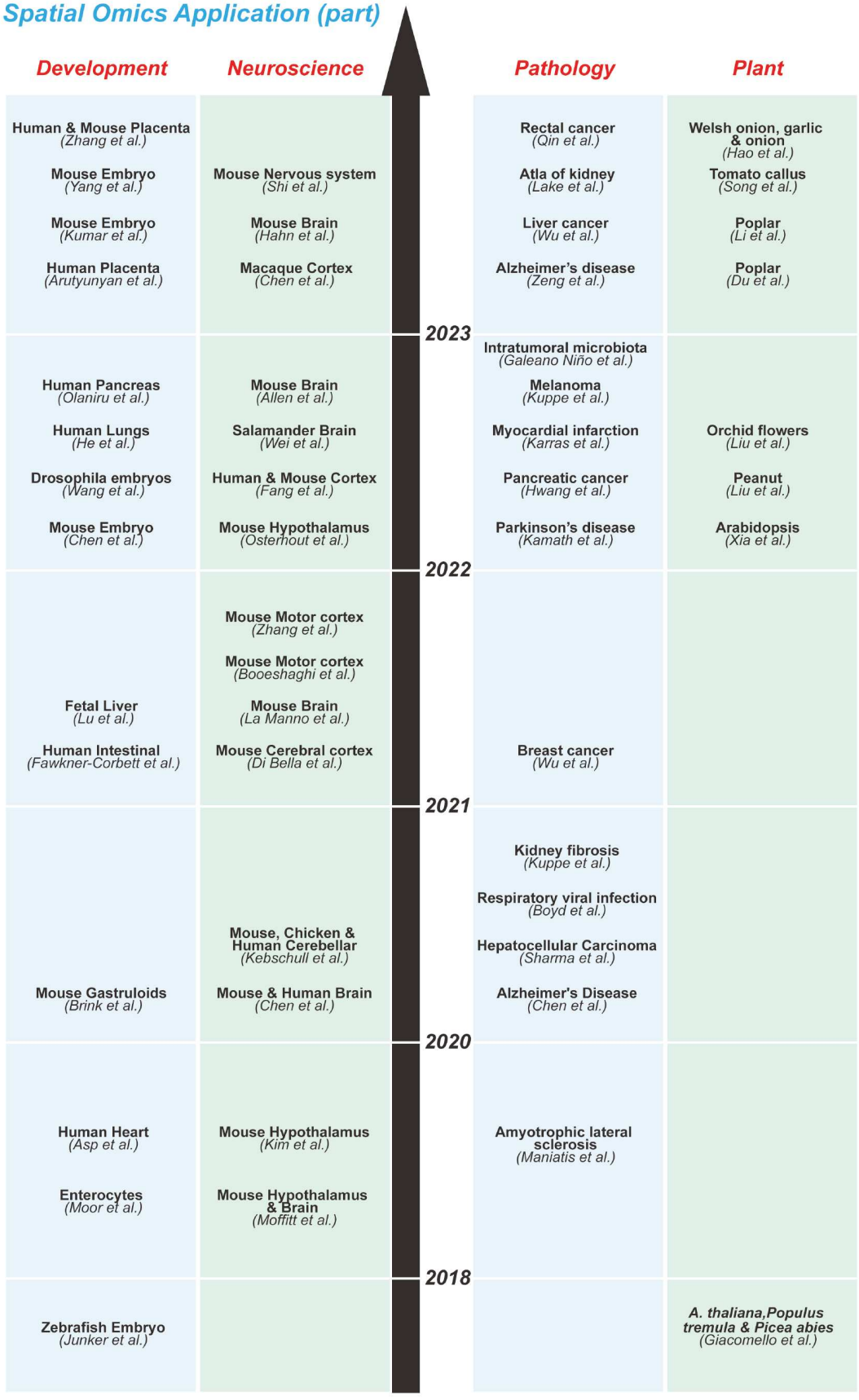

Figure 6. Timeline of the application of spatial omics methods in the fields of development, neuroscience, pathology, and plants.

https://doi.org/10.1007/s11427-024-2770-x

procedures, achieving simultaneous detection of proteins and transcriptomes on the same tissue section is relatively straightforward. This capability has enabled true spatial dual-omics technology encompassing proteomics and transcriptomics and provided crucial technical support for the rapid development and widespread application of spatial proteomics technology.

The method for detecting targeted proteins *in situ*, following specific binding with antibodies, relies heavily on the execution of the final imaging step. Therefore, this approach facilitates the direct spatial dimension detection of targeted proteins in imaging-based ST technologies. This method is frequently employed in conjunction with the existing RNA detection methods to integrate stable protein detection schemes. In imaging mass cytometry (Schulz et al., 2018), RNA is amplified *in situ* by using the RNAScope technique, and metal chelate antibodies are subsequently introduced to bind to the targeted proteins. Following staining, *in situ* imaging is conducted to achieve spatial dual-omics detection. CAD-HCR (Liu et al., 2022c) technology amplifies RNA signals *in situ* through the HCR, and fluorescence imaging is performed after achieving a specific binding between antibodies and proteins. STARmap PLUS technology, an extension of STARmap's RNA *in situ* detection, allows imaging of proteins after the *in-situ* binding of primary and secondary antibodies, thereby enabling spatial protein detection. This approach has been instrumental in the determination of spatial-specific expression of amyloid-beta and tau protein plaques, which are critical in the progression of Alzheimer's disease (AD) (Zeng et al., 2023b).

This approach involves pre-binding antibodies to nucleotides carrying unique barcode sequences and polyA tails, followed by specific binding to antigens. This process allows the capture of antigens alongside mRNA sequences by the polyT capture probes. High-throughput sequencing of the sequences carrying barcodes enables the reconstruction of their spatial positions. This CITE-seq-based strategy proves more suitable for sequencing-based spatial omics technology. Stereo-CITE-seq technology (Liao et al., 2023), built upon Stereo-seq, employs antibodies with antibody-derived tags (ADTs) sequences for targeted proteins. Capturing the ADTs allows synchronous quantification of transcriptomes and spatial protein detection. Other techniques employing the ADT strategy include DBiT-seq based on microfluidic channels (Zhang et al., 2020b), SM-Omics and SPOTS based on ST technology, and Spatial-CITE-seq based on Slide-seq (Liu et al., 2023e). In addition, an atomic effort has been made to profile spatially resolved translatomics (e.g., RIBOmap (Zeng et al., 2023a)), metabolomics (e.g., SEAM (Yuan et al., 2021), and scSpaMet (Hu et al., 2023)) at the single-cell level. Overall, sequencing-based spatial omics technologies exhibit stronger compatibility for capturing proteins and transcriptomes. Synchronous dual-omics analysis thus holds significant potential in pathological research.

In broader terms, spatial multi-omics technology is still in its early phases, and achieving simultaneous spatial detection of multiple omics on a single tissue section is a formidable challenge. The current workaround involves segregating the detection of different omics on adjacent tissue sections, followed by integration of the data for analysis. In the future, the pursuit of true spatial multi-omics on a single tissue section is bound to emerge as a crucial direction for technological advancement.

Comprehensive comparison of spatial omics technologies

Various spatial omics techniques are rapidly evolving, necessitating a thorough evaluation of their respective strengths and limitations for targeted applications in specific contexts. Comparison across categories of these techniques generally encompasses spatial resolution, detection efficiency, signal diffusion, and tissue area.

(1) Spatial resolution. Imaging-based spatial omics methodologies, as exemplified by MERFISH (Moffitt et al., 2016) and STARmap (Wang et al., 2018b), facilitate the detection of individual molecules within cells at the subcellular resolution level. This enables single-cell resolution through spatial segmentation of single cells by using cell staining techniques. Differentially, sequencing-based spatial omics approaches generally involve the capture of nucleic acid molecules via probes, with subsequent spatial position reconstruction achieved through barcoding. Consequently, the spatial resolution of the latter is contingent upon the probe density on the capture substrates. Presently, the majority of sequencing-based spatial omics methodologies function at a multicellular resolution, with each capture unit encompassing multiple cells, as exemplified by Visium (Suo et al., 2022) and DBiT-seq (Liu et al., 2020b). Nonetheless, certain techniques exhibit probe densities that can attain subcellular resolution and spatial single-cell precision can be achieved via cell image segmentation, as demonstrated in Stereo-seq (Chen et al., 2022a), Seq-Scope (Cho et al., 2021), and Pixel-seq (Fu et al., 2022).

(2) Detection efficiency. The evaluation criteria for detection efficiency currently vary between imaging-based and sequencing-based spatial omics techniques. smFISH is frequently employed to compare the detection efficiency owing to its high sensitivity, wherein the average number of transcripts detected per gene in each cell or region is assessed. Imaging-based spatial omics techniques generally exhibit a higher detection efficiency; for instance, MERFISH achieves an efficiency of approximately 95% (Moffitt and Zhuang, 2016). Conversely, sequencing-based spatial omics techniques often exhibit reduced detection efficiency, attributable to the type of capture and library sequencing methodologies used. For example, STs detect only 6.9% of the UMIs per unit area when compared with smFISH (Ståhl et al., 2016). Imaging-based spatial omics techniques typically target specific nucleic acid sequences, with optical constraints encountered during *in situ* imaging or sequencing, which limits the number of targeted nucleic acids that can be detected in a single experiment. Consequently, optimizing these optical limitations remains a primary focus toward enhancing the efficacy of imaging-based spatial omics techniques. Presently, both MERFISH (Xia et al., 2019) and seqFISH+ (Eng et al., 2019) can detect 10,000 genes. Theoretically, sequencing-based spatial omics techniques have the potential to capture the entirety of mRNA within cells by using polyT probes. Therefore, the primary objective in evaluating these techniques is to calculate and enhance the total number of reads captured per unit area or in a single cell. For instance, in an adult mouse coronal hemibrain section, Stereo-seq can detect 1,910 UMIs and 792 genes per cell.

(3) Signal diffusion. The spatial detection accuracy of RNA significantly influences the precision of detection outcomes. This parameter is extensively examined in sequencing-based spatial omics methodologies. A contributing factor to this scrutiny is the phenomenon wherein mRNA migrates a short distance post-

enzymatic reaction -permeabilization and is subsequently captured. The extent of this diffusion is contingent upon tissue type and the duration of the enzymatic reaction. For instance, in the comparative analysis of the *Slc17a7* marker gene within the olfactory bulb, both Slide-seq V1.5 and PIXEL-seq exhibit relatively robust performance with respect to diffusion. Conversely, when evaluating the Pmel marker gene in ocular tissues, Stereo-seq demonstrates a superior performance (You et al., 2024).

(4) Tissue area. Owing to the inherent limitations in detection methodologies, imaging-based spatial omics techniques often face challenges in executing large-scale detection within a single experiment. These methods are typically constrained in that they can detect a range of only a few hundred to a few thousand genes. Conversely, sequencing-based spatial omics techniques exhibit greater flexibility. The effective capture area of these techniques is contingent upon the probe arrangement on the capture substrate. For instance, Visium encompasses an area of 6.5 mm×6.5 mm (Denisenko et al., 2024). PIXEL-seq features an effective capture area of 75 mm×25 mm, and Stereo-seq can generate spatial omics data for monkey brain tissues with dimensions of 5 cm×3 cm (Chen et al., 2023a). The scalability and flexibility of these methodologies render them particularly suitable in the investigation of complex tissues and large-scale spatial gene expression patterns.

Analysis and application

With the emergence and widespread adaptation of spatial omics technology, the primary emphasis of spatial omics technology in scientific applications has gradually shifted from technical implementation to the analysis and processing of output data (Palla et al., 2022a). However, without an efficient, direct, and effective analysis approach for handling large volumes of generated data, it is difficult to distill core detection indicators from the massive datasets for guiding biological research. This hindrance can ultimately lead to data wastage and impede progress in the field of spatial omics research. Therefore, the data analysis stage has evolved into a limiting factor in spatial omics technology, encompassing aspects such as data segmentation, database integration, the development of multi-omics analysis methods, 3D reconstruction algorithms, and much more. The introduction and continuous improvement of efficient and practical analysis tools are poised to bring greater convenience and novel research methodologies to scientific research.

Spatial omics technology has ushered in new insights and research perspectives across diverse fields, spanning developmental biology, neuroscience, pathology, and botany. The foundational step in understanding biology often involves identifying the cell types and delineating cellular or genetic functional characteristics (Nolan et al., 2023). By providing a spatial dimension perspective, spatial omics offer a means to unearth novel cell types (Zeng, 2022). The integration of spatial features facilitates the establishment of connections between function and structure. During the developmental process, continuous spatial readability can significantly enhance the crosstalk analysis between differentiated cells and their regulation of cell differentiation by the surrounding cellular factors (Arutyunyan et al., 2023; He et al., 2022).

This approach facilitates the identification of cell subtypes that operate within distinct spatial locales. For instance, in the development of axolotl brain tissues, Wei et al. (2022) identified a particular cell type that can generate three distinct subpopulations, each occupying different spatial positions. Similarly, Zhang et al. (2023a) characterized three clusters of cells with subtly varying transcriptomes, which were mapped to specific regions in the human embryonic limb. The spatial relationships offer a critical perspective in cancer research, particularly in the elucidation of the advantageous interactions between cancer cells and surrounding immune cells in cancer cell migration (Ji et al., 2020; Karras et al., 2022). Notably, Erickson et al. (2022) identified distinct clonal patterns within tumors and adjacent benign tissues by analyzing spatial CNVs. This approach provides a framework for understanding the spatially continuous transition from benign to malignant tissues. In the domain of neuroscience, spatial omics methodologies offer significant advantages for uncovering the distinctive architecture and functions of brain tissues. For example, these techniques have delineated a spatial atlas of the mouse central nervous system at molecular resolution (Shi et al., 2023), delineated the cell type specificity within the primary motor cortex of the mouse (Zhang et al., 2021c), and highlighted the critical role of the spatial connectivity network of neurons in determining the specific functions of brain regions (Chen et al., 2023a; Fang et al., 2022a). In the realm of plant biology, the presence of cell walls poses challenges for conventional single-cell sequencing techniques. Spatial omics technology and leveraging of tissue sectioning and tissue clearing methods triumphs over the hurdles posed by cell walls (Li et al., 2023b; Xia et al., 2022). This approach enables spatial *in situ* signal capture, presenting breakthrough methodology for botanical research. The transcriptional profiles of upper and lower epidermal cells in *Arabidopsis* leaves display nuanced differences. For instance, Xia et al. (2022) effectively identified and differentiated these spatially distinct yet transcriptionally similar cell subtypes by analyzing the differential expression characteristics. In addition, Du et al. (2023) postulated the presence of two types of meristematic-like cell pools within secondary vascular tissues and corroborated the hypothesis through multiple methodological approaches. In summary, the emergence and development of spatial omics technology contributes to technological convenience and accessibility to various research fields, thereby fostering advancements in our current understanding of complex biological systems (Table S4).

Chapter 5. Computational challenges

With rapid advancements in scRNA-seq technologies, a flood of computational tools has been developed to handle the massive scRNA-seq data. Generally, these tools are developed for data preprocessing, batch-effect correction, data normalization, dimension reduction, feature selection, cell clustering, cell type annotation, graphic presentation, and other purposes. For example, Cell Ranger was developed for data preprocessing; Harmony (Lake et al., 2019) was introduced for batch-effect correction; and Seurat (Hao et al., 2021) was developed for normalization, dimension reduction, feature selection, and cell clustering. Although manual annotation with a predefined set of marker genes is the gold standard method for cell type annotation, a set of tools such as SingleR (Aran et al., 2019), CellAssign (Zhang et al., 2019a), and Garnett (Pliner et al., 2019) have been developed to annotate cells automatically to some extent. In addition, some useful tools have

been developed to infer cell trajectory (Qiu et al., 2017; Street et al., 2018), cell-cell communication (Browaeys et al., 2020; Cabello-Aguilar et al., 2020; Efremova et al., 2020; Jin et al., 2021), metabolic state (Wagner et al., 2021), cell-type-specific regulon (Ma et al., 2020a), and somatic mutation (Vu et al., 2019) from scRNA-seq data.

Computational analyses based on the scDNA-seq data mainly encompass mutation detection, mutation clustering, genotype inference, and phylogeny reconstruction. To address this concern, plenty of tools have been developed that aim to: (i) detect SNVs or CNVs, such as SCAN-SNV (Luquette et al., 2019), SeCNV (Ruohan et al., 2022), and SCOPE (Wang and Jiang, 2022; Wang et al., 2020a); (ii) perform mutation clustering and genotype inference, such as AMC (Yu and Du, 2022), BnpC (Borgsmüller et al., 2020), and SCClone (Yu et al., 2022); (iii) phylogeny reconstruction based on either or both of SNVs and CNVs such as ConDoR (Sashittal et al., 2023), Phertilizer (Weber et al., 2023), MEDALT (Wang et al., 2021a), BiTSC (Chen et al., 2022d), and CellPhy (Kozlov et al., 2022). In addition, some useful tools have been developed to detect doublets (e.g., doubletD (Weber et al., 2021)) or quantify intra-tumor heterogeneity (e.g., Dhaka (Rashid et al., 2021) and rcCAE (Yu et al., 2023c)) from scDNA-seq data.

Considering the high dimensionality and sparsity of single-cell epigenomic data, dimension reduction is a critical step. To address this concern, numerous analytical tools have been developed. A popular method is latent semantic indexing, which has been implemented in many tools, such as Signac (Stuart et al., 2021) and ArchR (Granja et al., 2021). In this method, the term frequency-inverse document frequency (TF-IDF) and singular value decomposition were used to transform the sparse count matrix. In addition, methods including spectral embedding (e.g., SnapATAC (Fang et al., 2021)), topic modeling (e.g., cisTopic (Bravo González-Blas et al., 2019)), and variational autoencoders (AEs) (e.g., SCALE (Xiong et al., 2019)) were introduced to reduce the dimension by transforming the sparse count matrix into a cell-cell similarity matrix, thereby defining a set of "topics" for similar features or encoding features via neural network. Owing to the wide application of single-cell chromatin accessibility assays, most of these tools were developed for this specific modality, with the possibility to be adapted and applied to other epigenomic data. Another critical step in the analysis of single-cell epigenomic data is to identify and characterize cell type-specific regulatory elements and their targeted genes. For scATAC-seq data, tools such as Signac (Stuart et al., 2021), ArchR (Granja et al., 2021), and SnapATAC (Fang et al., 2021) provide the functions to identify cluster-specific peaks and peak-to-gene links. For scDNAm-seq data yielded by snmC-seq (Luo et al., 2017), sn-m3C-seq or snmCAT-seq (Luo et al., 2022) and ALLCools (Liu et al., 2021) enables the identification of differentially methylated region (DMR) and DMR-to-gene links. For scHi-C data, SnapHiC (Yu et al., 2021b) has been developed to detect chromatin loops and predict target genes based on physical interactions.

With the accumulation of multi-layered modalities, the efficient integration of multi-omics datasets poses a huge challenge for computational biologists. To address this concern, a variety of algorithms were developed for the following purposes: (i) removing batch effects from multiple datasets, such as MNN (Haghverdi et al., 2018), Scarnorama (Hie et al., 2019), and Conos (Barkas et al., 2019); (ii) extract single-cell level latent variables for downstream analysis (e.g., clustering), such as VEGA (Seninge et al., 2021), scDHA (Tran et al., 2021), and scMM (Minoura et al., 2021); (iii) infer single-cell developmental progress, such as SPADE (Anchang et al., 2016) and DPT (Haghverdi et al., 2016); (iv) integrate multi-omics datasets, such as MARIO (Zhu et al., 2023), Seurat (Hao et al., 2021), LIGER (Liu et al., 2020a), GLUE (Cao and Gao, 2022), StabMap (Ghazanfar et al., 2024), Cobolt (Gong et al., 2021), and MultiVI (Ashuach et al., 2023). In addition, computational efforts were made to (i) introduce graph construction from single-cell spatial data, such as Squidpy (Palla et al., 2022b), Athena (Martinelli and Rapsomaniki, 2022), and SPEX; and (ii) detect metabolites from individual cells, such as SpaceM (Rappez et al., 2021).

Here, we have discussed five key computational challenges encountered in single-cell genomics.

From bulk tissues to single cells

Recent advancements in computational deconvolution methods have been proposed for estimating cell type abundances in tissue RNA-seq profiles (Figure 7) for identifying cell type-specific gene expression alterations and detecting rare cell populations. These methods are categorized based on their methodology, including linear models (Li et al., 2020; Monaco et al., 2019; Peng et al., 2019b; Shen-Orr et al., 2010; Wang et al., 2019b; Zhong et al., 2013), support vector regression (Newman et al., 2015; Newman et al., 2019), gene set enrichment approaches (Aran et al., 2017; Quon et al., 2013; Yoshihara et al., 2013), probabilistic models (Kang et al., 2019), and deep learning models (Chen et al., 2022c; de la Fuente et al., 2023; Menden et al., 2020). However, various benchmarking efforts have demonstrated that the performance of these deconvolution methods can be influenced by factors such as data volume, data transformation, and the composition of the reference matrix. For example, enrichment-based approaches and probabilistic approaches often impose limitations on the number of usable cell types (Im and Kim, 2023), while deep learning-based methods encounter challenges in comprehensively simulating crucial rare components (Zhang et al., 2023d). In addition, solid tissues frequently contain closely related cell types, and the gene expression correlation among these cell types introduces collinearity, thereby showcasing challenges in resolving their relative proportions in bulk data. The collinearity problem can be improved by selecting marker genes through support vector regression, as demonstrated in Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts (Newman et al., 2015) and BSEQ-sc (Baron et al., 2016) or by employing a tree-guided procedure that recursively focuses on closely related cell types, as exemplified in MuSiC (Wang et al., 2019b). Apart from reference-free deconvolution methods, such as DECODER (Peng et al., 2019b) and CDSeq (Kang et al., 2019), most methods rely on prior knowledge of cell types based on either cell type-specific gene expression signatures or scRNA-seq/snRNA-seq data obtained from the same tissues. When constructing a cell reference atlas, it is therefore important to integrate data from the same tissue samples, consider the effect of scRNA-seq and snRNA-seq, address data processing, standardize cell type annotation, and enhance cell type resolution (Maden et al., 2023). Overall, computational deconvolution can provide a convenient and efficient means for analyzing bulk data from a single-cell perspective.

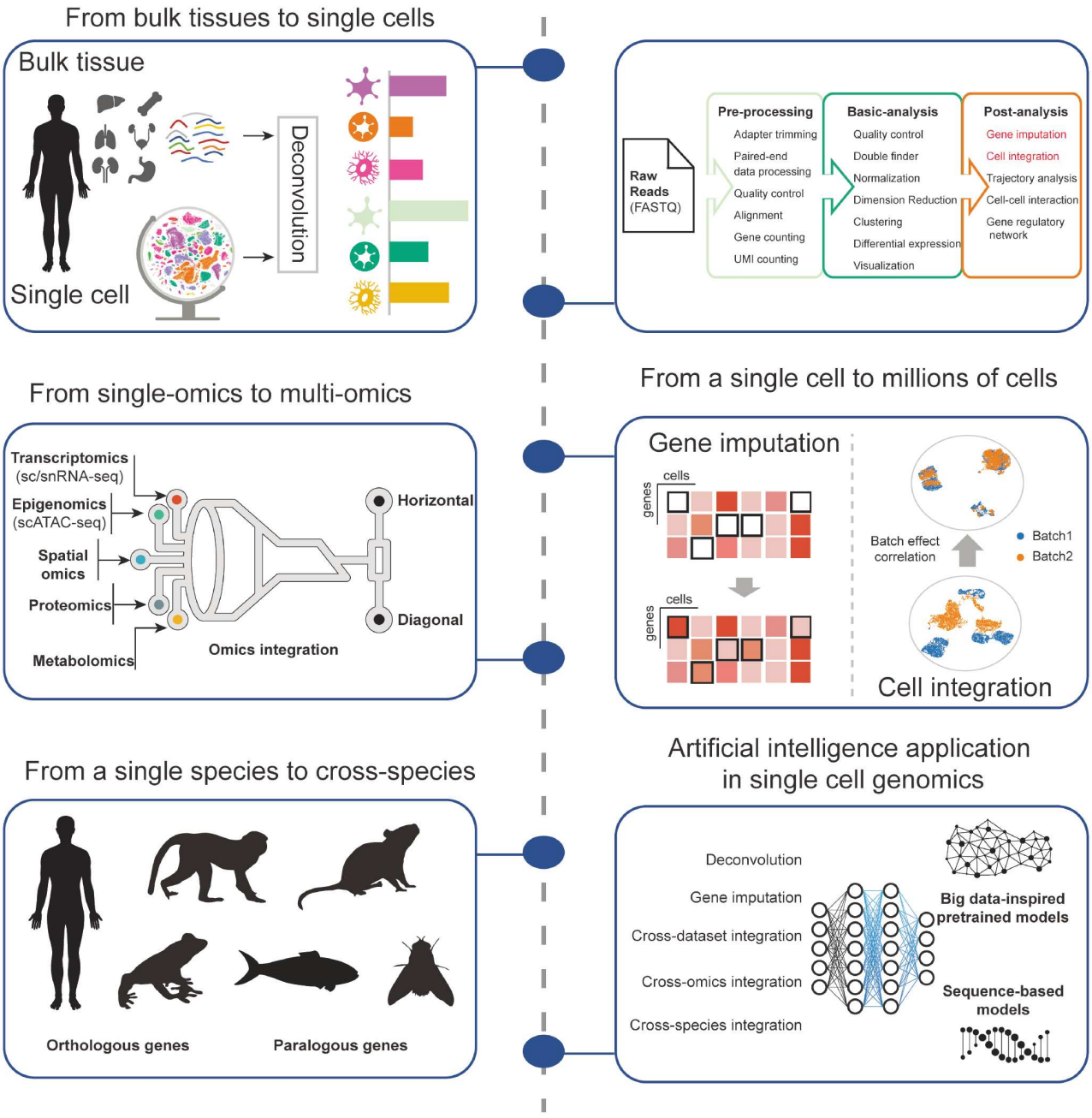

Figure 7. Key computational challenges presented in single-cell genomics analysis.

From a single cell to millions of cells

When compared with bulk RNA-seq analysis, scRNA-seq offers the advantage of exploring cellular processes with higher resolution. However, scRNA-seq is more vulnerable to various disturbances, including systematic differences derived from experimental procedures, dropout events, sequencing depth, and others. To address these challenges and conduct data preprocessing and noise removal, numerous bioinformatics algorithms are specifically dedicated to single-cell analysis. The general workflow of single-cell analysis typically involves quality control, alignment, feature quantification, normalization, dimension reduction, differential expression, and clustering analysis. As mentioned in some previous reports (Adil et al., 2021; Poirion et al., 2016), high levels of zero inflation pose a key challenge in analyzing scRNA-seq data. To address these dropout issues, cell imputation methods have been developed (Patruno et al., 2021). Model-based imputation methods utilize specific statistical models such as the Gamma-Gaussian mixture model (Li and Li, 2018), Poisson-Gamma mixture model (Huang et al., 2018), zero-inflated negative binomial model (Risso et al., 2018), and binomial model (Tang et al., 2020) to identify similar cells. Smooth-based imputation methods employ smoothing or diffusion techniques for determining gene expression values in cells with similar expression profiles, for example, DrImpute (Gong et al., 2018), Markov affinity-based graph imputation of cells

https://doi.org/10.1007/s11427-024-2770-x

(MAGIC) (van Dijk et al., 2018), and kNN-smoothing (Wagner et al., 2018). In addition, data reconstruction methods employ low-rank matrix-based approaches (Linderman et al., 2022; Mongia et al., 2019; Zhang and Zhang, 2018) or deep learning methods (Amodio et al., 2019; Arisdakessian et al., 2019; Deng et al., 2019; Eraslan et al., 2019; Lopez et al., 2018) to represent cells in a low-dimensional linear space or latent form for imputation. Recent studies have compared and evaluated systematic benchmarks of these scRNA-seq imputation methods (Cheng et al., 2023b; Hou et al., 2020). Although MAGIC and single-cell analysis via expression recovery (SAVER) (Huang et al., 2018) have outperformed other methods (Hou et al., 2020), they impute all zeros in the expression matrix, which increases the risk of overfitting for imputation (Zheng et al., 2023). Furthermore, the effectiveness of imputation varies across datasets (Cheng et al., 2023b). Therefore, it is essential to provide data-driven guidance for scRNA-seq imputation analysis based on the specific characteristics and challenges presented by each dataset.

Advancements in single-cell atlas mapping technology have facilitated the expansion of single-cell datasets in terms of both size and quantity (Angerer et al., 2017), which has helped establish common cell atlas consortium databases containing millions of single cells (Ye et al., 2024a). Consequently, single-cell genomics is powering big data science, thereby introducing new computational challenges and opportunities for data integration at the atlas level. To address the challenges of analyzing millions of cells, frameworks have been developed for single-cell data integration, such as the Python-based Scanpy (Wolf et al., 2018) and the R-based Seurat v5 (Hao et al., 2024b).

The major hurdle in achieving a common ontology of cell types for downstream analysis in single-cell atlas-level integration is the presence of batch effects (Eisenstein, 2020; Lähnemann et al., 2020) arising out of variations ranging from technical to biological factors (Luecken et al., 2022). Various batch effect correction (BEC) methods have been developed based on different models, ranging from early linear decomposition models to the more recent approaches such as mutual nearest neighbors (MNN)-based and deep learning models (Ryu et al., 2023). Bayes-based Combat (Johnson et al., 2007) and PCA-based Harmony (Korsunsky et al., 2019) are the popular linear decomposition models. A typical class of MNN-based BEC methods, such as MNNCorrect (Haghverdi et al., 2018), fastMNN (Haghverdi et al., 2018), and Scanorama (Hie et al., 2019), use MNN between batches as anchors. In addition, methods like scMerge (Lin et al., 2019) and sMNN (Yang et al., 2021) initially cluster each batch and then identify MNN between the clusters. Furthermore, MNN has been combined with non-negative matrix fractalization in methods such as Linked Inference of Genomic Experimental Relationships (LIGER) (Welch et al., 2018), K-nearest neighbors (KNN) in BBKNN (Polański et al., 2020), and deep learning models in deepMNN (Zou et al., 2021). In addition, deep learning-based BEC methods utilize a diverse model framework, with MMD-ResNet (Shekhar et al., 2016) employing a residual neural network (ResNet (He et al., 2016)); SAUCIE (Amodio et al., 2019) using a deep neural network with an AE workflow; scGen (Lotfollahi et al., 2019) implementing a deep neural network with variational AE (VAE); scVI (Lopez et al., 2018) combining VAE with a probabilistic approach. In addition, Seurat continuously updates its integration workflows. Seurat v2 uses canonical correlation analysis (CCA) for BEC, and a higher versions (Seurat v3 to v5) are used in "anchor-based" BEC methods, including anchor-based CCA and anchor-based reciprocal PCA (RPCA).

To provide an unbiased guide to the choice of the BEC method, benchmarking studies on integration methods have been conducted. In small-scale samples (comprising hundreds or thousands of cells), ComBat outperforms nonlinear methods with the quantitative metric k-nearest-neighbor batch-effect test (kBET), which evaluates batch-effect removal and dataset integration (Büttner et al., 2019). For medium-scale samples (hundreds of thousands of cells), Harmony, LIGER, and Seurat v3 are best recommended for batch integration (Tran et al., 2020) as they compare four benchmarks across 14 BCE methods (including kBET, local inverse Simpson's index (Korsunsky et al., 2019), average silhouette width (Rousseeuw, 1987), and adjusted rand index (Hubert and Arabie, 1985)). However, best practice recommendations vary for different sample scenarios, such as BBKNN and Harmony, exhibiting strong performance in mixed heterogeneous biodiversity samples (Chen et al., 2021b). In large-scale samples (millions of cells), deep learning-based BEC methods such as scVI and scGen perform well, especially in handling complex integration tasks (Luecken et al., 2022). Moreover, the runtime efficiency is a crucial consideration for large-scale data integration. Overall, benchmarking integrations such as scIB serve not only to aid researchers in selecting the available integration methods but also in assisting developers in building more efficient methods (Luecken et al., 2022).

Besides BEC methods, another crucial data integration issue involves the harmonization and automated labeling of cell types. For example, scANVI (Xu et al., 2021a) represents a semi-supervised method that extends the scVI model to enable automated annotation of cell types and states. CellTypist (Domínguez Conde et al., 2022) employs logistic regression models with stochastic gradient descent learning to identify cell types and provides built-in and pre-trained models for different tissues. As an upgraded version of CellTypist, CellHint (Xu et al., 2023a) focuses on cell type harmonization. By utilizing a predictive clustering tree, CellHint can efficiently evaluate cell-cell similarities and harmonize cell annotations across multiple datasets.

Integrating large-scale single-cell data into a common framework is therefore essential for constructing a unified cell census from a single cell to millions of cells. This comprehensive resource not only enhances our current understanding of human biology at the cellular level but also facilitates the exploration of diverse cellular functions and interactions in the human body.

From single-omics to multi-omics and spatial omics

Recent advancements in single-cell high-throughput technologies have enabled biologists to catalog various cell types across multiple omics layers, such as chromatin accessibility (scATAC-seq (Cusanovich et al., 2018)) and transcriptome (scRNA-seq), together providing a more comprehensive perspective to uncover the underlying mechanism for cell diversity as well as the changes occurring during disease development or progression. Unlike single-omics data, which provides information limited to a single molecular dimension, multi-omics data offers a multi-faceted view of cell identity by integrating information from multiple layers of cellular function. The integration creates a complex interplay of dependencies and relationships, thereby presenting unique challenges for analysis, such as data integra-

tion, data heterogeneity, dimensionality reduction, computational demands, and data interpretation.

Emerging single-cell protocols such as SHARE-seq (Kim et al., 2023b) and CITE-seq (Stoeckius et al., 2017) enable simultaneous profiling of multiple omics, offering a foundational basis for integrating multi-omics data at a single-cell resolution. To utilize and combine these datasets of multiple types, various novel computational methods that are both effective and precise have been proposed. Overall, these integration methods can be broadly classified into three types: horizontal, vertical, and diagonal data integration (Athaya et al., 2023), depending on cell pairing and anchor information. Vertical integration aims to integrate multimodal single-cell data within the same cells through the correlation of information from distinct modalities. The "weighted-nearest neighbor" (WNN) method (Hao et al., 2021) is an example of vertical integration that applies an unsupervised framework to learn cell-specific modality weights and construct a WNN graph for integrative analysis. WNN analysis has been leveraged to pair single-cell transcriptomes and ATAC-seq profiles, which benefits the multimodal definition of cell identity as well as the identification of putative regulatory programs. Despite the emergence of simultaneous assays, different omics are typically measured independently, which gives rise to unpaired data, requiring horizontal and diagonal data integration. Horizontal data integration relies on a set of shared features (such as genes) to link data from different modalities, including CCA (Butler et al., 2018) and LIGER (Welch et al., 2018). CCA has been explored to integrate scRNA-seq and scATAC-seq datasets by transferring chromatin accessibility to gene activity scores, while LIGER integrates gene expression and DNA methylation by assigning the methylated regions to the nearest gene. Furthermore, many multimodal deep learning (MDL) approaches adopt the diagonal data integration strategy. For each high-dimensional data type, MDL can generate a low-dimensional representation of single cells and convert it into a common latent space for integration by building on model architectures such as variational AEs (VAE), AE, convolutional neural network (CNN), and recurrent neural network. Some past studies integrated more than modality pairs, as exemplified by GLUE (Cao and Gao, 2022). As a deep learning framework combining AEs for each modality with a graph-based neural network, GLUE bridges the gaps among three different omics, including gene expression, chromatin accessibility, and DNA methylation, thereby enabling the effective triple-omics integration.

Another challenge in multimodal analysis is the need for integrating scRNA-seq data with ST data. The spatial information of cells plays a crucial role in determining their states and functions, which enables the reconstruction of transcriptional panoramas for tissue architecture. Although ST datasets contain transcriptome profiles along with spatial localization, they often lack high resolution for individual spots. On the other hand, scRNA-seq data retains individual cell heterogeneity but eliminates the spatial information due to cell isolation procedures. Therefore, there is an urgent need for new integrative approaches to leverage complementary signals from both scRNA and ST data. One application for integrative analysis of scRNA-seq and spatial expression data is in cellular devolution, which helps uncover the cellular heterogeneity for each captured spot in ST data. Methods designed for cell devolution tend to employ corresponding scRNA-seq datasets as references, including CARD (Ma and Zhou, 2022), Cell2location (Kleshchevnikov et al., 2022), and Tangram (Biancalani et al., 2021). Another integrative application is to reconstruct the spatial association of cells in scRNA-seq data by using ST data as references. For example, scSpace (Qian et al., 2023) can identify spatial cell subpopulations by using a transfer learning model and then reconstruct cells onto a pseudo-space.

In conclusion, computational methods for multimodal integration are fast gaining popularity in single-cell analyses and provide innovative insights into complicated cellular systems and biological processes.

From a single species to cross-species

Cell types are evolutionary units with the potential for independent evolutionary variations (Arendt et al., 2016). The increasing number of single-cell organism-level datasets from different species creates opportunities to explore evolutionary relationships between cell types across species. Cross-species analysis provides valuable insights into the origin and evolution of cell types, shedding light on species-specific expression patterns. Furthermore, many studies use model organisms, such as mice, zebrafish, and fruit flies, in the investigation of fundamental biological processes and disease mechanisms so as to obtain valuable information about how cell types execute their functions in both healthy and diseased bodies, thereby contributing to our understanding of biological processes across diverse organisms. Specifically, cross-species analysis, such as in the context of complex diseases like IgA nephropathy (IgAN) (Chen et al., 2024b), offers a powerful tool in the identification and understanding of shared disease mechanisms across species and pinpointing potential therapeutic targets.

MetaNeighbor (Crow et al., 2018) works based on the notion that cells of the same type share greater numbers of similar gene expression patterns than cells from different types; it has been applied to systematically assess the transcriptional similarity between species. The results of MetaNeighbor among humans, mice, and monkeys (Han et al., 2020; Qu et al., 2022) imply that the cell type similarity in orthologous gene expression transcends mammalian species divergence. Moreover, through MetaNeighbor analysis between vertebrates and invertebrates, researchers constructed a cross-species cell-type evolutionary tree at the organism level for tracing the origin of major cell lineages during animal evolution (Wang et al., 2021b) and further identified common pathways that change across typical life span (Wang et al., 2023b) to shed light on the dynamic changes in cell function and regulation across species. However, MetaNeighbor analysis has limited usage for remote species because they share a few one-to-one ortholog genes. To address this concern, Tarashansky et al. (2021) proposed the self-assembling manifold mapping (SAMap) algorithm that incorporates the relative contributions of homologous genes derived from reciprocal BLAST for cross-species cell mapping. Using SAMap, researchers identified 1,209 homologous cell-type pairs across eight species, and only 36.5% (441 of 1,209) could be recognized by MetaNeighbor analyses using one-to-one orthologous genes (Li et al., 2022c). In addition, the result of benchmarking strategies for cross-species integration revealed the superiority of SAMap over BEC methods in integrating whole-body or organism-level atlases among species (Song et al., 2023). Notably, cross-species deep learning methods such as CAME (Liu et al., 2023c) and SATURN (Rosen

https://doi.org/10.1007/s11427-024-2770-x

et al., 2024) were constructed recently and demonstrated better performance than SAMap in terms of benchmarking. To deal with homologous genes, SATURN uses the concept of "macrogenes," defined as groups of genes with similar protein embeddings, and CAME applies graph neural networks to make full use of homologous information.

However, these methods and all the efforts are largely limited to pairwise comparisons followed by integrative analyses for the identification of homologous cell-type pairs across species. Recently, Mah and Dunn (2024) employed principal components as phylogenetic features to reconstruct cell type evolution across species through cell phylogenies. Although cell phylogenies allow researchers to phylogenetically define cell type via clade membership, they remain restricted to one-to-one homologies shared by species. In addition, Wang et al. (2021b) inferred the phylogenetic tree with IQ-TREE (Nguyen et al., 2015), indicating that cell types from the same cell lineage tend to be clustered in vertebrate pairwise comparison, while, in invertebrates, cell types from the same species tend to cluster. Therefore, cross-species analysis utilizing traditional cell phylogenetic trees, especially for homologous gene selection in lower species, remains challenging. Future work should consider the incorporation of paralogous and orthologous genes in constructing cell phylogenies to address this issue.

Overall, cross-species analyses enhance our understanding of the evolution of cell types. This analysis not only informs insights into the genetic and functional relationships between different species but also paves the way for a better understanding of the fundamental principles underlying the diversity and adaptation of cell types across various organisms. From the transcriptional similarity of homologous genes to the construction of cell phylogenetic trees, cell type evolution is increasingly getting integrated into the evolutionary perspective.

AI application for single-cell genomics

Single-cell technologies generate large amounts of data, typically with characteristics of high dimensionality, sparsity, and complexity, which makes the analysis using conventional computational methods difficult and unfeasible. To address these challenges, researchers are turning to AI methods as potential alternatives.

The application of AI algorithms in the field of single-cell genomics is expected to offer significant advantages for analysis, including batch effect correction, denoising, and dimensionality reduction, making the single-cell genomics datasets easier to interpret and visualize. AI algorithms method can mainly use two types of data as input: non-sequential data (such as cell profiles and gene expression from scRNA-seq) and sequential data (such as DNA sequences of cCREs from scATAC-seq).

Currently, AI algorithms can holistically account for all stages of single-cell genomics analysis (Brendel et al., 2022; Erfanian et al., 2023). For example, the applications of the AI models mentioned earlier include several categories: deconvolution (such as TAPE (Chen et al., 2022c)), BEC (such as SAUCIE (Amodio et al., 2019) and scGen (Lotfollahi et al., 2019)), integration of multimodal omics (such as GLUE (Cao and Gao, 2022)), and cross-species analysis (such as CAME (Liu et al., 2023c)). Other applications include cell-cell communication analysis (such as DeepCCI (Yang et al., 2023a)) and RNA velocity (such as DeepVelo (Chen et al., 2022e)). At the epigenetic level, AI models are used to transfer labels from scRNA-seq to scATAC-seq data, as demonstrated by scGCN (Song et al., 2021), and to construct TF regulatory networks from single-cell ATAC-seq data, as exemplified by DeepTFni (Li et al., 2022b). At the spatial level, specially designed AI models are increasingly available (Zeng et al., 2022). Methods including DeepST (Xu et al., 2022a), STAGATE (Dong and Zhang, 2022), and DeLTA2.0 (O'Connor et al., 2022) have been designed to capture both gene expression profiles and spatial information. Furthermore, CellCharter (Varrone et al., 2024) has been developed to identify the cellular spatial niches related to tissue remodeling and cell plasticity. In addition, Graph Convolutional Neural Networks for Genes (GCNG) (Yuan and Bar-Joseph, 2020) and HoloNet (Li et al., 2023a) are commonly used to infer cell-cell interactions at a spatial resolution.

For a deeper exploration of gene regulatory programmers, it is essential to construct AI models with million-level single-cell data. Big data-inspired AI models hold promise for unraveling complex gene interactions and regulatory mechanisms, thereby aiding in the understanding of disease processes and potential therapeutic targets. Recently, Geneformer (Theodoris et al., 2023), a transformer-based model pretrained on a corpus of 30 million scRNA-seq data, has been used to facilitate predictions in settings with limited data using transfer learning. The pretrained Geneformer model demonstrated excellent performance in predicting gene dosage sensitivity. The fine-tuning procedure for downstream tasks includes predictions on chromatin dynamics and network dynamics. Specifically, *in silico* perturbation includes *in silico* deletion and *in silico* activation. The ability to simulate these perturbations *in silico* is thus crucial as it can provide a valuable reference point for guiding downstream experimental work, facilitating more efficient and effective biological research. Similarly, the AI model scGPT (Cui et al., 2024) is pretrained on over 10 million cells and enables multi-omics integration. Additionally, the transformer-based cross-species foundation model GeneCompass (Yang et al., 2023b) is pretrained on an extensive scRNA-seq dataset of more than 120 million cells from humans and mice. Notably, single-cell transformers models have exhibited versatility in single-cell genomics, as evidenced by the resources available on the web: https://github.com/theislab/single-cell-transformer-papers (scBERT) (Yang et al., 2022). These models can tackle a wide range of downstream tasks, including cell type annotation, gene expression imputation, genetic perturbation, cell clustering, gene regulatory network inference, simulation, and denoising. Current single-cell transformer models can be categorized into various groups. In conventional transformer architectures, models like scBERT, BioFormers (Fang et al., 2023), and CellLM (Zhao et al., 2023) utilize standard transformer architectures, often employing Masked Language Modeling (MLM) with different loss functions and objectives for gene expression prediction or cell representation learning. Decoder-based models such as scMulan (Bian et al., 2024) and tGPT (Shen et al., 2023) rely on decoder architectures for tasks such as conditional cell generation or batch integration. Multiencoder models such as GeneCompass and scFoundation (Hao et al., 2024a) utilize multiple encoders, often to handle different aspects of the data, such as genetic perturbation effect prediction. In AE models, the multitask learning model (Pang and Tegnér, 2020) employs an AE with two transformer encoders for cell clustering. The Universal Cell Embedding (Rosen et al., 2023) incorporates

specialized embeddings, which are designed to capture information about the gene expressions or chromosomal positions, enhancing their ability to learn meaningful representations of cells. iSEEEK (Shen et al., 2022) and Geneformer leverage rank-based ordering of gene expression values within their input representations, thereby contributing to their performance in gene-related tasks. scFormer (Li et al., 2024) is a great example of a task-specific model that is designed to address specific challenges in single-cell analysis by using a modified MLM approach coupled with other techniques. Models such as TOSICA (Chen et al., 2023b), scTransSort (Jiao et al., 2023), STGRNS (Xu et al., 2023c), CIForm (Xu et al., 2023b), and TransCluster (Song et al., 2022) represent a diverse category of unsupervised approaches, each with its own set of objectives and unique embedding strategies. Overall, big data-inspired AI models implement comprehensive analyses across various biological applications.

However, the abovementioned AI models operate on single-cell data independent of the knowledge about the sequence, such as the gene expression profiles. Sequence-based AI models have emerged recently, offering direct modeling of the genome sequentially. These models offer the potential to analyze and understand genetic information in a more comprehensive and detailed manner, thereby providing valuable insights for various applications in genomic studies and personalized therapies. Nvwa (Li et al., 2022c) is a groundbreaking CNN model using 13-kb DNA sequences as inputs. It creates a unified framework to analyze the regulatory process of gene expressions and elucidate the mechanism of how different cell types use the same DNA sequence to encode different genes. Nvwa effectively utilized AI-derived *cis*-regulatory elements as well as identified complex regulatory rules, presenting a novel strategy for studying regulatory grammar in diverse biological systems. Another model based on the CNN architecture, scBasset (Yuan and Kelley, 2022), inputs 1,344-bp DNA sequences from scATAC-seq data. scBasset achieves cell clustering, data denoising, data integration across assays, and TF activity inference. In particular, scBasset can predict perturbations including single nucleotide mutations as well as TF motif insertions in the non-coding regions. To bridge the gap between epigenetic regulation and gene expression, AI models such as Huatuo (Xiao et al., 2023), ExpectoSC (Sokolova et al., 2023), and seq2cell (Schwessinger et al., 2023) are constructed by integrating sequence-based epigenetic models to facilitate the prediction of cell type-specific expression directly from DNA sequences. Huatuo and ExpectoSC are constructed on the framework of CNN-based DeepSEA Beluga (Zhou et al., 2018), while transformer-based Enformer (Avsec et al., 2021) is used in seq2cell. Notably, these AI models promote the estimation of the effects of single nucleotide mutations on the gene expressions. However, Sasse et al. (2023) proposed that current sequence-based AI models often fail to accurately predict the direction of mutation effects on the gene expression. These authors suggest leveraging of diverse genomes and biochemical processes as inputs for future deep learning models.

Chapter 6. Advances in single-cell atlas mapping

Mouse cell atlas

The evolution and enhancement of throughput and sensitivity in single-cell technologies have facilitated the development of cell atlases across numerous species. Typically, a conventional cell atlas should, at a minimum, depict a comprehensive catalog of cell types in either single or multiple tissues across at least one dimension (i.e., transcriptome, genome, proteome, or metabolome). Considering the importance of the transcriptome, the majority of current cell atlases are constructed from the transcriptomic level. As a preeminent animal model, high-throughput single-cell RNA sequencing (scRNA-seq) methodologies have been primarily utilized for the pioneer mapping of murine cellular atlas (*Mus musculus*). The cellular compositions of diverse murine tissues serve as a reference for studies investigating developmental, aging, and disease mechanisms. Despite the efforts to characterize the single-cell landscapes across various tissues (Haber et al., 2017), it was not until 2018 that the MCA unveiled an exhaustive single-cell transcriptomic profile encompassing nearly half a million cells across more than 40 organs and tissues (Han et al., 2018b). This repository stands as a pivotal resource conducive to the discernment of cell types across predominantly studied organs as well as unexplored tissues. The Tabula Muris Consortium reported another comprehensive single-cell transcriptome atlas of 20 mouse organs (Tabula Muris et al., 2018). The use of a combination of high-sensitivity full-length alongside 3′-end methods has facilitated a more detailed characterization of TF patterns across diverse cell lineages. These studies serve as a pioneering landmark, defining the first large-scale cell atlas spanning multiple murine tissues at the gene level. In the realm of chromatin accessibility, Cusanovich et al. (2018) pioneered the first single-cell chromatin accessibility atlas on a genome-wide scale, encompassing 13 distinct adult murine tissues. In addition, the presence of regulatory elements, being cell type-specific, provides a robust framework purposed for pinpointing cell types that are predominantly associated with common human diseases and traits. Other representative studies focused on cell heterogeneity in different systems. For example, Fang et al. (2022b) integrated the stromal cell atlas across multiple tissues during the developmental stages. Zeisel et al. (2018) mapped the single-cell transcriptome atlas of the murine mammalian nervous system in major anatomical regions, including the brain, spinal cord, peripheral sensory organs, and the enteric and sympathetic nervous systems. In addition, Kalucka et al. (2020) constructed the first mouse endothelial cell atlas and established metabolic signatures in different vascular beds.

The murine embryo serves as a homologous model that is instrumental in unraveling the integral molecular events entwined with the process of cell development. After the release of the adult mouse cell atlas, another two studies constructed a single-cell transcriptome atlas in 2019, with a focus on mouse gastrulation and organogenesis. Pijuan-Sala et al. (2019) adopted 10x genomics technology to comprehensively delineate the transcriptomic landscape across nine distinct temporal points, spanning from 6.5 to 8.5 days postfertilization. In a complementary study, Cao et al. (2019) innovatively employed sci-RNA-seq3 technology to analyze approximately 2 million cells derived from 61 embryos, spanning stages from 9.5 to 13.5 days of gestation. Collectively, these two studies provided continuous and complete lineage-specific gene expression resources. As for another mouse developmental cell atlas, Fei et al. (2022) unveiled a comprehensive mouse cellular differentiation atlas that captures the temporal series of morphogenesis, extending from the early embryogenesis stage at E10.5 to

postnatal day 21. The Xbp1 was identified as a pivotal master regulator governing cell fate decisions in mice. Mittnenzweig et al. (2021) reported a single-embryo, single-cell-level mouse gastrulation (E6.5 to E8.1) atlas. Qiu et al. (2024) constructed the largest scRNA-seq dataset to date of mouse embryonic development and presented a comprehensive atlas (Ontogeny of Mouse) consisting of an impressive 12.4 million cells, covering critical stages spanning from gastrulation to the postnatal period. Other single-cell multi-omics studies have supplemented results from other dimensions, including *cis*-elements, histone modifications, and spatial information (Gorkin et al., 2020; Jiang et al., 2023b; Smith et al., 2022). In the dimension of STs, Peng et al. profiled the spatial transcription atlas of early gastrulation (embryonic day 5.5 (E5.5) and E7.5), as well as the brain development (Jiang et al., 2023a; Peng et al., 2019a). Other researchers adopted Slide-seq and Stereo-seq to portray mouse whole embryo organogenesis at the subcellular resolution, respectively (Chen et al., 2022a; Sampath Kumar et al., 2023). Altogether, these two studies covered the development stages from E8.5 to E16.5. The open-access ST databases greatly facilitated other researchers in exploring these resources. Some reports have already been published in the field of aging research. For instance, the "*Tabula Muris Senis*" highlights aging-related perturbations and systemic inflammation in multiple tissues (Almanzar et al., 2020). The spatial and single-cell atlas of the cycling and aging female reproductive tract sheds light on the role of fibroblasts and fibrosis in extracellular matrix reorganization during a reproductive lifespan (Winkler et al., 2024).

Human Cell Atlas (HCA)

Initiated in 2016, the HCA consortium signifies a worldwide collaborative endeavor, diligently working to characterize a cellular reference blueprint of all human tissues (Regev et al., 2018). Enlisting the utilization of avant-garde single-cell genomic methodologies, this collective effort aims to profoundly understand human health and various diseases to provide insights into the fundamental mechanisms shaping our physiological and pathological states. The first draft of the HCA aims to profile over 100 million human cells from all organs in men and women. In this draft, the data resource will mainly be based on the mature scRNA-seq (also including other omics) profiling of cell suspension (Lindeboom et al., 2021). The end version will be a comprehensive, super-resolution spatiotemporal multiscale omics database encapsulating the dynamic landscape of every single cell type within the human body. This seminal resource, which is readily accessible to the scientific community, promises to serve as a reference atlas of inestimable value not only for basic biological discoveries but also as an instrumental tool for an array of clinical applications. The development of the HCA community is also expected to facilitate the translation of technology from basic research to clinical application scenarios. In 2019, marking another significant milestone, the National Institutes of Health (NIH) instituted the Human Biomolecular Atlas Program (HuBMAP). This endeavor inaugurated an additional research consortium dedicated to charting the terrain of the human body at the single-cell resolution (Snyder et al., 2019). In the year 2023, the HuBMAP project rendered a report encapsulating its advancements, as signified by the depiction of single-cell transcriptomic atlases for the human kidney and intestines. Moreover, it explored new frontiers by providing ST atlases of the fetal-maternal interface, thereby adding another feather of accomplishment to their ongoing scientific exploits (Greenbaum et al., 2023; Hickey et al., 2023; Lake et al., 2023).

Research communities in China are keen participants who are making salient contributions toward shaping the trajectory of HCA research. In an epoch-making study published in 2020, Han et al. (2020) presented the world's foremost single-cell transcriptome landscape that spans multiple tissues—a colossal dataset comprising over 700,000 cells from an excess of 50 tissue types—thereby providing an insight into the comprehensive organ system of the human body. This research unprecedentedly encompasses data about a comprehensive range of adult tissues, embryonic structures, as well as diverse cell lines. In parallel, they unfolded a thorough single-cell transcriptional profiling study covering an expanse of 15 organs derived from a standard adult individual (He et al., 2020). Progressing along the chronological thread of research within the sphere of multi-organ cell atlas, the Tabula Sapiens Consortium has contributed significantly in the scientific realm, crafted another atlas that encompasses multiple organs at a single-cell transcriptomic level, and distinctly identified nearly 500 cell types (The Tabula Sapiens Consortium et al., 2022). Furthermore, Eraslan et al. (2022) employed single-nucleus RNA sequencing as an efficacious stratagem to scrutinize frozen, banked specimens pertaining to eight human organ types procured from 16 philanthropic donors. This sophisticated exploration proposed an innovative framework for the utilization of frozen tissue samples, thereby significantly augmenting the Genotype-Tissue Expression (GTEx) project's repertoire.

The Human Developmental Cell Atlas (HDCA), a significant component of the HCA, is launched with the distinct objective of formulating an exhaustive reference map at the single-cell level to capture the nuances of normal organogenesis and early developmental stages. This ambitious undertaking has laid a robust foundation to further the understanding of aging and cancer and to invigorate the field of regenerative medicine (Haniffa et al., 2021). Within this realm, the laboratories of Qiao and Tang have conducted a series of pioneer multi-omics studies, charting new territories in understanding human germline cells and embryo implantation processes (Li et al., 2017; Li et al., 2018; Zhou et al., 2019). With the advancement and popularization of scATAC-seq methods, an increasing number of HCA studies are concurrently focusing on the cell type-specific regulatory elements brought about by chromatin accessibility data. Past research has followed along various stages of human embryonic development. Two such in-depth studies have produced comprehensive transcriptome and spatial datasets that cover critical phases in human embryogenesis. These phases encompass the early gastrulation period, specifically 16- and 19-days postfertilization, as well as the organogenesis stage, occurring 4 to 6 weeks into fetal development. These studies have provided an intricate understanding of these crucial development stages (Tyser et al., 2021; Xu et al., 2023d). The group's Genomic Architecture of Cells in Tissues (GeACT) project, alongside a comparable preprint multi-omics study, employed high-detectability scRNA-seq and scATAC-seq approaches. They systematically profiled the cellular atlas of human fetuses during gestation, focusing on the window of 8–21 weeks (Tian et al., 2020; Yu et al., 2021a). Shendure's laboratory pioneered the creation of the most extensive fetal HCA to date (transcriptome and chromatin accessibility). This

groundbreaking work catalogs 4 million individual cells derived from a spectrum of 28 fetuses, with estimated postconceptual age ranging from 72 to 129 days (Cao et al., 2020a; Domcke et al., 2020). Zhang et al. (2021b) pioneered the inaugural scATAC-seq atlas, elucidating over one million nuclei from a diverse set of 222 distinct cell types across 30 different adult human tissues. These authors also performed an integrative analysis of previous published fetal datasets and their study provided a research paradigm to identify the correlation of fetal and adult human cCREs with human traits and diseases. With the advancements of ST technology in recent years, these methodologies have now been incorporated into the mapping of HCAs to validate key molecular feature events during developmental processes with temporal and spatial resolution. Pan et al. (2023), for instance, reported the inaugural spatiotemporal transcriptome atlas of human embryos post-gastrulation. This pioneering research effort revealed organ-centric regulons that are postulated as potential determinants of cellular lineage. Moreover, they succeeded in categorizing the genes of specific cell types associated with developmental irregularities and infections, thereby contributing significantly to our current state of knowledge about embryonic development and disease pathology.

The advancements in single-cell atlases of various human organ tissues are indeed remarkable and captivating. Recently, Jiao's group (Li et al., 2023e) combined scRNA-seq and Stereo-seq to build the spatiotemporal developmental atlas of multiple human brain regions from 6–23 gestational weeks. This is the most extensive spatiotemporal developmental transcriptome atlas of the human brain to date, covering the greatest period and area; it uncovered the instrumental roles of diverse interactions between different types of glial cells and neurons in the regional specialization process of neuron subtypes. Notably, the Brain Research through Advancing Innovative Neurotechnology's Initiative-Cell Census Network (BICCN) alliance, propelled by the NIH, has achieved a series of progressions in the exploration of cell types and functions in humans, non-human primates, and rodent brains. This project revolves around the construction and comparative analyses of single-cell atlases for these species' brain maps. These multi-omics works cover the comprehensive brain cell types in different anatomic regions (Braun et al., 2023; Chartrand et al., 2023; Johansen et al., 2023; Jorstad et al., 2023a; Jorstad et al., 2023b; Kim et al., 2023a; Lee et al., 2023; Li et al., 2023f; Siletti et al., 2022; Tian et al., 2023b; Velmeshev et al., 2023; Weninger and Arlotta, 2023). Collectively, these efforts represent significant contributions to further our understanding of human biology and pave the way for future research in areas such as disease mechanisms, diagnostics, and therapeutics (Figure 8).

Cell atlas of other organisms

The nonhuman primate serves as another important model organism in the field of biomedical research. It has proven to be particularly instrumental in investigating human physiology, disease progression, and, notably, the intricacies of organ aging. The BGI in China generated the first nonhuman primate cell atlas (NHPCA) (Han et al., 2022). In their study, they provided an expansive cell transcriptomic atlas that incorporated data from more than 1 million cells spanning 45 tissues collected from adult nonhuman primates, specifically *Macaca fascicularis*. Significantly, the authors further established a high-definition ST atlas that intricately mapped the single-nucleus transcriptomic landscape of the entire macaque cortex (Chen et al., 2023a). A comprehensive cross-species transcriptomic data comparison comprising humans, macaques, and mice underscored the presence of cell types exclusive to primates, predominantly enriched within the fourth layer of cortical tissues.

Zhai et al. (2022) provided a comprehensive report on the gastrulation and early organogenesis cellular atlas of monkeys, which notably bolstered our comprehension of primate development. Their study demonstrated the transcriptomic characteristics pertinent to diverse cell types during primitive streak expansion, somitogenesis, the genesis of the gut tube, neural tube patterning, and neural crest differentiation periods. Building upon this body of research, another study combined scRNA-seq and scATAC-seq to construct a detailed map of gene expression and chromatin accessibility at the single-cell level across 16 organs in cynomolgus monkeys (Qu et al., 2022). Cross-species comparative investigations, encompassing mice, monkeys, and humans, have proposed a distinctly elevated equivalence in immune-related cellular characteristics, which is predominantly shared between monkeys and humans. In another critical brain region, the adult and aging cellular atlases of the hippocampus have also been reported (Hao et al., 2022; Zhang et al., 2021a). The authors unearthed the onset of proinflammatory reactions in aged microglial and oligodendrocyte cells to cultivate an adverse microenvironment impeding neurogenesis. Moreover, the group led by Liu has made notable strides in recent years by unveiling a remarkable series of primate tissue aging cell atlases. They assembled and established a unique single-cell transcriptome atlas of numerous aging primate organs, inclusive of the cardiopulmonary system, ovaries, liver, testes, muscles, arteries, and the pancreas (Huang et al., 2023a; Jing et al., 2023; Li et al., 2021a; Ma et al., 2021; Wang et al., 2020b; Yang et al., 2024; Zhang et al., 2020c). Most aging-associated transcriptional alterations, which have been identified specific to certain cell types, underscored an amplified systemic inflammation as a characteristic signature of primate aging. Moreover, this group ingeniously formulated tissue-specific therapeutic strategies geared toward intervening in the degenerative processes related to aging.

Within the domains of piscine and amphibious studies, organisms such as zebrafish, *Xenopus*, and axolotl serve as quintessential models, offering valuable insights into the intricate processes of embryonic development and holistic tissue regeneration. Jiang et al. (2021) unveiled a comprehensive zebrafish cellular landscape, representing over 250,000 unique cells that span across both embryonic and adult life stages. The authors shed light on the distinctive characteristics of the blastema population, which are a crucial component of the zebrafish's extraordinary caudal fin regeneration. Briggs et al. (2018) derived a comparative cell atlas of *Xenopus* and zebrafish development to identify embryonic cell states. The first single-cell transcriptome atlas of *Xenopus*, including multiple larval and adult organs, was reported by Liao et al. (2022). In their study, the authors discovered a ubiquitous regulatory mechanism encompassing the phenotype of antigen processing and presentation, which were consistently manifested in cells derived from disparate embryonic layers during metamorphosis. The Axolotl serves as another paradigm in the realm of developmental and regenerative biology. Notable single-cell RNA-sequencing studies investigating Axolotl limb regeneration

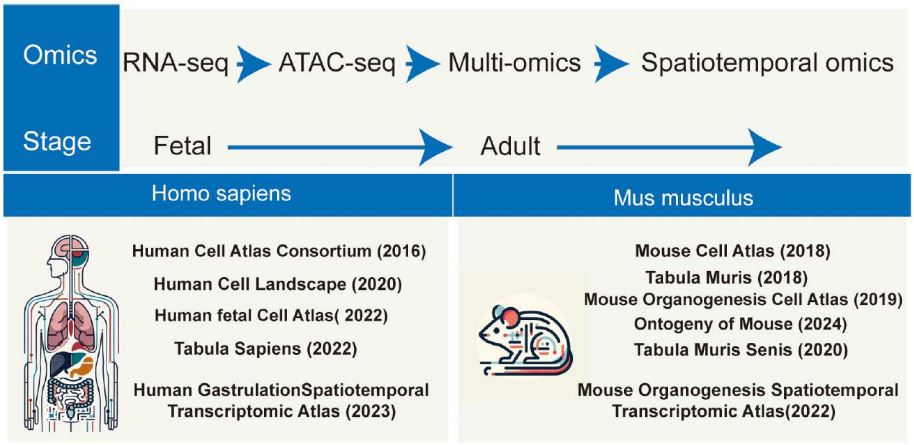

Figure 8. Mapping of human and mouse cell atlases at the single-cell level.

uncovered distinctive subsets of stem cells and identified pivotal regulatory activity within fibroblasts linked to the instigation of regeneration (Gerber et al., 2018; Lin et al., 2021b). Furthermore, the inaugural atlas of axolotl cells provided a comprehensive mapping of transcriptomic alterations that occurred throughout the course of metamorphosis and development (Ye et al., 2022). In the most recent advancements, an ST dataset indicated an injury-induced cluster of ependymal cells that manifested as a progenitor cell population with the ability to replace lost neurons (Wei et al., 2022). This research further delineated the presence of a state resembling immature neurons at the injury site.

The fruit fly (*Drosophila melanogaster*) is a classic model organism in the study of neurobiology and genetics. Li et al. (Li et al., 2022a; Lu et al., 2023) diligently constructed a cross-tissue single-cell transcriptome atlas encompassing both adult and aging flies that featured the gene signatures and correlating TFs manifesting across various cell types. Serving as a point of reference, these cellular signatures catalyze further explorations into genetic perturbations and disease modeling studies. Furthermore, the aging fly cell atlas sheds light on the evolution of age-related gene expression and the dynamic changes in pathways, such as the noticeable shift in lineage toward the fat body and the incidence of apoptosis in muscles. In a noteworthy addition, the authors established an aging clock model that ingeniously allows the prediction of an animal's age purely based on scRNA-seq data. Fly brain cell atlas has successfully deciphered the single-cell accessible chromatin and transcriptome atlas throughout adulthood and during the aging process, thereby identifying comprehensive regulatory regions associated with the processes of developmental trajectories, neurogenesis, reprogramming, and maturation (Davie et al., 2018; Janssens et al., 2022b). The comprehensive whole-body cell landscapes encompassing zebrafish, flies, and earthworms have been painstakingly assembled thereafter, paving the way for a genomic sequence-grounded, cross-species deep learning model. This state-of-the-art model holds the potential to accurately predict cellular lineages for creating new avenues for future research (Li et al., 2022c). Certain cell atlas mappings of terrestrial invertebrates have been recorded in the existing scientific publications. The cell atlas of the earthworm manifested a more comprehensive array of major cellular lineages than those examined previously (Shao et al., 2020). At the L2 larval stage, the developmental cell atlas of *C. elegans* incorporated extensive whole-body cellular representations encapsulated within a singular organism (Cao et al., 2017). Meanwhile, the lineage-resolved molecular atlas of *C. elegans* during embryogenesis and subsequent larval development evidenced a pattern of lineage priming within distinct cells right up unto the precipice of terminal fate determination (Packer et al., 2019). In addition, some review articles have encapsulated the application and advancements of single-cell technologies within the realm of plant studies (Bawa et al., 2022; Yu et al., 2023b).

Chapter 7. Single-cell insights for translational medicine

Single-cell genomics has emerged as a powerful tool in deciphering the complexities of human diseases, particularly in the field of cancer. By providing unprecedented resolution and insights into cellular heterogeneity, single-cell genomics has revolutionized our current understanding of disease mechanisms, treatment response, and patient stratification, as well as drug discovery and therapy selection in the context of precision medicine. In this section, we have highlighted how single-cell genomics can be translated into clinical applications and discussed its transformative impact on disease understanding and treatment strategies (Figure 9).

Transformative impact of single-cell genomics on cancer treatment

Hyperfine resolution of single-cell genomic: better targeting cancer cells in tumor progression

As a leading cause of death across the world, cancer poses formidable dangers with a widespread menace in mechanisms that are yet to be disentangled. Dissecting cancer mechanisms is of paramount importance in understanding the complexities of cancer biology, with pivotal implications in diagnosis, treatment strategies, and prognostic predictions. Thanks to the unprecedented aid of single-cell genomics, finer profiling of cancer cells provides deeper insights into the treatment and mechanisms of cancer, which encompass several key aspects such as cancer initiation and progression, cancer metastasis, and cancer heterogeneity.

Cancer initiation and progression are complicated processes

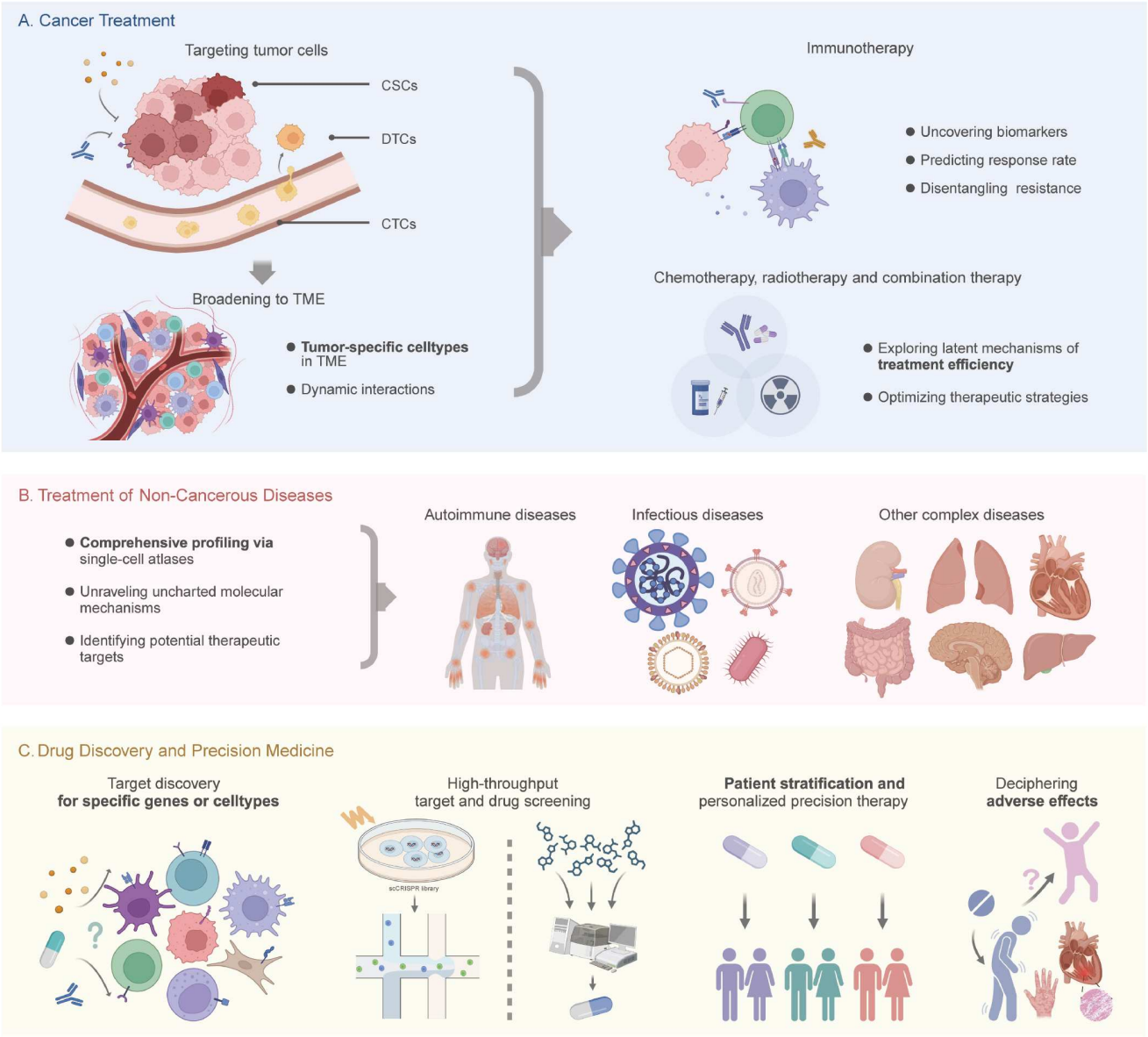

Figure 9. Clinical applications of single-cell genomics in cancer treatment (A), treatment of non-cancer diseases (B), and drug discovery and precision medicine (C).

that involve the acquisition of genetic and epigenetic alterations by normal cells resulting in the development of malignant properties and, finally, modifying the surrounding host cells into the tumor-supportive tumor microenvironment (TME). Cancer driver genes, whose mutations are more likely to induce cancer than those of passenger genes (Greenman et al., 2007), and cancer stem cells (CSCs), which arise by mutation from normal stem cells and maintain the self-renewal capability (Jordan et al., 2006), are thought to be two key initiators of tumorigenesis (Hanahan, 2022). Single-cell DNA sequencing (scDNA-seq) and its mapping to scRNA-seq can identify somatic mutations, copy number alterations, and chromosomal rearrangements in individual tumor cells, thereby revealing the phylogenetic relationships and temporal order of mutational events (Funnell et al., 2022). Single-cell omics technologies also illuminate the isolation and identification of CSCs, which helps understand the specific mutations and epigenetic changes that confer a selective advantage, resulting in the initiation and propagation of cancer (Frank et al., 2021; Lawson et al., 2015; Yang et al., 2017).

Single-cell omics technology has the potential to unveil the intricate cellular and molecular dynamics of cancer cells underlying cancer metastasis, which refers to the spread of cancer cells from the primary tumor site to distant organs or tissues—a complex process encompassing multiple biological events involving invasion, intravasation, dissemination, extravasation, and colonization (de Visser and Joyce, 2023; Gerstberger et al., 2023). By employing single-cell spatial, transcriptomic, and epigenomic analyses, several recent studies can successfully elucidate the immune landscape and intricate cellular interaction network within metastatic tumors (Karimi et al., 2023; Zhang et al., 2023b). Furthermore, CTCs and disseminated tumor cells (DTCs) represent tumor cells that respectively detach from the

https://doi.org/10.1007/s11427-024-2770-x

primary tumor and enter the bloodstream or lymphatic system (Lin et al., 2021a) as well as those that have already shed from the primary tumor and migrated to other tissues or organs (Kang and Pantel, 2013). Understanding the mechanisms of how CTCs and DTCs survive physical shear stress, evade anoikis, and evade clearance by immune cells is crucial in the field of metastasis research (Gerstberger et al., 2023; Ring et al., 2023). Single-cell genomics facilitates the characterization of CTCs at the cellular level in the context of their rare nature and uncovering their heterogeneity, plasticity, survival strategies, and metastatic capacity (Liu et al., 2023d; Reza et al., 2021; Suvilesh et al., 2022; Ye et al., 2023). In addition, the employment of single-cell sequencing technology for the detection of CTCs has made a noteworthy impact on facilitating early-stage diagnosis and prognostic monitoring of oncology.

Tumor heterogeneity, the diversity of cancer cells within a tumor or among different tumors, serves as a pivotal determinant in tumor evolution, treatment response, and drug resistance and can arise from genetic variations, epigenetic modifications, or TME factors (Prasetyanti and Medema, 2017). Deciphering tumor heterogeneity is paramount for the selection of targeted therapies determined based on molecular features. In early-stage investigations, tumor heterogeneity was originally defined based on discrepancies in driver gene mutations among distinct tumor types, reflecting inconsistent mutation patterns and frequencies (Wood et al., 2007). This concept gradually evolved to encompass the heterogeneity of mutation profiles within the same tumor, implying intratumoral diversity of neoplastic cells (Gerlinger et al., 2012), such as the absence of EGFR mutations in some lung cancer cells, which limits the efficacy of EGFR-targeted therapies in eradicating most cancer cell population. Subsequently, with the advancements in diverse single-cell technologies, the integration and temporal analysis of numerous single-cell pan-cancer atlases have given a more nuanced definition of tumor heterogeneity encompassing the comprehensive illustration of differential composition and dynamic interplay in the TME across distinct tumor types, disparate tumor regions, and among different patients (Barkley et al., 2022; Cheng et al., 2021b; Tang et al., 2023a; Zheng et al., 2021). Moreover, single-cell multi-omics technologies, such as scATAC-seq, scTrio-seq, and CITE-seq, provide supplementary perspectives for disentangling the intricacies of tumor heterogeneity (Bian et al., 2018; Satpathy et al., 2019). Hence, beyond the mere characterization of cancer cell genetic features, unraveling the heterogeneity and intricate interactions within the TME via single-cell genomics assumes a critical and indispensable role.

Unlocking new frontiers through single-cell genomics: the potential of targeting tumor-specific TME cell subtypes

Our understanding of cancer has evolved over the course of time, shifting from a narrow focus concentrated on cancer cells to a broader understanding encompassing the intricate surrounding ecosystem composed of non-malignant cells and the extracellular matrix, termed the TME. Despite their profound impact on the discovery of cancer driver genes, traditional technologies for bulk samples, which treat the tumor sample as a cohesive unit and primarily capture reads from cancer cells, fail to portray the detailed components and tumor-associated signals of the holistic tumor ecosystem. Thanks to their capability of deciphering cancer genomes at a hyperfine resolution, burgeoning single-cell technology has elevated the cancer research paradigm to reveal the cellular components and diversity of TME, some of which possess antitumor characteristics and others play pro-tumor roles.

Dissecting the anti-tumor response of various T cell types in the context of TME is the key to enhancing the efficacy of cancer treatment, which, therefore, entails rigorous and comprehensive profiling across multiple dimensions. When combined with TCR sequencing, scRNA-seq empowers the identification of various T cell subtypes in pan-cancer landscapes (Zheng et al., 2021). For example, exhausted $CD8^+$ T cells (Tex), featured with reduced cytotoxicity and effector function (Blank et al., 2019; Chen and Flies, 2013; Wherry, 2011), have been identified to demonstrate preferential enrichment in various tumors (Oliveira et al., 2021; Zhang et al., 2018; Zheng et al., 2017a), with LAYN as a signature gene exerting negatively regulation functions. Meanwhile, the $CXCL13^+$ activated precursors of exhausted $CD8^+$ T cells have been associated with tumor-reactive features and a better prognosis (Guo et al., 2018; Liu et al., 2022b), while $CXCL13^+CD4^+$ T cells, also denoted as T helper 1 (Th1) like T cells, seem to have anti-tumor potentials (Zhang et al., 2021d; Zheng et al., 2022), both of which collectively play a predictive role in immune checkpoint blockade (ICB) response (Liu et al., 2022b). Another subtype of $CD4^+$ T cells, regulatory T cells (Tregs), have been characterized into two clusters in single-cell profiling of nonsmall cell lung cancer (NCSLC) based on the expression pattern of TNFRSF9 (4-1BB), whose $TNFRSF9^+$ cluster highly express genes such as *IL1R*, which is associated with immunosuppressive functions (Guo et al., 2018).

Myeloid cells, which are comprised of several distinct populations, account for substantial proportions and mass of immune cells and are acknowledged as comparable contributors to tumor progression and metastasis (Sender et al., 2023), thereby attracting continuous scrutiny as potential targets for cancer therapeutic interventions (Engblom et al., 2016). However, it is imperative to note that not all subtypes of myeloid cells within the TME exhibit pro-tumor function, thereby necessitating a pressing demand for the application of single-cell genomics to dissect their heterogeneity and delineate the immunosuppressive clusters for more precise therapeutic targets (Cheng et al., 2021b). Multiple discrete single-cell studies have autonomously identified several pivotal subtypes of myeloid cells, such as $C1QC^+$ and $SPP1^+$ macrophages, which exhibit inflammatory and angiogenic signatures, respectively (Zhang et al., 2020a). Notably, neither of these subtypes aligns with the conventional M1 and M2 dichotomous phenotypes, thus underscoring the indispensability of employing single-cell technologies for precise macrophage classification. In hepatocellular carcinoma (HCC), single-cell analysis has contributed to the delineation of specific subgroups within tumor-associated neutrophils (TANs), such as $CCL4^+$, $SPP1^+$, and $PD\text{-}L1^+$ TANs, which have demonstrated pro-tumor attributes and potential as therapeutic targets (Xue et al., 2022). Also designated as mregDC, $LAMP3^+$ DC has been shown to possess the potential to migrate to lymph nodes and regulate multiple T cells and NK cells through crosstalk (Maier et al., 2020; Zhang et al., 2019c), such as by regulating the ICB response through cellular triads formed by mregDC, $CXCL13^+$ Th, and $PD\text{-}1^{hi}$ Progenitor $CD8^+$ T cells (Magen et al., 2023), thereby potentially regulating the anti-tumor functions of NK cells (Tang et al., 2023a).

Immune and stromal cells other than T cells and myeloid cells also hold paramount significance in the process of cancer

procession. A recent pan-cancer research presented a panorama of human natural killer cells by analyzing their heterogeneity and function in tumors from 716 patients with 24 cancer types, uncovering a subset denoted TaNK cells, which are enriched in tumors and exhibit anti-tumor function and unfavorable prognosis (Tang et al., 2023a). Another research demonstrated that LRRC15$^+$ cancer-associated fibroblasts (CAFs), as influenced by TGF-β, promote tumor growth and exert a direct inhibitory effect on CD8$^+$ T cell function to restrict the ability to respond to checkpoint blockade therapies (Krishnamurty et al., 2022).

Uncovering immunotherapy biomarkers and prospectively enhancing response rates with single-cell genomics

ICB therapies, such as anti-PD-1/L1 treatment, have demonstrated significant efficacy in controlling tumor progression in some cancer patients (Sharma et al., 2023). However, the response rates vary among different tumor types, and a considerable proportion of patients do not benefit from these treatments (Nishino et al., 2017). Moreover, some of the currently identified biomarkers, such as tumor mutational burden, mismatch repair (MMR) deficiency, and PD-L1 expression, do not consistently predict treatment outcomes across all cancer types. Consequently, there is a pressing need for the establishment of precise biomarkers that can accurately forecast the therapeutic response to ICB to enable optimal patient selection and maximize efficacy with minimal toxicity. By enabling high-resolution analyses of dynamic changes in the TME before and after treatment, single-cell sequencing technology can facilitate the exploration of immunotherapy mechanisms and help identify novel and effective biomarkers. Notably, a past study investigating PD-1 antibody treatment in basal cell carcinoma revealed the concept of clonal replacement, wherein the therapy improved the numbers of tumor-specific CD8 T cells with novel clonotypes (Yost et al., 2019). Another study in NSCLC through scRNA-seq and scTCR-seq analysis identified an expansion of precursor-exhausted T cells among treatment responders, originating from local expansion or peripheral blood T cell infiltration and replenishment, which was defined as clonal revival, expanding upon the theory of clonal replacement (Liu et al., 2022a). Subsequently, through a meta-analysis of single-cell data from patients undergoing ICB treatment in diverse cancers, researchers identified an elevation of tumor-specific CXCL13$^+$CD8$^+$ T cells in responders (Liu et al., 2022b; Zhang et al., 2021d). This past study also implied that CXCL13 can serve as a potential unified biomarker for tumor-reactive CD8$^+$ T cells and ICB response.

The TME encompasses a dynamic and intricate interplay among stromal cells, immune cells, and tumor cells, which plays a pivotal role in tumor progression and response or resistance to ICB therapy. By leveraging the high-resolution capabilities of single-cell analysis, researchers can discern distinct cellular subpopulations associated with immune therapy resistance, thereby shedding light on strategies to enhance ICB response rates or facilitate the development of novel immunotherapeutic interventions. For instance, through single-cell profiling, two clusters, namely ecm-myCAF and TGFβ-myCAF, have been identified to interact with Tregs and exert influential roles in ICB resistance (Kieffer et al., 2020); in another investigation, the association between iCAFs and ICB response was confirmed (Ma et al., 2023a). Notably, in gastric cancer, scRNA-seq analysis revealed the presence of fibroblast and endothelial subpopulations linked to ICB resistance, thereby opening up new avenues for immunotherapeutic exploration (Kang et al., 2022). Among the immune constituents within the TME, myeloid cells dominate, with specific subsets, including tumor-associated macrophages (TAMs), exhibiting suppressive properties that contribute to ICB resistance. In an in-depth scrutiny of HCC, single-cell data uncovered PD-L1$^+$ TANs as a potential immunotherapeutic target, considering their capacity to impede T cell cytotoxicity (Xue et al., 2022). A multi-dimensional single-cell analysis revealed the crucial involvement of SPP1$^+$ TAMs through the SPP1-CD44 axis and the targeting potential of enhancing the efficacy of immunotherapy in driving tumor immune evasion in AFP-positive HCC (APHC) (He et al., 2023).

Optimizing therapeutic strategies for chemotherapy, radiotherapy, and combination therapies

The mechanisms underlying the efficacy of chemotherapy and radiotherapy in clinical cancer treatment have extended beyond the cytotoxic effects to now include the modulation of the anti-tumor immune microenvironment (Galluzzi et al., 2015). Leveraging single-cell genomics methodologies has facilitated the exploration of latent mechanisms, thereby advancing the formulation of enhanced therapeutic strategies for chemotherapy, radiotherapy, and combination treatments, thereby potentially optimizing the anti-tumor outcomes. In gemcitabine plus cisplatin treatment of nasopharyngeal carcinoma, for instance, single-cell analysis showed that chemotherapy induces B cells to stimulate the expansion of T_{FH} and T_H1 through IL-1β. Remarkably, IL-1β has been identified as a potential biomarker for exploring the treatment efficacy, promisingly enhancement of patient management, and informing the selection of therapeutic approaches (Lv et al., 2023). An exploration of chemotherapy data in breast cancer confirmed the post-chemotherapy enrichment of ICOSL$^+$ B cells, a subset that enhances chemotherapy efficacy by eliciting an anti-tumor T cell immune response with its functionality intricately regulated by the complement-CR2 signal (Lu et al., 2020). This past study also posits CD55 as a potential target for enhancing chemotherapy effectiveness. Single-cell analysis has facilitated the identification of pre-existing chemoresistant genotypes within tumors prior to neoadjuvant chemotherapy, thereby presenting potential opportunities for predictive efficacy assessment before treatment (Kim et al., 2018). In the context of radiochemotherapy (RCT), single-cell sequencing has delineated the nuanced responses of specific TME subpopulations, revealing features such as MHC-II upregulation in epithelial cells and an augmentation of naïve CD4$^+$ cells, CD16$^+$ NK cells, and FCN1$^+$ M-MDSC cells after RCT (Liu et al., 2023a).

In recent years, the widespread adoption of combined immunotherapy with chemotherapy, radiotherapy, and other modalities has demonstrated a notable enhancement in the overall response rates and survival rates, surpassing the outcomes achieved with individual approaches (Meric-Bernstam et al., 2021). Single-cell technology has advanced the exploration of pertinent combination therapies, revealing diverse novel combinations for various cancer types and facilitating the investigation of mechanisms that can contribute to enhanced therapeutic efficacy across malignancies such as NSCLC (Cascone et al., 2023) and breast cancer (Bassez et al., 2021; Zhang et al., 2021d).

Treatment of non-cancerous diseases

In addition to its applications in cancer treatment, single-cell genomics holds significant promise in the diagnosis and treatment of a wide range of non-cancerous diseases, encompassing viral inflammatory conditions, immunological diseases, and neurological ailments. By leveraging the power of single-cell genomics, researchers can now unravel the intricate cellular heterogeneity and functional diversity inherent in these diseases, thereby shedding light on the underlying molecular mechanisms and identifying the potential therapeutic targets. This approach not only deepens our understanding of the complex cellular landscapes associated with various non-cancerous ailments but also paves the way for the development of precision medicine strategies.

Autoimmune diseases, which encompass >80 distinct disorders such as type 1 diabetes mellitus, inflammatory bowel disease, multiple sclerosis, rheumatoid arthritis, and systemic lupus erythematosus, are characterized by dysregulated responses of B cells and T cells toward self-antigens, which triggers the generation of autoantibodies and self-reactive T cells (Pisetsky, 2023). The intricate nature of autoimmune diseases poses formidable challenges, with marked inter-patient heterogeneity and a limited understanding of the underlying pathogenic mechanisms (Harroud and Hafler, 2023). Single-cell genomics technologies offer valuable insights by enabling the identification of TCR and BCR sequences from individual lymphocytes, which empowers researchers to decipher the intricacies of immune-related diseases and explore novel therapeutic avenues. Furthermore, the comprehensive profiling of immune disorders using large-scale single-cell atlases, along with the integration of healthy reference atlases, contributes significantly to the advancement of patient stratification and precision medicine in autoimmune diseases (King et al., 2021; Rood et al., 2022), such as in systemic lupus erythematosus (Arazi et al., 2019; Perez et al., 2022; Stewart et al., 2020), multiple sclerosis (Deng et al., 2021; Falcão et al., 2018), type 1 diabetes (Chiou et al., 2021; Hrovatin et al., 2023), Crohn's disease (Mukherjee et al., 2023), and rheumatoid arthritis (Nathan et al., 2022; Yamada et al., 2023; Zhang et al., 2019b), as well as other immunological diseases such as hypersensitivity syndrome (Kim et al., 2020).

Many infectious diseases, particularly viral infections, elicit infection-induced tissue inflammation, which causes immune dysregulation and profound clinical manifestations (Hill, 1998). To comprehend the intricate immunopathogenesis underlying these diseases, the utilization of large-scale single-cell atlases has emerged as an invaluable tool, providing a comprehensive depiction of tissue inflammation and enabling the identification of distinct immune subsets (Rood et al., 2022). Notably, our understanding of the immunological mechanisms governing certain infectious diseases remains limited. However, through the application of single-cell genomics analysis, we can unravel the complex interplay between pathogens and immune cells, thereby elucidating disease mechanisms and unveiling novel therapeutic targets. For instance, in the unprecedented global crisis of COVID-19, the integration of an extensive single-cell landscape encompassing various disease severities and stages uncovered an array of immune characteristics associated with COVID-19 infection (Ren et al., 2021). These atlases also furnish intricate details about the cellular and molecular responses elicited by the virus of interest, thereby delineating disparities in immune reactions to COVID-19 across diverse ethnicities and distinct cellular subpopulations (Aquino et al., 2023). Pioneering research utilizing scRNA-seq analysis has unraveled the intricate interplay between COVID-19 and myeloid cells through C-type lectin receptors and Tweety family member 2, which collectively contribute to immune hyperactivation (Lu et al., 2021). Moreover, the powerful combination of scBCR-seq and scRNA-seq techniques facilitates an effective identification of potent neutralizing antibodies, thereby offering promising therapeutic interventions and preventive strategies against prevalent infectious diseases (Cao et al., 2020c).

In addition to the conventional disease therapies available, single-cell technologies have exhibited a promising potential in enhancing assisted reproductive technologies (ART) as a crucial tool in infertility treatment. The rapid advancement in single-cell multi-omics has facilitated the generation of high-resolution single-cell atlases of preimplantation human embryos (Li et al., 2018), along with providing profound insights into the maternal-fetal interface interactions during the early gestation period (Vento-Tormo et al., 2018). These innovations have offered essential theoretical and technical foundations for the progressive evolution of ART. For instance, a clinical study on preimplantation embryo DNA methylation profiling demonstrated that embryos with optimal methylation states are associated with better clinical outcomes in ART (Gao et al., 2023).

Regarding other prevalent complex diseases, the application of single-cell atlases collected for diverse disease tissues and organs provides a heightened resolution for comparative analysis between pathological and healthy tissues. These cell atlases proved instrumental in unveiling novel cellular subpopulations expressing disease-relevant genes, deciphering alterations in cellular composition, and elucidating intricate intercellular crosstalk that underlies the pathogenic mechanisms across organs, such as in the liver (Aizarani et al., 2019), kidney (Schreibing and Kramann, 2022), lung (Sikkema et al., 2023), brain (Garcia et al., 2022; Kanton et al., 2019; Winkler et al., 2022), heart (Asp et al., 2019; Litviňuková et al., 2020), and gastrointestinal tract (Fawkner-Corbett et al., 2021). Furthermore, the integration of multi-organ single-cell atlases facilitates a comprehensive elucidation of the heterogeneity in the impact of disease-associated genes on distinct tissues, as well as tissue-specificity of the gene expression, cellular subtypes, and compositional distributions (Elmentaite et al., 2022; Eraslan et al., 2022; Mulder et al., 2021; The Tabula Sapiens Consortium et al., 2022).

Drug discovery and precision medicine by single-cell technology

Target discovery: broadening the focus from cancer cells to holistic TME

One of the key strategies in the pursuit of target discovery involves the direct targeting of driver genes in neoplastic cells. Preliminary analyses and screenings conducted through conventional genomic methods often lack the requisite granularity, necessitating substantial experimental efforts and temporal investments to obtain an expansive list of potential targets. Single-cell technologies offer an efficacious approach in this direction. For instance, leveraging scRNA-seq and scDNA-seq for

intratumor heterogeneity dissection and the analysis of epithelial-mesenchymal transition (EMT)-related genes is a strategic screening approach for tumor cell genes, which facilitates the discovery of new targets and innovative treatment strategies (Gavish et al., 2023; Pellegrino et al., 2018). Within the spectrum of other pathological conditions, the convergence of genome-wide association studies (GWAS) with single-cell expression quantitative trait loci (sc-eQTL) (Yazar et al., 2022), STING-seq (Morris et al., 2023), sc-linker framework (Jagadeesh et al., 2022), and other singular cell sequencing modalities facilitates the comprehensive integration of RNA sequences and human genetics. By delving into the molecular intricacies of GWAS mutations, this approach probes into the causal relationships among cell types, risk loci, and the associated diseases, thereby elucidating potential therapeutic target genes or cells.

Therapeutic approaches exclusively targeting cancer cells may encounter challenges in accessing certain targets, as evidenced by the limited efficacy of CAR-T in solid tumors. The burgeoning field of single-cell genomics, when coupled with advancements in integrative analyses, proffers a novel paradigm for target discovery, which pivots on the identification of therapeutically relevant subtypes of TME cells through single-cell omics so as to uncover pertinent genes and modulation of their cellular interactions, consequently skewing the balance in favor of a tumoricidal TME. This approach encompasses functional augmentation or recruitment of anti-tumorigenic cellular subtypes such as CXCL13$^+$ precursors of exhausted CD8$^+$ T cells (Liu et al., 2022b), CXCL13$^+$CD4$^+$ T cells (Zhang et al., 2021d), and the diminution of pro-tumorigenic cellular subtypes like activated Treg (Guo et al., 2018), SPP1$^+$ macrophages (Zhang et al., 2020a), TaNK (Tang et al., 2023a), LRRC15$^+$ CAFs (Krishnamurty et al., 2022), or the obstruction of their interactive pathways, potentially offering a more holistic and nuanced approach to neoplastic therapy.

Furthermore, various cell therapies can be harnessed across a spectrum of diseases including cancer (Weber et al., 2020), and the construction of a tumor antigen map through single-cell genomics furnishes potential targets for diverse cell therapies in cancer (Lareau et al., 2021b). Methods affiliated with single-cell genomics can thus foster the development of diverse cell or immune therapies, including deciphering of biological characteristics and adverse reaction mechanisms of CAR-T and unearthing potential novel strategies via single-cell technologies and methodologies such as scRNA-seq, scTCR-seq, scATAC-seq, CyTOF, and CITE-seq (Huang et al., 2023b).

Single-cell high-throughput target and drug screening

The potential targets identified through preliminary analysis necessitate further screening to ascertain their functional relevance to certain disease biology via genetically encoded perturbations or preclinical models. Perturbations induced by specific candidate genes often elicit distinct transcriptional changes across heterogeneous cellular subpopulations or tissue organs. Conventional methodologies, such as the use of a combination of pooled CRISPR screening approach with bulk RNA-seq, may obscure the nuanced functionality of these genes in minority cell populations (Bock et al., 2022). To surmount this limitation, emerging scCRISPR-seq such as perturb-seq (Dixit et al., 2016), which integrates with scRNA-seq, as well as ECCITE-seq, which has expanded CRISPR to multimodal single-cell screens (Papalexi et al., 2021), have been devised. These high-throughput methodologies empower the detection of the functional impact of multiple targeted genes across cellular subtypes, enabling streamlined and accurate screening of candidate targets. To further alleviate the experimental burden, several deep learning frameworks leveraging single-cell perturbation data have been developed. Models like GEARS, which incorporate graph neural networks, predict the downstream effects of multiple gene perturbations, even those not experimentally perturbed, thereby guiding experimental design and mitigating costs (Roohani et al., 2024). Another approach, known as CINEMA-OT, employs causal independent effect module attribution based on causal inference to undertake single-cell perturbation analysis and prognosticate therapeutic efficacy (Dong et al., 2023).

Following the identification of potential drug candidates, the subsequent steps involve high-throughput screening (HTS) of drug reagents. Conventional HTS methods, though effective, are often associated with high costs and significant time investments, as each screening step can only be conducted using a single dosage and under certain specific biological conditions (Macarron et al., 2011). The advent of single-cell technology, when integrated with HTS, allows for simultaneous testing under multiple dosages and diverse conditions (Van de Sande et al., 2023). Moreover, a range of new single-cell methods have emerged to enhance the cost-effectiveness and efficiency of HTS in combination with scRNA-seq, either by reducing the cost or increasing the throughput. For example, the Sci-Plex method, which employs Nuclear hashing and single-cell combinatorial indexing sequencing (sci-RNA-seq), can capture the global transcriptional response of cancer cells to thousands of independent perturbations at a single-cell resolution within a single experiment (Srivatsan et al., 2020).

Patient stratification and personalized precision therapy

The intricate heterogeneity observed among distinct tumors necessitates the precise stratification of different cancer patients during therapy selection, as it is a critical step in enhancing the efficacy and safety of cancer treatment. Current studies on tumor stratification, which rely on conventional genomics approaches, are primarily based on genetic or epigenetic profiles as well as alterations in oncogenic pathways (Guinney et al., 2015; Hoadley et al., 2018). The use of enhanced resolution in single-cell studies, especially in analyses of large-scale data, enables better-refined molecular stratification of cancer patients, which incorporates not only cancer cells but also immune cells in the TME, thereby guiding better targeting strategies for accurate, precise, and personalized immunotherapy. For instance, a classification of pan-cancer samples based on the composition of tumor-infiltrating T cells from a large-scale pan-cancer landscape of T cells can have distinct clinical implications and preferences for certain cancer types. These findings suggest that patient groups of cancer types featured with a high proportion of exhausted T cells and low resident memory T cells have worse overall survival and that cancer types enriched in the group of high numbers of exhausted T cells and low Treg, such as melanoma, are indicative of better response to ICB therapy (Zheng et al., 2021). T cell-based patient stratifications are associated with the efficacies of T cell-targeted therapy; as such, the stratifications considering integration of other TME cells might provide a broader perspective for deciphering the complexities in therapy selection for immunotherapy targeting a wider

range of cells (Joanito et al., 2022; Leader et al., 2021). In addition, the intricate nature of dynamic cellular interactions within the TME necessitates surpassing the limitations of categorizing single-cell types or individual cancer types, thereby pressing for further leaps in accurate patient stratifications based on large-scale single-cell integration of all TME components across numerous cancer types, which can ultimately guide the personalized selection of immunotherapy regimens.

The challenge posed by pan-essential cancer targets may engender inadvertent toxicities in patients (Chang et al., 2021). Patient stratification through single-cell genomics, hinging on the distinct enrichment profiles of diverse cell subtypes within patients, can allow the implementation of patient-specific targeted pharmacotherapies or treatment modalities, whose variance can be comparatively assessed by leveraging identified biomarkers. Patient stratification paradigms have now extended beyond the realm of cancer and have found applications in diverse diseases. In Crohn's disease, for example, scRNA-seq has unveiled the cellular module termed GIMATS, which is intricately linked to anti-TNF therapy resistance (Martin et al., 2019). Similarly, in human chronic inflammatory skin disease, a sophisticated tool that leverages single-cell characteristics has been devised for the purposeful classification of rash samples (Liu et al., 2022d).

Deciphering adverse effects of immunotherapy

In recent years, the landscape of cancer treatment has been significantly altered by the advent of immunotherapeutic strategies, with ICB standing out as a pivotal player therein. However, the clinical efficacy of these approaches is intricately entwined with the emergence of immune-related adverse events (irAEs), which affect a diverse array of organs, including the gastrointestinal tract, skin, liver, kidneys, and lungs, thereby introducing a formidable challenge in patient care and management (Martins et al., 2019). Despite the generally lower incidence of irAEs in comparison to that in conventional treatments, their nuanced and multifaceted characteristics impose substantial hurdles in the realm of clinical management, which, to some extent, constrain the progressive advancement of immunotherapies (Gangadhar and Vonderheide, 2014). The application of single-cell sequencing and analysis has emerged as a pivotal approach for determining the intricate mechanisms of irAEs and for propelling the mitigation of adverse side effects. Single-cell genomics research has elucidated the pivotal role of α-myosin as a significant autoantigen in immune checkpoint inhibitor-associated myocarditis (ICI-MC), which acts as a potent stimulus for clonal T-cell expansion and incites autoimmune inflammatory responses (Axelrod et al., 2022). Notably, peripheral α-myosin-specific T-cells have been identified as a promising biomarker for predicting myocarditis risk. Another study highlighted the enrichment of Temra $CD8^+$ cells in patients with ICI-MC and an upregulation of chemokines CCL5/CCL4/CCL4L2 (Zhu et al., 2022). In checkpoint inhibitor-induced colitis (+CPI colitis), comprehensive single-cell analysis of colitis data elucidated a pronounced proliferation and accumulation of cytotoxic effector $CD8^+$ T cells, indicating their potential implication in the pathogenesis of +CPI colitis while also identifying several cytokines, chemokines, and surface receptors as prospective therapeutic targets for any associated side effects (Luoma et al., 2020).

Single-cell genomics has thus emerged as a transformative approach with broad implications in cancer therapeutics as well as in the treatment of other complex diseases. By broadening the investigative perspective, such as dissecting TME and patient stratification, single-cell genomics can facilitate drug target discovery and screening, thereby propelling the efficacy of precision medicine. The ongoing exploration of the vast potential of single-cell genomics within the realm of healthcare holds significant promise toward heralding imminent transformative changes in clinical practices.

Chapter 8. Perspectives, challenges, and opportunities of single-cell genomics

With synergistic advances made in experimental and computational techniques in the fields of single-cell and spatial genomics, it has been revealed, in the context of both physiology and pathology, how heterogeneous the transcriptome of individual cells can be in a seemingly homogeneous cell population (Kelsey et al., 2017; Lein et al., 2017; Stubbington et al., 2017). Despite these advances, several challenges remain to be resolved in the successful application of single-cell genomics in the future. First, single-cell multi-omics has become an indispensable approach to elucidate the complexity of cellular processes. Although numerous techniques have been developed so far, there is much scope for improvement in the protocols to address issues such as limited coverage so as to capture more informative and complete genomic data (Baysoy et al., 2023). Second, spatial genomics has revolutionized our current understanding of cellular processes by providing the exact location of cells. Although several spatially resolved single-cell multi-omics techniques have been developed so far, most of these methods are low-throughput and have low sensitivity, and hence cannot capture the complete spatial architecture of tissue with most of the molecular information (Baysoy et al., 2023). Third, the cost of single-cell spatial genomics is very high, and there is a need to develop cost-effective and scalable methods for their routine application. Another important challenge in this area arises from data integration. Matched data from single-cell multi-omics are relatively easy to handle since distinct modalities can be tethered to a single cell. However, future improvements are required for the integration of unmatched data, considering the excessive single-cell data available with singular modality that has been generated over the past decade (Argelaguet et al., 2021). To move forward to successful clinical translation, there is a need to bridge the gap between research and clinical application for improving patient outcomes. Finally, since single-cell and spatial genomics methods amass huge amounts of genomic information, ethical issues related to data privacy and application in clinical settings would also need to be considered.

Increase data coverage

Recent genomic technologies for studying single-cell gene expression (scRNA-seq), chromatin accessibility (scATAC-seq), and chromatin interactions (scHi-C, Dip-C) have tremendously advanced our understanding of cellular function and diversity. However, there is a major obstacle in this direction—data sparsity. Although the current single-cell data coverage is generally adequate for identifying cell types and clustering, it lacks the depth needed for a thorough characterization of individual cells. The additional data required include their

intricate properties, specific functions, response to external stimuli and drugs, and their differentiation or reprogramming paths. To fully harness the potential of these technologies and to aspire toward creating cellular reference maps of organisms and detailing the position, function, and characteristics of every individual cell, the coverage of single-cell data must be increased significantly.

In scRNA-seq, the utilization of random priming and *in situ* polyadenylation have enabled full-length coverage and detection of total RNA (McKellar et al., 2023; Salmen et al., 2022; Xu et al., 2023f). In addition, the application of diverse single-molecule labeling strategies is poised to elevate the detectable spectrum of RNA (Mahat et al., 2024; Wu et al., 2024b; Yao et al., 2023; Ye et al., 2024b). These innovative methods can extend their utility to STs or *in situ* sequencing, thereby enhancing the sensitivity of single-cell and spatial genomic methods.

In sn-ATAC-seq, different strategies may have the potential to increase the sensitivity. For instance, Tn5 transposase has been applied to insert T7 promoters (T7P) into genomic DNA (gDNA) to facilitate *in vitro* linear amplification and thereby reduce PCR-introduced bias during exponential amplification (Bartlett et al., 2021; Chen et al., 2017a; Lake et al., 2018; Sos et al., 2016). More recently, Mulqueen et al. implemented a method using a single-adaptor transposition strategy (Mulqueen et al., 2021) called s3-ATAC that utilizes a forward-adaptor containing a uracil base. By employing a uracil-intolerant polymerase for gap-filling, this approach prevents the incorporation of a forward adapter sequence on both ends. Subsequently, the reverse adapter is introduced with a polymerase-based adaptor switch. In this step, a reverse primer with locked nucleic acid is incorporated into the template by using a uracil-processing polymerase. In addition, the adoption of long-read strategies in single-cell analyses, such as scNanoHi-C (Li et al., 2023c), has shown promise, demonstrating approximately nearly complete coverage of genomic regions (99.3% of 100-kb pair genomic bins).

In spatial and *in situ* sequencing, 10x Genomics Xenium *in situ* gene expression profiling has expanded the target panel from 300 to 5,000 genes for single-cell spatial imaging (https://www.10xgenomics.com/platforms/xenium). Deep-STARmap and Deep-RIBOmap enabled single-cell *in situ* transcriptional and translational gene expression of 200-μm thick tissue blocks (Sui et al., 2024). Besides single-molecular DNA sequencing (Pacbio, Nanopore) and smFISH-based strategy (Vizgen, NanoString, 10x Genomics Xenium), single-molecule protein sequencing has emerged as a promising way for integrated omics (Reed et al., 2022). Subtle cellular structure analyzing methods for nuclei, mitochondria (Guo et al., 2023), and ribosome (VanInsberghe et al., 2021; Zeng et al., 2023a) provides more elaborate information about cytogenetics and cell kinetics.

Cumulatively, these innovations have significantly improved the complexity of libraries for individual cells, thereby enhancing detection sensitivity and disclosing subtle cellular variations. Advancements along these lines are, therefore, poised to further deepen our insights into the myriad of biological intricacies inherent in a single cell.

Increase multi-omics of single-cell mapping

Cellular mechanisms of biological functions involve the central dogma of genetic information converting from DNA to RNA to protein. The traditional mono-omics methods can not reveal the direct causality between different types of molecular information. Multi-omics technologies, however, represent a pivotal advancement in molecular biology that is designed to capture multiple types of molecular information within the same biological samples by a single experiment. To understand the molecular hierarchy of genome structures and functions, methods like Chromatin Interaction Analysis with Paired-End Tag (ChIA-PET) (Fullwood et al., 2009) have been developed to detect TF binding sites and associated chromatin interactions, while Chromatin Interaction Analysis with ATAC-seq (ChIATAC) (Chai et al., 2023) was created to simultaneously measure chromatin accessibility and interactions in cell populations. However, despite the great strides accomplished and the known potential, two primary challenges persist. First, Hi-C-related data is inherent to limited resolution, which hampers the detailed characterization of interactions between *cis*-regulatory elements and their target gene promoters. Second, multi-omics single-cell studies involving 3D genome analysis have limited throughput. Particularly, the latter represents a serious bottleneck, constraining our ability to perform comprehensive and detailed studies at scale and, thus, slowing the pace of gathering insights into the 3D genomic organization across a diverse range of cell types and conditions.

The next frontier in multi-omics research lies in overcoming current limitations to achieve high-resolution mapping of chromatin interactions, particularly to gene-centric contacts involving specific promoter-enhancer interactions that directly modulate the gene expression for determining cell identity. By combining this approach with RNA expression information within the same cell, researchers can investigate the relationship between the genome structure with transcription. Additionally, incorporating protein information may provide mechanistic insights that span from DNA to RNA to protein synthesis—a process that is at the heart of the central dogma of molecular biology.

Equally critical is tackling the chemical incompatibilities inherent in multi-modal methodologies, which often lead to reduced detection capabilities when compared with that with single-modality analyses. In these directions, it is foreseeable that additional modalities of molecular information can be included in the future multi-omics single-cell methods for simultaneously mapping RNA expression and interactions among DNA-DNA, DNA-proteins (e.g., histone and TF), DNA-RNA, and RNA-RNA. Such multi-omics molecular information derived in the same cell will greatly enhance our ability to establish the causal relationship between 3D epigenomic states to their functional outcomes in transcription regulation and cellular functions. To support widespread, large-scale applications of multi-omics research, systematic comparison and standardization of multiple methodological variations will be essential. Major projects such as ENCODE, BICCN, and 4DN have used multi-omics technologies to build comprehensive atlases (BRAIN Initiative Cell Census Network (BICCN), 2021; Dekker et al., 2017; Peng et al., 2021).

Multi-modal datasets can act as a "molecular" bridge (Hao et al., 2024b) for connecting single-cell datasets across modalities. Although the existing mono-modality single-cell research has provided a deep dive into specific biological aspects, its insights remain isolated, which is limited to the scope of each modality. Encompassing information from disparate modalities, multi-modal datasets can enable effective translation and integration

within these modalities, free from the reliance on constraining biological presumptions. In summary, advancements in these directions will not only refine our current understanding of the cellular machinery but also enhance our ability to decipher the molecular underpinnings of a healthy state and a diseased state.

Integrate single-cell multi-omics with spatial omics

For multi-cellular organisms, cellular function is highly dependent on the context of spatial localization within a given tissue or organ. Although recent advancements in imaging and sequence-based spatial technologies have revolutionized our understanding of how gene expression patterns are influenced by the spatial arrangement of cells within tissues, spatial omics technologies are currently constrained by limited capturing area and low chemistry efficiency. As a result, spatial omics data have a reduced capacity to capture genomic and transcriptomic features when compared with single-cell analyses. To capitalize on the existing single-cell consortia data (Regev et al., 2017), efforts have been made to integrate the scRNA-seq data with the ST data. Despite these efforts, the existing computational tools for integration often lack precision; they typically estimate the cell types within each spatial spot rather than obtaining a single-cell resolution (Wan et al., 2023). Advancing computational tools is critical in enhancing single-cell analyses, as it enables researchers to fully leverage spatial context toward understanding cellular functions and interactions. This development is key to unlocking the full potential of spatially informed single-cell biology.

Recently, Russell et al. (2024) introduced a groundbreaking method that involved tagging tissues with spatial barcodes, followed by the recovery of the desired nuclei for robust single-cell RNA-seq analysis. These spatial barcodes are ingeniously designed to integrate seamlessly with downstream single-cell processing workflows. The spatially barcoded nuclei can then be input into commercially available single-cell microfluidic platforms, thereby bridging the single-cell measurements with spatial genomics insights. Although the Slide-tag technique currently faces challenges due to sample loss, which limits analysis to only a subset of the tissue, it represents a significant step forward in linking single-cell analysis with spatial context, thereby enhancing our understanding of cellular functions within their native environments.

Alternatively, an appealing strategy to link spatial omics data with single-cell multi-omics data, or conversely, to assign single-cell multi-omics information with spatial localization in the tissue context, is through computational integration. For matched ST data and single-cell multi-omics datasets with RNA, 3D genome, and epigenome modalities from the same biological sample, the shared spatial RNA and the single-cell RNA data can function as an information bridge to connect tissue anatomic location with the 3D epigenomic information. Therefore, future development in this direction by leveraging the combined capabilities of single-cell precision and spatial resolution is positioned at the cutting edge of scientific discovery. Technically, future directions for spatial omics may involve click chemistry, chemical labeling, and a computational-based "optics-free" system for single-molecular imaging reconstruction (Weinstein et al., 2019). The application foreground of spatial omics may focus on the 3D construction of a large number of clinical samples (especially formalin-fixed paraffin-embedded tissues) for biomarker detection. High-resolution, ease-of-use, and cost-efficient open-source spatial omics frameworks are expected to lead to the widespread application of these technologies in medical organizations (Schott et al., 2024).

AI for single-cell science

On one hand, single-cell technologies produce highly systematic data, which makes them well-suited for AI training. On the other hand, AI has been providing revolutionary applications in single-cell genomics, thereby elevating the field to new heights. It has revolutionized data preprocessing, batch-effect correction, data normalization, dimension reduction, feature selection, cell clustering, cell type annotation, graphic presentation, and other aspects. We have provided here an overview of the broad scope of AI applications in single-cell science, including (i) advanced deep learning models (Chen et al., 2022c; de la Fuente et al., 2023; Menden et al., 2020) that enable deconvolution from bulk to single-cell data, thereby capturing individual cell behaviors with precision; (ii) AI methods (Amodio et al., 2019; Arisdakessian et al., 2019; Deng et al., 2019; Eraslan et al., 2019; Lopez et al., 2018) can represent cells in a low-dimensional linear space or latent form to fill in any missing or unseen gene expression values, thereby enhancing the overall data quality; (iii) AI-driven algorithms like SAUCIE (Amodio et al., 2019), scGen (Lotfollahi et al., 2019), and scVI (Lopez et al., 2018) correct batch effects to align datasets from different experimental conditions; (iv) by utilizing techniques such as GPTCelltype (Hou and Ji, 2024), AI automatically classifies or labels different cell types based on the gene expression profiles; (v) AI models like GLUE (Cao and Gao, 2022) incorporate multiple layers of biological data to provide deeper insights into cellular processes; (vi) AI models like CAME (Liu et al., 2023c) and SATURN (Rosen et al., 2024) integrate datasets across multiple species to understand cell type evolution and conservation; (vii) AI methods like DeepST (Xu et al., 2022a), STAGATE (Dong and Zhang, 2022), and DeLTA2.0 (O'Connor et al., 2022) capture both the gene expression profiles and spatial information, thereby enriching our current understanding of cellular organization within tissues; (viii) AI techniques such as DeepCCI (Yang et al., 2023a) enable the comprehensive exploration of interaction between different cell types to reveal complex cellular dynamics; (ix) AI techniques like DeepVelo (Chen et al., 2022e) enable reconstruction of temporal trajectories of cellular states, thereby unraveling dynamic processes such as differentiation and development; (x) AI models such as Nvwa (Li et al., 2022c) decodes the secrets of regulatory grammars, which plays a pivotal role in elucidating core regulatory mechanisms governing gene expression and providing valuable insights into cell fate determination; (xi) pre-trained models like Geneformer has an impressive success in the context of gene regulatory network inference in the field of single-cell science. Overall, AI is shaping single-cell data analysis (Ma and Xu, 2022).

AI models have become a powerful tool in the field of spatial data, with their applications ranging from cell type identification to functional perdition. These applications leverage AI's ability to analyze complex spatial patterns and relationships within biological systems. The key areas where AI can significantly contribute include (i) models like DeepST and STAGATE leverage graph-based methods to reconstruct spatial relationships, which provides a more comprehensive understanding of spatial organization; (ii) AI algorithms like DSTG (Song and Su, 2021)

and GTAD (Zhang et al., 2023c) utilize graph-based convolutional network to perform spot deconvolution, parsing spot-based data into single-cell level expression profiles; (iii) models such as SpaGCN (Hu et al., 2021) and GLISS (Zhu and Sabatti, 2020) integrate gene expression, spatial location, and histology to identify spatial domains and spatially variable genes through the application of graph convolutional networks; (iv) models like GCNG and HoloNet are commonly employed to infer cell-cell interactions at a spatial resolution. Notably, graph-based models have emerged as a powerful tool for spatial data analysis. The ability to represent complex spatial relationships within tissues as graphical topological structures has unlocked new possibilities for understanding tissue organization and function to greater depths.

Future AI models within the field of single-cell science are expected to leverage multi-omics data even further so as to enhance our understanding of gene regulatory rules and syntax. The evolution of AI architectures from CNN-based to attention-based models is a testament to the power of the latter in dealing with complex, long-range relationships. Current specialized transformer models like scGPT, scMoFormer (Tang et al., 2023b), and scCLIP (Xiong et al., 2023) have been constructed for single-cell omics prediction. More importantly, systems biology will witness considerable benefits from AI models inspired by big data, which will shift it from a mere integrative subject to one with heightened predictive capabilities.

Clinical translations

In addition to their profound impact on our current understanding of disease pathogenesis, drug discovery, and precision medicine, single-cell and spatial genomics have the potential to be transformative in developing targeted therapies and bridging the gap between research and clinical applications for improved patient-related clinical outcomes.

Previous studies have suggested that the majority of variants identified through GWAS are situated in the extensive non-coding regions of the human genome, which comprises approximately 91% of the total genome (Claringbould and Zaugg, 2021). These non-coding regions do not directly affect protein coding, implying an enhancer-mediated regulatory role. However, GWAS studies often fail to link variants to specific target genes and do not specify the functions of these variants in specific cell types or tissues. As such, linking GWAS variants to their target genes is crucial for developing targeted therapies.

Single-cell multi-omics 3D genome mapping technologies can enable precise mapping of enhancers at single-cell resolution (Liu et al., 2023f). This precise mapping will, in turn, facilitate the connection of GWAS loci with their target genes, thereby enabling the development of drugs targeting these genes and allowing the enhancers to serve as direct targets for putative therapy. For example, in diseases where enhancers drive the expression of undesirable gene products, editing the enhancer sequence via CRISPR editing techniques presents a promising treatment option. This therapy has been pioneered in transfusion-dependent β-thalassemia and sickle cell disease (Frangoul et al., 2021). Unlike targeting specific proteins or enzymes, targeting disease-related enhancers offers greater cell type specificity and tissue specificity, with the promise of minimizing the disruption of gene functions of other cell types.

In clinical practice, the application of single-cell genomics can benefit the oncological, hematological, and immunological aspects of diagnostic biomarkers and therapeutic target detection (Lim et al., 2023), especially for epigenetic targeted therapy

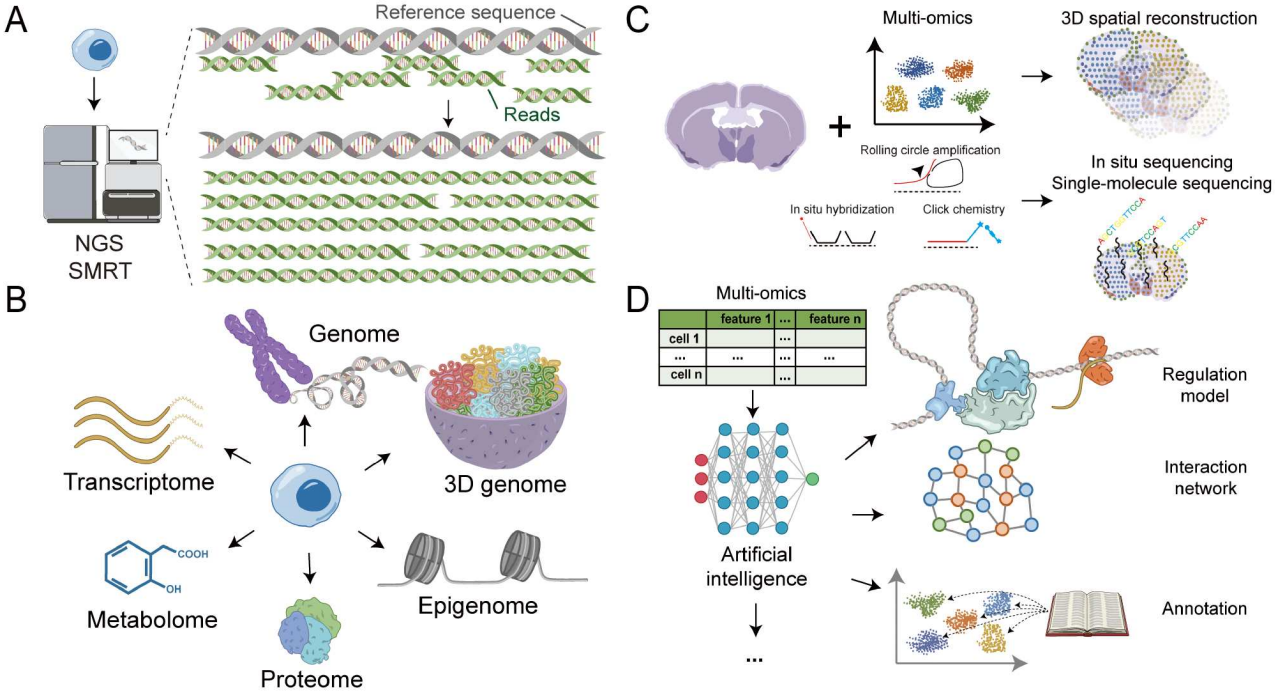

Figure 10. Future directions in single-cell and spatial genomics. A, Increasing data coverage in single cells (NGS, next-generation sequencing; SMRT, single-molecule real-time sequencing). B, Enabling of multi-modal detection in single cells. C, Integration of single-cell genomic data with spatial genomic and *in situ* sequencing data. D, Leveraging AI for single-cell and spatial genomics.

https://doi.org/10.1007/s11427-024-2770-x

(Yamagishi et al., 2024). Multimodal single-cell genomics approaches have enabled wholescale sequencing of small, frozen clinical specimens (Wang et al., 2023c). Spatial genomics of clinical FFPE samples can further provide a retrospective pathology landscape for tumor treatment (Jackson et al., 2020). Technical advances in single-cell and spatial genomics have provided solutions for different clinical samples. The cell-type specific regulatory events in disease can inspire cutting-edge therapies, including genome editing, oncolytic virus (Ling et al., 2023), cell and immunotherapy, and epigenetic targeted small molecules.

In the forthcoming years, the advent of single-cell and spatial genomics is poised to facilitate the establishment of the regulatory atlas, detailing cell type-specificity and the developmental stages' specificity. Fine-tuning the activity of regulators or regulatory elements can empower clinicians to achieve precise control over gene expression, thereby expanding the possibilities for targeted treatments. This approach may emphasize a broader regulatory genomic landscape, thereby laying a novel groundwork for future drug development.

Conclusion

With improvements in throughput, sensitivity, and modalities, single-cell and spatial technologies are expected to drive a paradigm shift in all facets of biological investigation. In addition, the synergistic integration of AI stands as a cornerstone, poised to bolster these efforts to enable more advanced analysis and interpretation of genomic data (Figure 10). Together, these innovations promise to usher in a new era of discovery, with the potential to reveal new biological insights and translate them into clinical applications for improved human health and disease management.

Summary

Recent advancements in single-cell and spatial genomics have revolutionized the research on cellular states and heterogeneity. Innovations in single-cell multi-omics and spatial genomics have opened up ways to integrate molecular data with the mapping of cellular features within tissue contexts. These technologies have enabled the construction of detailed cellular atlases, offering new insights into tissue organization and disease mechanisms. In addition, improved computational tools have enhanced data integration, while systematic data from single-cell and spatial genomics have paved the way for digital cell atlases across various systems. Importantly, AI-driven predictive foundation models are emerging to enable *in silico* perturbation screening and synthetic regulatory network design to advance our current understanding of complex diseases and guide targeted therapeutic strategies.

Compliance and ethics

The authors declare that they have no conflict of interest.

Acknowledgement

This work was supported by the Key R&D Program of Zhejiang (2024SSYS0022), the National Natural Science Foundation of China (31930028, 91842301, 32370608, 32250710678, 32200073, 32250710678), Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (2021R01012), "Pioneer" R&D programs of Zhejiang Province (2024C03005), STI2030-Major Projects (2021ZD0202502), the National Key Research and Development Program of China (2022YFC3400400, 2018YFA0800503, 2018YFA0107804) and Beijing Municipal Science and Technology Commission (Z221100007022002).

Supporting information

The supporting information is available online at https://doi.org/10.1007/s11427-024-2770-x. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

Adil, A., Kumar, V., Jan, A.T., and Asger, M. (2021). Single-cell transcriptomics: current methods and challenges in data acquisition and analysis. Front Neurosci 15, 591122.

Ai, S., Xiong, H., Li, C.C., Luo, Y., Shi, Q., Liu, Y., Yu, X., Li, C., and He, A. (2019). Profiling chromatin states using single-cell itChIP-seq. Nat Cell Biol 21, 1164–1172.

Aizarani, N., Saviano, A., Sagar, A., Mailly, L., Durand, S., Herman, J.S., Pessaux, P., Baumert, T.F., and Grün, D. (2019). A human liver cell atlas reveals heterogeneity and epithelial progenitors. Nature 572, 199–204.

Almanzar, N., Antony, J., Baghel, A.S., Bakerman, I., Bansal, I., Barres, B.A., Beachy, P.A., Berdnik, D., Bilen, B., Brownfield, D., et al. (2020). A single-cell transcriptomic atlas characterizes ageing tissues in the mouse. Nature 583, 590–595.

Alon, S., Goodwin, D.R., Sinha, A., Wassie, A.T., Chen, F., Daugharthy, E.R., Bando, Y., Kajita, A., Xue, A.G., Marrett, K., et al. (2021). Expansion sequencing: spatially precise *in situ* transcriptomics in intact biological systems. Science 371, eaax2656.

Amaral, P., Carbonell-Sala, S., De La Vega, F.M., Faial, T., Frankish, A., Gingeras, T., Guigo, R., Harrow, J.L., Hatzigeorgiou, A.G., Johnson, R., et al. (2023). The status of the human gene catalogue. Nature 622, 41–47.

Amodio, M., van Dijk, D., Srinivasan, K., Chen, W.S., Mohsen, H., Moon, K.R., Campbell, A., Zhao, Y., Wang, X., Venkataswamy, M., et al. (2019). Exploring single-cell data with deep multitasking neural networks. Nat Methods 16, 1139–1145.

Anchang, B., Hart, T.D.P., Bendall, S.C., Qiu, P., Bjornson, Z., Linderman, M., Nolan, G.P., and Plevritis, S.K. (2016). Visualization and cellular hierarchy inference of single-cell data using SPADE. Nat Protoc 11, 1264–1279.

Angerer, P., Simon, L., Tritschler, S., Wolf, F.A., Fischer, D., and Theis, F.J. (2017). Single cells make big data: new challenges and opportunities in transcriptomics. Curr Opin Syst Biol 4, 85–91.

Angermueller, C., Clark, S.J., Lee, H.J., Macaulay, I.C., Teng, M.J., Hu, T.X., Krueger, F., Smallwood, S.A., Ponting, C.P., Voet, T., et al. (2016). Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. Nat Methods 13, 229–232.

Aquino, Y., Bisiaux, A., Li, Z., O'Neill, M., Mendoza-Revilla, J., Merkling, S.H., Kerner, G., Hasan, M., Libri, V., Bondet, V., et al. (2023). Dissecting human population variation in single-cell responses to SARS-CoV-2. Nature 621, 120–128.

Aran, D., Hu, Z., and Butte, A.J. (2017). xCell: Digitally portraying the tissue cellular heterogeneity landscape. Genome Biol 18, 220.

Aran, D., Looney, A.P., Liu, L., Wu, E., Fong, V., Hsu, A., Chak, S., Naikawadi, R.P., Wolters, P.J., Abate, A.R., et al. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. Nat Immunol 20, 163–172.

Arazi, A., Rao, D.A., Berthier, C.C., Davidson, A., Liu, Y., Hoover, P.J., Chicoine, A., Eisenhaure, T.M., Jonsson, A.H., Li, S., et al. (2019). The immune cell landscape in kidneys of patients with lupus nephritis. Nat Immunol 20, 902–914.

Arendt, D., Musser, J.M., Baker, C.V.H., Bergman, A., Cepko, C., Erwin, D.H., Pavlicev, M., Schlosser, G., Widder, S., Laubichler, M.D., et al. (2016). The origin and evolution of cell types. Nat Rev Genet 17, 744–757.

Argelaguet, R., Cuomo, A.S.E., Stegle, O., and Marioni, J.C. (2021). Computational principles and challenges in single-cell data integration. Nat Biotechnol 39, 1202–1215.

Arisdakessian, C., Poirion, O., Yunits, B., Zhu, X., and Garmire, L.X. (2019). DeepImpute: An accurate, fast, and scalable deep neural network method to impute single-cell RNA-seq data. Genome Biol 20, 211.

Arrastia, M.V., Jachowicz, J.W., Ollikainen, N., Curtis, M.S., Lai, C., Quinodoz, S.A., Selck, D.A., Ismagilov, R.F., and Guttman, M. (2022). Single-cell measurement of higher-order 3D genome organization with scSPRITE. Nat Biotechnol 40, 64–73.

Arutyunyan, A., Roberts, K., Troulé, K., Wong, F.C.K., Sheridan, M.A., Kats, I., Garcia-Alonso, L., Velten, B., Hoo, R., Ruiz-Morales, E.R., et al. (2023). Spatial multiomics map of trophoblast development in early pregnancy. Nature 616, 143–151.

Ashuach, T., Gabitto, M.I., Koodli, R.V., Saldi, G.A., Jordan, M.I., and Yosef, N. (2023). MultiVI: Deep generative model for the integration of multimodal data. Nat Methods 20, 1222–1231.

Asp, M., Giacomello, S., Larsson, L., Wu, C., Fürth, D., Qian, X., Wärdell, E., Custodio, J., Reimegård, J., Salmén, F., et al. (2019). A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. Cell 179, 1647–1660. e19.

Athaya, T., Ripan, R.C., Li, X., and Hu, H. (2023). Multimodal deep learning approaches for single-cell multi-omics data integration. Brief Bioinf 24, bbad313.
Avsec, Ž., Agarwal, V., Visentin, D., Ledsam, J.R., Grabska-Barwinska, A., Taylor, K. R., Assael, Y., Jumper, J., Kohli, P., and Kelley, D.R. (2021). Effective gene expression prediction from sequence by integrating long-range interactions. Nat Methods 18, 1196–1203.
Axelrod, M.L., Meijers, W.C., Screever, E.M., Qin, J., Carroll, M.G., Sun, X., Tannous, E., Zhang, Y., Sugiura, A., Taylor, B.C., et al. (2022). T cells specific for α-myosin drive immunotherapy-related myocarditis. Nature 611, 818–826.
Bagnoli, J.W., Ziegenhain, C., Janjic, A., Wange, L.E., Vieth, B., Parekh, S., Geuder, J., Hellmann, I., and Enard, W. (2018). Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. Nat Commun 9, 2937.
Bai, Z., Deng, Y., Kim, D., Chen, Z., Xiao, Y., and Fan, R. (2020). An integrated dielectrophoresis-trapping and nanowell transfer approach to enable double-sub-poisson single-cell RNA sequencing. ACS Nano 14, 7412–7424.
Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res 21, 381–395.
Barkas, N., Petukhov, V., Nikolaeva, D., Lozinsky, Y., Demharter, S., Khodosevich, K., and Kharchenko, P.V. (2019). Joint analysis of heterogeneous single-cell RNA-seq dataset collections. Nat Methods 16, 695–698.
Barkley, D., Moncada, R., Pour, M., Liberman, D.A., Dryg, I., Werba, G., Wang, W., Baron, M., Rao, A., Xia, B., et al. (2022). Cancer cell states recur across tumor types and form specific interactions with the tumor microenvironment. Nat Genet 54, 1192–1201.
Baron, M., Veres, A., Wolock, S.L., Faust, A.L., Gaujoux, R., Vetere, A., Ryu, J.H., Wagner, B.K., Shen-Orr, S.S., Klein, A.M., et al. (2016). A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. Cell Syst 3, 346–360.e4.
Bartlett, D.A., Dileep, V., Handa, T., Ohkawa, Y., Kimura, H., Henikoff, S., and Gilbert, D.M. (2021). High-throughput single-cell epigenomic profiling by targeted insertion of promoters (TIP-seq). J Cell Biol 220, e202103078.
Bartosovic, M., Kabbe, M., and Castelo-Branco, G. (2021). Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. Nat Biotechnol 39, 825–835.
Bassez, A., Vos, H., Van Dyck, L., Floris, G., Arijs, I., Desmedt, C., Boeckx, B., Vanden Bempt, M., Nevelsteen, I., Lambein, K., et al. (2021). A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. Nat Med 27, 820–832.
Bawa, G., Liu, Z., Yu, X., Qin, A., and Sun, X. (2022). Single-cell RNA sequencing for plant research: insights and possible benefits. Int J Mol Sci 23, 4497.
Baysoy, A., Bai, Z., Satija, R., and Fan, R. (2023). The technological landscape and applications of single-cell multi-omics. Nat Rev Mol Cell Biol 24, 695–713.
Bhattarai-Kline, S., Lear, S.K., Fishman, C.B., Lopez, S.C., Lockshin, E.R., Schubert, M. G., Nivala, J., Church, G.M., and Shipman, S.L. (2022). Recording gene expression order in DNA by CRISPR addition of retron barcodes. Nature 608, 217–225.
Bian, H., Chen, Y., Dong, X., Li, C., Hao, M., Chen, S., Hu, J., Sun, M., Wei, L., and Zhang, X. (2024). scMulan: a multitask generative pre-trained language model for single-cell analysis. In: Ma, J., ed. Research in Computational Molecular Biology. RECOMB 2024. Lecture Notes in Computer Science. Cham: Springer. 479–482.
Bian, S., Hou, Y., Zhou, X., Li, X., Yong, J., Wang, Y., Wang, W., Yan, J., Hu, B., Guo, H., et al. (2018). Single-cell multiomics sequencing and analyses of human colorectal cancer. Science 362, 1060–1063.
Biancalani, T., Scalia, G., Buffoni, L., Avasthi, R., Lu, Z., Sanger, A., Tokcan, N., Vanderburg, C.R., Segerstolpe, Å., Zhang, M., et al. (2021). Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. Nat Methods 18, 1352–1362.
Bianchi, A., Scherer, M., Zaurin, R., Quililan, K., Velten, L., and Beekman, R. (2022). scTAM-seq enables targeted high-confidence analysis of DNA methylation in single cells. Genome Biol 23, 229.
Blank, C.U., Haining, W.N., Held, W., Hogan, P.G., Kallies, A., Lugli, E., Lynn, R.C., Philip, M., Rao, A., Restifo, N.P., et al. (2019). Defining 'T cell exhaustion'. Nat Rev Immunol 19, 665–674.
Blattman, S.B., Jiang, W., Oikonomou, P., and Tavazoie, S. (2020). Prokaryotic single-cell RNA sequencing by *in situ* combinatorial indexing. Nat Microbiol 5, 1192–1201.
Bock, C., Datlinger, P., Chardon, F., Coelho, M.A., Dong, M.B., Lawson, K.A., Lu, T., Maroc, L., Norman, T.M., Song, B., et al. (2022). High-content CRISPR screening. Nat Rev Methods Primers 2, 8.
Borgsmüller, N., Bonet, J., Marass, F., Gonzalez-Perez, A., Lopez-Bigas, N., Beerenwinkel, N., and Mathelier, A. (2020). BnpC: Bayesian non-parametric clustering of single-cell mutation profiles. Bioinformatics 36, 4854–4859.
Borm, L.E., Mossi Albiach, A., Mannens, C.C.A., Janusauskas, J., Özgün, C., Fernández-García, D., Hodge, R., Castillo, F., Hedin, C.R.H., Villablanca, E.J., et al. (2023). Scalable *in situ* single-cell profiling by electrophoretic capture of mRNA using EEL FISH. Nat Biotechnol 41, 222–231.
BRAIN Initiative Cell Census Network (BICCN) (2021). A multimodal cell census and atlas of the mammalian primary motor cortex. Nature 598, 86–102.
Braun, E., Danan-Gotthold, M., Borm, L.E., Lee, K.W., Vinsland, E., Lönnerberg, P., Hu, L., Li, X., He, X., Andrusivová, Ž., et al. (2023). Comprehensive cell atlas of the first-trimester developing human brain. Science 382, eadf1226.
Bravo González-Blas, C., Minnoye, L., Papasokrati, D., Aibar, S., Hulselmans, G., Christiaens, V., Davie, K., Wouters, J., and Aerts, S. (2019). cisTopic: *cis*-regulatory topic modeling on single-cell ATAC-seq data. Nat Methods 16, 397–400.
Brendel, M., Su, C., Bai, Z., Zhang, H., Elemento, O., and Wang, F. (2022). Application of deep learning on single-cell RNA sequencing data analysis: a review. Genomics Proteomics Bioinf 20, 814–835.
Bressan, D., Battistoni, G., and Hannon, G.J. (2023). The dawn of spatial omics. Science 381, eabq4964.
Briggs, J.A., Weinreb, C., Wagner, D.E., Megason, S., Peshkin, L., Kirschner, M.W., and Klein, A.M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. Science 360, eaar5780.
Browaeys, R., Saelens, W., and Saeys, Y. (2020). NicheNet: modeling intercellular communication by linking ligands to target genes. Nat Methods 17, 159–162.
Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., and Greenleaf, W.J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523, 486–490.
Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 36, 411–420.
Büttner, M., Miao, Z., Wolf, F.A., Teichmann, S.A., and Theis, F.J. (2019). A test metric for assessing single-cell RNA-seq batch correction. Nat Methods 16, 43–49.
Cabello-Aguilar, S., Alame, M., Kon-Sun-Tack, F., Fau, C., Lacroix, M., and Colinge, J. (2020). SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics. Nucleic Acids Res 48, e55.
Cao, J., Cusanovich, D.A., Ramani, V., Aghamirzaie, D., Pliner, H.A., Hill, A.J., Daza, R.M., McFaline-Figueroa, J.L., Packer, J.S., Christiansen, L., et al. (2018). Joint profiling of chromatin accessibility and gene expression in thousands of single cells. Science 361, 1380–1385.
Cao, J., O'Day, D.R., Pliner, H.A., Kingsley, P.D., Deng, M., Daza, R.M., Zager, M.A., Aldinger, K.A., Blecher-Gonen, R., Zhang, F., et al. (2020a). A human cell atlas of fetal gene expression. Science 370, eaba7721.
Cao, J., Zhou, W., Steemers, F., Trapnell, C., and Shendure, J. (2020b). Sci-fate characterizes the dynamics of gene expression in single cells. Nat Biotechnol 38, 980–988.
Cao, J., Packer, J.S., Ramani, V., Cusanovich, D.A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S.N., Steemers, F.J., et al. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. Science 357, 661–667.
Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature 566, 496–502.
Cao, Y., Bai, Y., Yuan, T., Song, L., Fan, Y., Ren, L., Song, W., Peng, J., An, R., Gu, Q., et al. (2023). Single-cell bisulfite-free 5mC and 5hmC sequencing with high sensitivity and scalability. Proc Natl Acad Sci USA 120, e2310367120.
Cao, Y., Su, B., Guo, X., Sun, W., Deng, Y., Bao, L., Zhu, Q., Zhang, X., Zheng, Y., Geng, C., et al. (2020c). Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of convalescent patients' B cells. Cell 182, 73–84.e16.
Cao, Z.J., and Gao, G. (2022). Multi-omics single-cell data integration and regulatory inference with graph-linked embedding. Nat Biotechnol 40, 1458–1466.
Cardozo Gizzi, A.M., Cattoni, D.I., Fiche, J.B., Espinola, S.M., Gurgo, J., Messina, O., Houbron, C., Ogiyama, Y., Papadopoulos, G.L., Cavalli, G., et al. (2019). Microscopy-based chromosome conformation capture enables simultaneous visualization of genome organization and transcription in intact organisms. Mol Cell 74, 212–222.e5.
Carter, B., Ku, W.L., Kang, J.Y., Hu, G., Perrie, J., Tang, Q., and Zhao, K. (2019). Mapping histone modifications in low cell number and single cells using antibody-guided chromatin tagmentation (ACT-seq). Nat Commun 10, 3747.
Cascone, T., Leung, C.H., Weissferdt, A., Pataer, A., Carter, B.W., Godoy, M.C.B., Feldman, H., William Jr, W.N., Xi, Y., Basu, S., et al. (2023). Neoadjuvant chemotherapy plus nivolumab with or without ipilimumab in operable non-small cell lung cancer: the phase 2 platform NEOSTAR trial. Nat Med 29, 593–604.
Chai, H., Tjong, H., Li, P., Liao, W., Wang, P., Wong, C.H., Ngan, C.Y., Leonard, W.J., Wei, C.L., and Ruan, Y. (2023). ChIATAC is an efficient strategy for multi-omics mapping of 3D epigenomes from low-cell inputs. Nat Commun 14, 213.
Chamorro González, R., Conrad, T., Stöber, M.C., Xu, R., Giurgiu, M., Rodriguez-Fos, E., Kasack, K., Brückner, L., van Leen, E., Helmsauer, K., et al. (2023). Parallel sequencing of extrachromosomal circular DNAs and transcriptomes in single cancer cells. Nat Genet 55, 880–890.

Chang, L., Deng, E., Wang, J., Zhou, W., Ao, J., Liu, R., Su, D., and Fan, X. (2023a). Single-cell third-generation sequencing-based multi-omics uncovers gene expression changes governed by ecDNA and structural variants in cancer cells. Clin Transl Med 13, e1351.
Chang, L., Ruiz, P., Ito, T., and Sellers, W.R. (2021). Targeting pan-essential genes in cancer: challenges and opportunities. Cancer Cell 39, 466–479.
Chang, T., Han, W., Jiang, M., Li, J., Liao, Z., Tang, M., Zhang, J., Shen, J., Chen, Z., Fei, P., et al. (2023b). Rapid and signal crowdedness-robust *in situ* sequencing through hybrid block coding. Proc Natl Acad Sci USA 120, e2309227120.
Charlton, J., Downing, T.L., Smith, Z.D., Gu, H., Clement, K., Pop, R., Akopian, V., Klages, S., Santos, D.P., Tsankov, A.M., et al. (2018). Author Correction: Global delay in nascent strand DNA methylation. Nat Struct Mol Biol 25, 355.
Chartrand, T., Dalley, R., Close, J., Goriounova, N.A., Lee, B.R., Mann, R., Miller, J.A., Molnar, G., Mukora, A., Alfiler, L., et al. (2023). Morphoelectric and transcriptomic divergence of the layer 1 interneuron repertoire in human versus mouse neocortex. Science 382, eadf0805.
Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Qiu, X., Yang, J., Xu, J., Hao, S., et al. (2022a). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. Cell 185, 1777–1792.e21.
Chen, A., Sun, Y., Lei, Y., Li, C., Liao, S., Meng, J., Bai, Y., Liu, Z., Liang, Z., Zhu, Z., et al. (2023a). Single-cell spatial transcriptome reveals cell-type organization in the macaque cortex. Cell 186, 3726–3743.e24.
Chen, C., Xing, D., Tan, L., Li, H., Zhou, G., Huang, L., and Xie, X.S. (2017a). Single-cell whole-genome analyses by Linear Amplification via Transposon Insertion (LIANTI). Science 356, 189–194.
Chen, H., Fang, X., Shao, J., Zhang, Q., Xu, L., Chen, J., Mei, Y., Jiang, M., Wang, Y., Li, Z., et al. (2024a). Pan-cancer single-nucleus total RNA sequencing using snHH-Seq. Adv Sci 11, e2304755.
Chen, H., Liao, Y., Zhang, G., Sun, Z., Yang, L., Fang, X., Sun, H., Ma, L., Fu, Y., Li, J., et al. (2021a). High-throughput Microwell-seq 2.0 profiles massively multiplexed chemical perturbation. Cell Discov 7, 107.
Chen, J., Suo, S., Tam, P.P., Han, J.D.J., Peng, G., and Jing, N. (2017b). Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq. Nat Protoc 12, 566–580.
Chen, J., Xu, H., Tao, W., Chen, Z., Zhao, Y., and Han, J.D.J. (2023b). Transformer for one stop interpretable cell type annotation. Nat Commun 14, 223.
Chen, K., and Wang, Z. (2023). A micropillar array based microfluidic device for rare cell detection and single-cell proteomics. Methods Protoc 6, 80.
Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S., and Zhuang, X. (2015). Spatially resolved, highly multiplexed RNA profiling in single cells. Science 348, aaa6090.
Chen, L., and Flies, D.B. (2013). Molecular mechanisms of T cell co-stimulation and co-inhibition. Nat Rev Immunol 13, 227–242.
Chen, S., Lake, B.B., and Zhang, K. (2019). High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. Nat Biotechnol 37, 1452–1457.
Chen, W., Guillaume-Gentil, O., Rainer, P.Y., Gäbelein, C.G., Saelens, W., Gardeux, V., Klaeger, A., Dainese, R., Zachara, M., Zambelli, T., et al. (2022b). Live-seq enables temporal transcriptomic recording of single cells. Nature 608, 733–740.
Chen, W., Zhao, Y., Chen, X., Yang, Z., Xu, X., Bi, Y., Chen, V., Li, J., Choi, H., Ernest, B., et al. (2021b). A multicenter study benchmarking single-cell RNA sequencing technologies using reference samples. Nat Biotechnol 39, 1103–1114.
Chen, X., Miragaia, R.J., Natarajan, K.N., and Teichmann, S.A. (2018). A rapid and robust method for single cell chromatin accessibility profiling. Nat Commun 9, 5345.
Chen, X., Wang, T., Chen, L., Zhao, Y., Deng, Y., Shen, W., Li, L., Yin, Z., Zhang, C., Cai, G., et al. (2024b). Cross-species single-cell analysis uncovers the immunopathological mechanisms associated with IgA nephropathy progression. JCI Insight 9, e173651.
Chen, Y., Wang, Y., Chen, Y., Cheng, Y., Wei, Y., Li, Y., Wang, J., Wei, Y., Chan, T.F., and Li, Y. (2022c). Deep autoencoder for interpretable tissue-adaptive deconvolution and cell-type-specific gene analysis. Nat Commun 13, 6735.
Chen, Z., Gong, F., Wan, L., and Ma, L. (2022d). *BiTSC2*: Bayesian inference of tumor clonal tree by joint analysis of single-cell SNV and CNA data. Brief Bioinf 23, bbac092.
Chen, Z., King, W.C., Hwang, A., Gerstein, M., and Zhang, J. (2022e). *DeepVelo*: single-cell transcriptomic deep velocity field learning with neural ordinary differential equations. Sci Adv 8, eabq3745.
Cheng, H., Pui, H.P., Lentini, A., Kolbeinsdóttir, S., Andrews, N., Pei, Y., Reinius, B., Deng, Q., and Enge, M. (2021a). Smart3-ATAC: a highly sensitive method for joint accessibility and full-length transcriptome analysis in single cells. bioRxiv, 2021.2012.2002.470912.
Cheng, M., Jiang, Y., Xu, J., Mentis, A.F.A., Wang, S., Zheng, H., Sahu, S.K., Liu, L., and Xu, X. (2023a). Spatially resolved transcriptomics: a comprehensive review of their technological advances, applications, and challenges. J Genet Genomics 50, 625–640.
Cheng, S., Li, Z., Gao, R., Xing, B., Gao, Y., Yang, Y., Qin, S., Zhang, L., Ouyang, H., Du, P., et al. (2021b). A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. Cell 184, 792–809.e23.
Cheng, Y., Ma, X., Yuan, L., Sun, Z., and Wang, P. (2023b). Evaluating imputation methods for single-cell RNA-seq data. BMC Bioinf 24, 302.
Chiou, J., Geusz, R.J., Okino, M.L., Han, J.Y., Miller, M., Melton, R., Beebe, E., Benaglio, P., Huang, S., Korgaonkar, K., et al. (2021). Interpreting type 1 diabetes risk with genetics and single-cell epigenomics. Nature 594, 398–402.
Cho, C.S., Xi, J., Si, Y., Park, S.R., Hsu, J.E., Kim, M., Jun, G., Kang, H.M., and Lee, J.H. (2021). Microscopic examination of spatial transcriptome using Seq-Scope. Cell 184, 3559–3572.e22.
Chung, H., Melnikov, A., McCabe, C., Drokhlyansky, E., Wittenberghe, N.V., Magee, E. M., Waldman, J., Spira, A., Chen, F., Mazzilli, S., et al. (2022). SnFFPE-Seq: towards scalable single nucleus RNA-Seq of formalin-fixed paraffin-embedded (FFPE) tissue. bioRxiv, 2022.2008.2025.505257.
Chung, H., Parkhurst, C.N., Magee, E.M., Phillips, D., Habibi, E., Chen, F., Yeung, B.Z., Waldman, J., Artis, D., and Regev, A. (2021). Joint single-cell measurements of nuclear proteins and RNA *in vivo*. Nat Methods 18, 1204–1212.
Claringbould, A., and Zaugg, J.B. (2021). Enhancers in disease: molecular basis and emerging treatment strategies. Trends Mol Med 27, 1060–1073.
Clark, I.C., Fontanez, K.M., Meltzer, R.H., Xue, Y., Hayford, C., May-Zhang, A., D'Amato, C., Osman, A., Zhang, J.Q., Hettige, P., et al. (2023). Microfluidics-free single-cell genomics with templated emulsification. Nat Biotechnol 41, 1557–1566.
Clark, S.J., Argelaguet, R., Kapourani, C.A., Stubbs, T.M., Lee, H.J., Alda-Catalinas, C., Krueger, F., Sanguinetti, G., Kelsey, G., Marioni, J.C., et al. (2018). scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. Nat Commun 9, 781.
Codeluppi, S., Borm, L.E., Zeisel, A., La Manno, G., van Lunteren, J.A., Svensson, C.I., and Linnarsson, S. (2018). Spatial organization of the somatosensory cortex revealed by osmFISH. Nat Methods 15, 932–935.
Crow, M., Paul, A., Ballouz, S., Huang, Z.J., and Gillis, J. (2018). Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. Nat Commun 9, 884.
Cui, H., Wang, C., Maan, H., Pang, K., Luo, F., Duan, N., and Wang, B. (2024). scGPT: toward building a foundation model for single-cell multi-omics using generative AI. Nat Methods 21, 1470–1480.
Cuomo, A.S.E., Nathan, A., Raychaudhuri, S., MacArthur, D.G., and Powell, J.E. (2023). Single-cell genomics meets human genetics. Nat Rev Genet 24, 535–549.
Cusanovich, D.A., Hill, A.J., Aghamirzaie, D., Daza, R.M., Pliner, H.A., Berletch, J.B., Filippova, G.N., Huang, X., Christiansen, L., DeWitt, W.S., et al. (2018). A single-cell atlas of *in vivo* mammalian chromatin accessibility. Cell 174, 1309–1324.e18.
Dar, D., Dar, N., Cai, L., and Newman, D.K. (2021). Spatial transcriptomics of planktonic and sessile bacterial populations at single-cell resolution. Science 373, eabi4882.
Datlinger, P., Rendeiro, A.F., Boenke, T., Senekowitsch, M., Krausgruber, T., Barreca, D., and Bock, C. (2021). Ultra-high-throughput single-cell RNA sequencing and perturbation screening with combinatorial fluidic indexing. Nat Methods 18, 635–642.
Datlinger, P., Rendeiro, A.F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J., Schuster, L.C., Kuchler, A., Alpar, D., and Bock, C. (2017). Pooled CRISPR screening with single-cell transcriptome readout. Nat Methods 14, 297–301.
Davie, K., Janssens, J., Koldere, D., De Waegeneer, M., Pech, U., Kreft, Ł., Aibar, S., Makhzami, S., Christiaens, V., Bravo González-Blas, C., et al. (2018). A single-cell transcriptome atlas of the aging *Drosophila* brain. Cell 174, 982–998.e20.
de la Fuente, J., Legarra, N., Serrano, G., Osta, A.G., Kalari, K.R., Fernandez-Granda, C., Ochoa, I., and Hernaez, M. (2023). Sweetwater: an interpretable and adaptive autoencoder for efficient tissue deconvolution. arXiv, 231111991.
de Visser, K.E., and Joyce, J.A. (2023). The evolving tumor microenvironment: from cancer initiation to metastatic outgrowth. Cancer Cell 41, 374–403.
Dekker, J., Belmont, A.S., Guttman, M., Leshyk, V.O., Lis, J.T., Lomvardas, S., Mirny, L. A., O'Shea, C.C., Park, P.J., Ren, B., et al. (2017). The 4D nucleome project. Nature 549, 219–226.
Deng, C.C., Hu, Y.F., Zhu, D.H., Cheng, Q., Gu, J.J., Feng, Q.L., Zhang, L.X., Xu, Y.P., Wang, D., Rong, Z., et al. (2021). Single-cell RNA-seq reveals fibroblast heterogeneity and increased mesenchymal fibroblasts in human fibrotic skin diseases. Nat Commun 12, 3709.
Deng, Y., Bao, F., Dai, Q., Wu, L.F., and Altschuler, S.J. (2019). Scalable analysis of cell-type composition from single-cell transcriptomics using deep recurrent learning. Nat Methods 16, 311–314.
Deng, Y., Bartosovic, M., Kukanja, P., Zhang, D., Liu, Y., Su, G., Enninful, A., Bai, Z., Castelo-Branco, G., and Fan, R. (2022a). Spatial-CUT&Tag: spatially resolved chromatin modification profiling at the cellular level. Science 375, 681–686.

Deng, Y., Bartosovic, M., Ma, S., Zhang, D., Kukanja, P., Xiao, Y., Su, G., Liu, Y., Qin, X., Rosoklija, G.B., et al. (2022b). Spatial profiling of chromatin accessibility in mouse and human tissues. Nature 609, 375–383.
Denisenko, E., de Kock, L., Tan, A., Beasley, A.B., Beilin, M., Jones, M.E., Hou, R., Muirí, D.Ó., Bilic, S., Mohan, G.R.K.A., et al. (2024). Spatial transcriptomics reveals discrete tumour microenvironments and autocrine loops within ovarian cancer subclones. Nat Commun 15, 2860.
Dey, S.S., Kester, L., Spanjaard, B., Bienko, M., and van Oudenaarden, A. (2015). Integrated genome and transcriptome sequencing of the same cell. Nat Biotechnol 33, 285–289.
Dixit, A., Parnas, O., Li, B., Chen, J., Fulco, C.P., Jerby-Arnon, L., Marjanovic, N.D., Dionne, D., Burks, T., Raychowdhury, R., et al. (2016). Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. Cell 167, 1853–1866.e17.
Domcke, S., Hill, A.J., Daza, R.M., Cao, J., O'Day, D.R., Pliner, H.A., Aldinger, K.A., Pokholok, D., Zhang, F., Milbank, J.H., et al. (2020). A human cell atlas of fetal chromatin accessibility. Science 370, 809.
Domínguez Conde, C., Xu, C., Jarvis, L.B., Rainbow, D.B., Wells, S.B., Gomes, T., Howlett, S.K., Suchanek, O., Polanski, K., King, H.W., et al. (2022). Cross-tissue immune cell analysis reveals tissue-specific features in humans. Science 376, eabl5197.
Dong, K., and Zhang, S. (2022). Deciphering spatial domains from spatially resolved transcriptomics with an adaptive graph attention auto-encoder. Nat Commun 13, 1739.
Dong, M., Wang, B., Wei, J., de O. Fonseca, A.H., Perry, C.J., Frey, A., Ouerghi, F., Foxman, E.F., Ishizuka, J.J., Dhodapkar, R.M., et al. (2023). Causal identification of single-cell experimental perturbation effects with CINEMA-OT. Nat Methods 20, 1769–1779.
Du, J., Wang, Y., Chen, W., Xu, M., Zhou, R., Shou, H., and Chen, J. (2023). High-resolution anatomical and spatial transcriptome analyses reveal two types of meristematic cell pools within the secondary vascular tissue of poplar stem. Mol Plant 16, 809–828.
Efremova, M., Vento-Tormo, M., Teichmann, S.A., and Vento-Tormo, R. (2020). CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc 15, 1484–1506.
Eisenstein, M. (2020). Single-cell RNA-seq analysis software providers scramble to offer solutions. Nat Biotechnol 38, 254–257.
Elmentaite, R., Domínguez Conde, C., Yang, L., and Teichmann, S.A. (2022). Single-cell atlases: Shared and tissue-specific cell types across human organs. Nat Rev Genet 23, 395–410.
Emmert-Buck, M.R., Bonner, R.F., Smith, P.D., Chuaqui, R.F., Zhuang, Z., Goldstein, S.R., Weiss, R.A., and Liotta, L.A. (1996). Laser capture microdissection. Science 274, 998–1001.
Eng, C.H.L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.C., et al. (2019). Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. Nature 568, 235–239.
Engblom, C., Pfirschke, C., and Pittet, M.J. (2016). The role of myeloid cells in cancer therapies. Nat Rev Cancer 16, 447–462.
Eraslan, G., Drokhlyansky, E., Anand, S., Fiskin, E., Subramanian, A., Slyper, M., Wang, J., Van Wittenberghe, N., Rouhana, J.M., Waldman, J., et al. (2022). Single-nucleus cross-tissue molecular reference maps toward understanding disease gene function. Science 376, eabl4290.
Eraslan, G., Simon, L.M., Mircea, M., Mueller, N.S., and Theis, F.J. (2019). Single-cell RNA-seq denoising using a deep count autoencoder. Nat Commun 10, 390.
Erfanian, N., Heydari, A.A., Feriz, A.M., Iañez, P., Derakhshani, A., Ghasemigol, M., Farahpour, M., Razavi, S.M., Nasseri, S., Safarpour, H., et al. (2023). Deep learning applications in single-cell genomics and transcriptomics data analysis. Biomed Pharmacother 165, 115077.
Erhard, F., Baptista, M.A.P., Krammer, T., Hennig, T., Lange, M., Arampatzi, P., Jürges, C.S., Theis, F.J., Saliba, A.E., and Dölken, L. (2019). scSLAM-seq reveals core features of transcription dynamics in single cells. Nature 571, 419–423.
Erickson, A., He, M., Berglund, E., Marklund, M., Mirzazadeh, R., Schultz, N., Kvastad, L., Andersson, A., Bergenstråhle, L., Bergenstråhle, J., et al. (2022). Spatially resolved clonal copy number alterations in benign and malignant tissue. Nature 608, 360–367.
Falcão, A.M., van Bruggen, D., Marques, S., Meijer, M., Jäkel, S., Agirre, E., Samudyata, E., Floriddia, E.M., Vanichkina, D.P., ffrench-Constant, C., et al. (2018). Disease-specific oligodendrocyte lineage cells arise in multiple sclerosis. Nat Med 24, 1837–1844.
Fan, H.C., Fu, G.K., and Fodor, S.P.A. (2015a). Combinatorial labeling of single cells for gene expression cytometry. Science 347, 1258367.
Fan, X., Zhang, X., Wu, X., Guo, H., Hu, Y., Tang, F., and Huang, Y. (2015b). Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. Genome Biol 16, 148.
Fang, L., Chen, Q., Wei, C.H., Lu, Z., and Wang, K. (2023). Bioformer: an efficient transformer language model for biomedical text mining. arXiv, 2302.01588.
Fang, R., Preissl, S., Li, Y., Hou, X., Lucero, J., Wang, X., Motamedi, A., Shiau, A.K., Zhou, X., Xie, F., et al. (2021). Comprehensive analysis of single cell ATAC-seq data with SnapATAC. Nat Commun 12, 1337.
Fang, R., Xia, C., Close, J.L., Zhang, M., He, J., Huang, Z., Halpern, A.R., Long, B., Miller, J.A., Lein, E.S., et al. (2022a). Conservation and divergence of cortical cell organization in human and mouse revealed by MERFISH. Science 377, 56–62.
Fang, X., Jiang, M., Zhou, M., Shao, J., Fang, X., Wang, J., Fu, Y., Liao, Y., Chen, H., Wang, R., et al. (2022b). Elucidating the developmental dynamics of mouse stromal cells at single-cell level. Life Med 1, 45–48.
Fawkner-Corbett, D., Antanaviciute, A., Parikh, K., Jagielowicz, M., Gerós, A.S., Gupta, T., Ashley, N., Khamis, D., Fowler, D., Morrissey, E., et al. (2021). Spatiotemporal analysis of human intestinal development at single-cell resolution. Cell 184, 810–826.e23.
Fei, L., Chen, H., Ma, L., E, W., Wang, R., Fang, X., Zhou, Z., Sun, H., Wang, J., Jiang, M., et al. (2022). Systematic identification of cell-fate regulatory programs using a single-cell atlas of mouse development. Nat Genet 54, 1051–1061.
Femino, A.M., Fay, F.S., Fogarty, K., and Singer, R.H. (1998). Visualization of single RNA transcripts *in situ*. Science 280, 585–590.
Fiskin, E., Lareau, C.A., Ludwig, L.S., Eraslan, G., Liu, F., Ring, A.M., Xavier, R.J., and Regev, A. (2022). Single-cell profiling of proteins and chromatin accessibility using PHAGE-ATAC. Nat Biotechnol 40, 374–381.
Flyamer, I.M., Gassler, J., Imakaev, M., Brandão, H.B., Ulianov, S.V., Abdennur, N., Razin, S.V., Mirny, L.A., and Tachibana-Konwalski, K. (2017). Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature 544, 110–114.
Frangoul, H., Altshuler, D., Cappellini, M.D., Chen, Y.S., Domm, J., Eustace, B.K., Foell, J., de la Fuente, J., Grupp, S., Handgretinger, R., et al. (2021). CRISPR-Cas9 gene editing for sickle cell disease and β-thalassemia. N Engl J Med 384, 252–260.
Frank, M.H., Wilson, B.J., Gold, J.S., and Frank, N.Y. (2021). Clinical implications of colorectal cancer stem cells in the age of single-cell omics and targeted therapies. Gastroenterology 160, 1947–1960.
Fu, X., Sun, L., Dong, R., Chen, J.Y., Silakit, R., Condon, L.F., Lin, Y., Lin, S., Palmiter, R.D., and Gu, L. (2022). Polony gels enable amplifiable DNA stamping and spatial transcriptomics of chronic pain. Cell 185, 4621–4633.e17.
Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H., et al. (2009). An oestrogen-receptor-α-bound human chromatin interactome. Nature 462, 58–64.
Funnell, T., O'Flanagan, C.H., Williams, M.J., McPherson, A., McKinney, S., Kabeer, F., Lee, H., Salehi, S., Vázquez-García, I., Shi, H., et al. (2022). Single-cell genomic variation induced by mutational processes in cancer. Nature 612, 106–115.
Galluzzi, L., Buqué, A., Kepp, O., Zitvogel, L., and Kroemer, G. (2015). Immunological effects of conventional chemotherapy and targeted anticancer agents. Cancer Cell 28, 690–714.
Gangadhar, T.C., and Vonderheide, R.H. (2014). Mitigating the toxic effects of anticancer immunotherapy. Nat Rev Clin Oncol 11, 91–99.
Gao, W., Ku, W.L., Pan, L., Perrie, J., Zhao, T., Hu, G., Wu, Y., Zhu, J., Ni, B., and Zhao, K. (2021). Multiplex indexing approach for the detection of DNase I hypersensitive sites in single cells. Nucleic Acids Res 49, e56.
Gao, Y., Yi, L., Zhan, J., Wang, L., Yao, X., Yan, J., Jian, S., Gao, L., Farangez, M., Gao, M., et al. (2023). A clinical study of preimplantation DNA methylation screening in assisted reproductive technology. Cell Res 33, 483–485.
Garcia, F.J., Sun, N., Lee, H., Godlewski, B., Mathys, H., Galani, K., Zhou, B., Jiang, X., Ng, A.P., Mantero, J., et al. (2022). Single-cell dissection of the human brain vasculature. Nature 603, 893–899.
Gavish, A., Tyler, M., Greenwald, A.C., Hoefflin, R., Simkin, D., Tschernichovsky, R., Galili Darnell, N., Somech, E., Barbolin, C., Antman, T., et al. (2023). Hallmarks of transcriptional intratumour heterogeneity across a thousand tumours. Nature 618, 598–606.
Gehring, J., Hwee Park, J., Chen, S., Thomson, M., and Pachter, L. (2020). Highly multiplexed single-cell RNA-seq by DNA oligonucleotide tagging of cellular proteins. Nat Biotechnol 38, 35–38.
Gerber, T., Murawala, P., Knapp, D., Masselink, W., Schuez, M., Hermann, S., Gac-Santel, M., Nowoshilow, S., Kageyama, J., Khattak, S., et al. (2018). Single-cell analysis uncovers convergence of cell identities during axolotl limb regeneration. Science 362, eaaq0681.
Gerlach, J.P., van Buggenum, J.A.G., Tanis, S.E.J., Hogeweg, M., Heuts, B.M.H., Muraro, M.J., Elze, L., Rivello, F., Rakszewska, A., van Oudenaarden, A., et al. (2019). Combined quantification of intracellular (phospho-)proteins and transcriptomics from fixed single cells. Sci Rep 9, 1469.
Gerlinger, M., Rowan, A.J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., et al. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J

Med 366, 883–892.
Gerstberger, S., Jiang, Q., and Ganesh, K. (2023). Metastasis. Cell 186, 1564–1579.
Ghazanfar, S., Guibentif, C., and Marioni, J.C. (2024). Stabilized mosaic single-cell data integration using unshared features. Nat Biotechnol 42, 284–292.
Giannoukos, G., Ciulla, D.M., Huang, K., Haas, B.J., Izard, J., Levin, J.Z., Livny, J., Earl, A.M., Gevers, D., Ward, D.V., et al. (2012). Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes. Genome Biol 13, R23.
Gierahn, T.M., Wadsworth Ii, M.H., Hughes, T.K., Bryson, B.D., Butler, A., Satija, R., Fortune, S., Love, J.C., and Shalek, A.K. (2017). Erratum: Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. Nat Methods 14, 752.
Giolai, M., Verweij, W., Lister, A., Heavens, D., Macaulay, I., and Clark, M.D. (2019). Spatially resolved transcriptomics reveals plant host responses to pathogens. Plant Methods 15, 114.
Goldstein, L.D., Chen, Y.J.J., Wu, J., Chaudhuri, S., Hsiao, Y.C., Schneider, K., Hoi, K. H., Lin, Z., Guerrero, S., Jaiswal, B.S., et al. (2019). Massively parallel single-cell B-cell receptor sequencing enables rapid discovery of diverse antigen-reactive antibodies. Commun Biol 2, 304.
Gong, B., Zhou, Y., and Purdom, E. (2021). Cobolt: Integrative analysis of multimodal single-cell sequencing data. Genome Biol 22, 351.
Gong, W., Kwak, I.Y., Pota, P., Koyano-Nakagawa, N., and Garry, D.J. (2018). DrImpute: Imputing dropout events in single cell RNA sequencing data. BMC Bioinf 19, 220.
Gopalan, S., Wang, Y., Harper, N.W., Garber, M., and Fazzio, T.G. (2021). Simultaneous profiling of multiple chromatin proteins in the same cells. Mol Cell 81, 4736–4746.e5.
Gorkin, D.U., Barozzi, I., Zhao, Y., Zhang, Y., Huang, H., Lee, A.Y., Li, B., Chiou, J., Wildberg, A., Ding, B., et al. (2020). An atlas of dynamic chromatin landscapes in mouse fetal development. Nature 583, 744–751.
Granja, J.M., Corces, M.R., Pierce, S.E., Bagdatli, S.T., Choudhry, H., Chang, H.Y., and Greenleaf, W.J. (2021). Author Correction: ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. Nat Genet 53, 935.
Greenbaum, S., Averbukh, I., Soon, E., Rizzuto, G., Baranski, A., Greenwald, N.F., Kagel, A., Bosse, M., Jaswa, E.G., Khair, Z., et al. (2023). A spatially resolved timeline of the human maternal-fetal interface. Nature 619, 595–605.
Greenman, C., Stephens, P., Smith, R., Dalgliesh, G.L., Hunter, C., Bignell, G., Davies, H., Teague, J., Butler, A., Stevens, C., et al. (2007). Patterns of somatic mutation in human cancer genomes. Nature 446, 153–158.
Gross, A., Schoendube, J., Zimmermann, S., Steeb, M., Zengerle, R., and Koltay, P. (2015). Technologies for single-cell isolation. Int J Mol Sci 16, 16897–16919.
Grosselin, K., Durand, A., Marsolier, J., Poitou, A., Marangoni, E., Nemati, F., Dahmani, A., Lameiras, S., Reyal, F., Frenoy, O., et al. (2019). High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. Nat Genet 51, 1060–1066.
Gu, H., Raman, A.T., Wang, X., Gaiti, F., Chaligne, R., Mohammad, A.W., Arczewska, A., Smith, Z.D., Landau, D.A., Aryee, M.J., et al. (2021). Smart-RRBS for single-cell methylome and transcriptome analysis. Nat Protoc 16, 4004–4030.
Gu, W., Crawford, E.D., O'Donovan, B.D., Wilson, M.R., Chow, E.D., Retallack, H., and DeRisi, J.L. (2016). Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. Genome Biol 17, 41.
Guinney, J., Dienstmann, R., Wang, X., de Reyniès, A., Schlicker, A., Soneson, C., Marisa, L., Roepman, P., Nyamundanda, G., Angelino, P., et al. (2015). The consensus molecular subtypes of colorectal cancer. Nat Med 21, 1350–1356.
Guo, C., Kong, W., Kamimoto, K., Rivera-Gonzalez, G.C., Yang, X., Kirita, Y., and Morris, S.A. (2019). CellTag Indexing: genetic barcode-based sample multiplexing for single-cell genomics. Genome Biol 20, 90.
Guo, F., Li, L., Li, J., Wu, X., Hu, B., Zhu, P., Wen, L., and Tang, F. (2017). Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. Cell Res 27, 967–988.
Guo, H., Zhu, P., Wu, X., Li, X., Wen, L., and Tang, F. (2013). Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. Genome Res 23, 2126–2135.
Guo, X., Xu, W., Zhang, W., Pan, C., Thalacker-Mercer, A.E., Zheng, H., and Gu, Z. (2023). High-frequency and functional mitochondrial DNA mutations at the single-cell level. Proc Natl Acad Sci USA 120, e2201518120.
Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., Kang, B., Liu, Z., Jin, L., Xing, R., et al. (2018). Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. Nat Med 24, 978–985.
Gyllborg, D., Langseth, C.M., Qian, X., Choi, E., Salas, S.M., Hilscher, M.M., Lein, E.S., and Nilsson, M. (2020). Hybridization-based *in situ* sequencing (HybISS) for spatially resolved transcriptomics in human and mouse brain tissue. Nucleic Acids Res 48, e112.
Haber, A.L., Biton, M., Rogel, N., Herbst, R.H., Shekhar, K., Smillie, C., Burgin, G., Delorey, T.M., Howitt, M.R., Katz, Y., et al. (2017). A single-cell survey of the small intestinal epithelium. Nature 551, 333–339.
Hagemann-Jensen, M., Ziegenhain, C., and Sandberg, R. (2022). Scalable single-cell RNA sequencing from full transcripts with Smart-seq3xpress. Nat Biotechnol 40, 1452–1457.
Haghverdi, L., Büttner, M., Wolf, F.A., Buettner, F., and Theis, F.J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. Nat Methods 13, 845–848.
Haghverdi, L., Lun, A.T.L., Morgan, M.D., and Marioni, J.C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat Biotechnol 36, 421–427.
Hainer, S.J., Bošković, A., McCannell, K.N., Rando, O.J., and Fazzio, T.G. (2019). Profiling of pluripotency factors in single cells and early embryos. Cell 177, 1319–1329.e11.
Hamashima, K., Wong, K.W., Sam, T.W., Teo, J.H.J., Taneja, R., Le, M.T.N., Li, Q.J., Hanna, J.H., Li, H., and Loh, Y.H. (2023). Single-nucleus multiomic mapping of m^6A methylomes and transcriptomes in native populations of cells with sn-m^6A-CT. Mol Cell 83, 3205–3216.e5.
Han, K.Y., Kim, K.T., Joung, J.G., Son, D.S., Kim, Y.J., Jo, A., Jeon, H.J., Moon, H.S., Yoo, C.E., Chung, W., et al. (2018a). SIDR: simultaneous isolation and parallel sequencing of genomic DNA and total RNA from single cells. Genome Res 28, 75–87.
Han, L., Wei, X., Liu, C., Volpe, G., Zhuang, Z., Zou, X., Wang, Z., Pan, T., Yuan, Y., Zhang, X., et al. (2022). Cell transcriptomic atlas of the non-human primate Macaca fascicularis. Nature 604, 723–731.
Han, L., Zi, X., Garmire, L.X., Wu, Y., Weissman, S.M., Pan, X., and Fan, R. (2014). Co-detection and sequencing of genes and transcripts from the same single cells facilitated by a microfluidics platform. Sci Rep 4, 6485.
Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., Saadatpour, A., Zhou, Z., Chen, H., Ye, F., et al. (2018b). Mapping the mouse cell atlas by microwell-seq. Cell 173, 1307.
Han, X., Zhou, Z., Fei, L., Sun, H., Wang, R., Chen, Y., Chen, H., Wang, J., Tang, H., Ge, W., et al. (2020). Construction of a human cell landscape at single-cell level. Nature 581, 303–309.
Hanahan, D. (2022). Hallmarks of cancer: new dimensions. Cancer Discov 12, 31–46.
Haniffa, M., Taylor, D., Linnarsson, S., Aronow, B.J., Bader, G.D., Barker, R.A., Camara, P.G., Camp, J.G., Chédotal, A., Copp, A., et al. (2021). A roadmap for the Human Developmental Cell Atlas. Nature 597, 196–205.
Hao, M., Gong, J., Zeng, X., Liu, C., Guo, Y., Cheng, X., Wang, T., Ma, J., Zhang, X., and Song, L. (2024a). Large-scale foundation model on single-cell transcriptomics. Nat Methods 21, 1481–1491.
Hao, Y., Hao, S., Andersen-Nissen, E., Mauck Iii, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell 184, 3573–3587.e29.
Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., et al. (2024b). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat Biotechnol 42, 293–304.
Hao, Z.Z., Wei, J.R., Xiao, D., Liu, R., Xu, N., Tang, L., Huang, M., Shen, Y., Xing, C., Huang, W., et al. (2022). Single-cell transcriptomics of adult macaque hippocampus reveals neural precursor cell populations. Nat Neurosci 25, 805–817.
Harroud, A., and Hafler, D.A. (2023). Common genetic factors among autoimmune diseases. Science 380, 485–490.
Hartmann, F.J., Mrdjen, D., McCaffrey, E., Glass, D.R., Greenwald, N.F., Bharadwaj, A., Khair, Z., Verberk, S.G.S., Baranski, A., Baskar, R., et al. (2021). Single-cell metabolic profiling of human cytotoxic T cells. Nat Biotechnol 39, 186–197.
Hashimshony, T., Senderovich, N., Avital, G., Klochendler, A., de Leeuw, Y., Anavy, L., Gennert, D., Li, S., Livak, K.J., Rozenblatt-Rosen, O., et al. (2016). CEL-Seq2: Sensitive highly-multiplexed single-cell RNA-Seq. Genome Biol 17, 77.
Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep 2, 666–673.
Hayashi, T., Ozaki, H., Sasagawa, Y., Umeda, M., Danno, H., and Nikaido, I. (2018). Single-cell full-length total RNA sequencing uncovers dynamics of recursive splicing and enhancer RNAs. Nat Commun 9, 619.
He, H., Chen, S., Fan, Z., Dong, Y., Wang, Y., Li, S., Sun, X., Song, Y., Yang, J., Cao, Q., et al. (2023). Multi-dimensional single-cell characterization revealed suppressive immune microenvironment in AFP-positive hepatocellular carcinoma. Cell Discov 9, 60.
He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas. 770–778.
He, P., Lim, K., Sun, D., Pett, J.P., Jeng, Q., Polanski, K., Dong, Z., Bolt, L., Richardson, L., Mamanova, L., et al. (2022). A human fetal lung cell atlas uncovers proximal-distal gradients of differentiation and key regulators of epithelial fates. Cell 185, 4841–4860.e25.
He, S., Wang, L.H., Liu, Y., Li, Y.Q., Chen, H.T., Xu, J.H., Peng, W., Lin, G.W., Wei, P.

P., Li, B., et al. (2020). Single-cell transcriptome profiling of an adult human cell atlas of 15 major organs. Genome Biol 21, 294.
Hendriks, G.J., Jung, L.A., Larsson, A.J.M., Lidschreiber, M., Andersson Forsman, O., Lidschreiber, K., Cramer, P., and Sandberg, R. (2019). NASC-seq monitors RNA synthesis in single cells. Nat Commun 10, 3138.
Hickey, J.W., Becker, W.R., Nevins, S.A., Horning, A., Perez, A.E., Zhu, C., Zhu, B., Wei, B., Chiu, R., Chen, D.C., et al. (2023). Organization of the human intestine at single-cell resolution. Nature 619, 572–584.
Hie, B., Bryson, B., and Berger, B. (2019). Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. Nat Biotechnol 37, 685–691.
Hill, A.V.S. (1998). The immunogenetics of human infectious diseases. Annu Rev Immunol 16, 593–617.
Hoadley, K.A., Yau, C., Hinoue, T., Wolf, D.M., Lazar, A.J., Drill, E., Shen, R., Taylor, A.M., Cherniack, A.D., Thorsson, V., et al. (2018). Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. Cell 173, 291–304.e6.
Hochgerner, H., Lönnerberg, P., Hodge, R., Mikes, J., Heskol, A., Hubschle, H., Lin, P., Picelli, S., La Manno, G., Ratz, M., et al. (2017). STRT-seq-2i: dual-index 5′ single cell and nucleus RNA-seq on an addressable microwell array. Sci Rep 7, 16327.
Homberger, C., Hayward, R.J., Barquist, L., Vogel, J., and Laub, M.T. (2023). Improved bacterial single-cell RNA-seq through automated MATQ-seq and Cas9-based removal of rRNA reads. mBio 14, e0355722.
Honda, M., Oki, S., Kimura, R., Harada, A., Maehara, K., Tanaka, K., Meno, C., and Ohkawa, Y. (2021). High-depth spatial transcriptome analysis by photo-isolation chemistry. Nat Commun 12, 4416.
Hou, W., and Ji, Z. (2024). Assessing GPT-4 for cell type annotation in single-cell RNA-seq analysis. Nat Methods 21, 1462–1465.
Hou, W., Ji, Z., Ji, H., and Hicks, S.C. (2020). A systematic evaluation of single-cell RNA-sequencing imputation methods. Genome Biol 21, 218.
Hou, Y., Guo, H., Cao, C., Li, X., Hu, B., Zhu, P., Wu, X., Wen, L., Tang, F., Huang, Y., et al. (2016). Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. Cell Res 26, 304–319.
Hrovatin, K., Bastidas-Ponce, A., Bakhti, M., Zappia, L., Büttner, M., Salinno, C., Sterr, M., Böttcher, A., Migliorini, A., Lickert, H., et al. (2023). Delineating mouse β-cell identity during lifetime and in diabetes with a single cell atlas. Nat Metab 5, 1615–1637.
Hu, J., Li, X., Coleman, K., Schroeder, A., Ma, N., Irwin, D.J., Lee, E.B., Shinohara, R. T., and Li, M. (2021). SpaGCN: integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. Nat Methods 18, 1342–1351.
Hu, K.H., Eichorst, J.P., McGinnis, C.S., Patterson, D.M., Chow, E.D., Kersten, K., Jameson, S.C., Gartner, Z.J., Rao, A.A., and Krummel, M.F. (2020). ZipSeq: barcoding for real-time mapping of single cell transcriptomes. Nat Methods 17, 833–843.
Hu, P., Zhang, W., Xin, H., and Deng, G. (2016a). Single cell isolation and analysis. Front Cell Dev Biol 4, 116.
Hu, T., Allam, M., Cai, S., Henderson, W., Yueh, B., Garipcan, A., Ievlev, A.V., Afkarian, M., Beyaz, S., and Coskun, A.F. (2023). Single-cell spatial metabolomics with cell-type specific protein profiling for tissue systems biology. Nat Commun 14, 8260.
Hu, W., Zeng, H., Shi, Y., Zhou, C., Huang, J., Jia, L., Xu, S., Feng, X., Zeng, Y., Xiong, T., et al. (2022). Single-cell transcriptome and translatome dual-omics reveals potential mechanisms of human oocyte maturation. Nat Commun 13, 5114.
Hu, Y., Huang, K., An, Q., Du, G., Hu, G., Xue, J., Zhu, X., Wang, C.Y., Xue, Z., and Fan, G. (2016b). Simultaneous profiling of transcriptome and DNA methylome from a single cell. Genome Biol 17, 88.
Huang, D., Zuo, Y., Zhang, C., Sun, G., Jing, Y., Lei, J., Ma, S., Sun, S., Lu, H., Zhang, X., et al. (2023a). A single-nucleus transcriptomic atlas of primate testicular aging reveals exhaustion of the spermatogonial stem cell reservoir and loss of Sertoli cell homeostasis. Protein Cell 14, 888–907.
Huang, M., Wang, J., Torre, E., Dueck, H., Shaffer, S., Bonasio, R., Murray, J.I., Raj, A., Li, M., and Zhang, N.R. (2018). SAVER: Gene expression recovery for single-cell RNA sequencing. Nat Methods 15, 539–542.
Huang, S., Wang, X., Wang, Y., Wang, Y., Fang, C., Wang, Y., Chen, S., Chen, R., Lei, T., Zhang, Y., et al. (2023b). Deciphering and advancing CAR T-cell therapy with single-cell sequencing technologies. Mol Cancer 22, 80.
Hubert, L., and Arabie, P. (1985). Comparing partitions. J Classification 2, 193–218.
Hwang, B., Lee, J.H., and Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med 50, 1–14.
Im, Y., and Kim, Y. (2023). A comprehensive overview of RNA deconvolution methods and their application. Mol Cells 46, 99–105.
Imdahl, F., Vafadarnejad, E., Homberger, C., Saliba, A.E., and Vogel, J. (2020). Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. Nat Microbiol 5, 1202–1206.
International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. Nature 431, 931–945.
Islam, S., Kjällquist, U., Moliner, A., Zajac, P., Fan, J.B., Lönnerberg, P., and Linnarsson, S. (2011). Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res 21, 1160–1167.
Jackson, H.W., Fischer, J.R., Zanotelli, V.R.T., Ali, H.R., Mechera, R., Soysal, S.D., Moch, H., Muenst, S., Varga, Z., Weber, W.P., et al. (2020). The single-cell pathology landscape of breast cancer. Nature 578, 615–620.
Jagadeesh, K.A., Dey, K.K., Montoro, D.T., Mohan, R., Gazal, S., Engreitz, J.M., Xavier, R.J., Price, A.L., and Regev, A. (2022). Identifying disease-critical cell types and cellular processes by integrating single-cell RNA-sequencing and human genetics. Nat Genet 54, 1479–1492.
Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., et al. (2014). Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science 343, 776–779.
Jaitin, D.A., Weiner, A., Yofe, I., Lara-Astiaso, D., Keren-Shaul, H., David, E., Salame, T.M., Tanay, A., van Oudenaarden, A., and Amit, I. (2016). Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. Cell 167, 1883–1896.e15.
Janssens, D.H., Otto, D.J., Meers, M.P., Setty, M., Ahmad, K., and Henikoff, S. (2022a). CUT&Tag2for1: a modified method for simultaneous profiling of the accessible and silenced regulome in single cells. Genome Biol 23, 81.
Janssens, J., Aibar, S., Taskiran, I.I., Ismail, J.N., Gomez, A.E., Aughey, G., Spanier, K. I., De Rop, F.V., González-Blas, C.B., Dionne, M., et al. (2022b). Decoding gene regulation in the fly brain. Nature 601, 630–636.
Ji, A.L., Rubin, A.J., Thrane, K., Jiang, S., Reynolds, D.L., Meyers, R.M., Guo, M.G., George, B.M., Mollbrink, A., Bergenstråhle, J., et al. (2020). Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. Cell 182, 497–514.e22.
Jiang, F., Zhou, X., Qian, Y., Zhu, M., Wang, L., Li, Z., Shen, Q., Wang, M., Qu, F., Cui, G., et al. (2023a). Simultaneous profiling of spatial gene expression and chromatin accessibility during mouse brain development. Nat Methods 20, 1048–1057.
Jiang, M., Xiao, Y., E, W., Ma, L., Wang, J., Chen, H., Gao, C., Liao, Y., Guo, Q., Peng, J., et al. (2021). Characterization of the zebrafish cell landscape at single-cell resolution. Front Cell Dev Biol 9, 743421.
Jiang, S., Huang, Z., Li, Y., Yu, C., Yu, H., Ke, Y., Jiang, L., and Liu, J. (2023b). Single-cell chromatin accessibility and transcriptome atlas of mouse embryos. Cell Rep 42, 112210.
Jiang, Y.R., Zhu, L., Cao, L.R., Wu, Q., Chen, J.B., Wang, Y., Wu, J., Zhang, T.Y., Wang, Z.L., Guan, Z.Y., et al. (2023c). Simultaneous deep transcriptome and proteome profiling in a single mouse oocyte. Cell Rep 42, 113455.
Jiao, L., Wang, G., Dai, H., Li, X., Wang, S., and Song, T. (2023). scTransSort: transformers for intelligent annotation of cell types by gene embeddings. Biomolecules 13, 611.
Jin, S., Guerrero-Juarez, C.F., Zhang, L., Chang, I., Ramos, R., Kuan, C.H., Myung, P., Plikus, M.V., and Nie, Q. (2021). Inference and analysis of cell-cell communication using CellChat. Nat Commun 12, 1088.
Jing, Y., Zuo, Y., Yu, Y., Sun, L., Yu, Z., Ma, S., Zhao, Q., Sun, G., Hu, H., Li, J., et al. (2023). Single-nucleus profiling unveils a geroprotective role of the FOXO3 in primate skeletal muscle aging. Protein Cell 14, 497–512.
Joanito, I., Wirapati, P., Zhao, N., Nawaz, Z., Yeo, G., Lee, F., Eng, C.L.P., Macalinao, D.C., Kahraman, M., Srinivasan, H., et al. (2022). Single-cell and bulk transcriptome sequencing identifies two epithelial tumor cell states and refines the consensus molecular classification of colorectal cancer. Nat Genet 54, 963–975.
Johansen, N., Somasundaram, S., Travaglini, K.J., Yanny, A.M., Shumyatcher, M., Casper, T., Cobbs, C., Dee, N., Ellenbogen, R., Ferreira, M., et al. (2023). Interindividual variation in human cortical cell type abundance and expression. Science 382, eadf2359.
Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8, 118–127.
Jordan, C.T., Guzman, M.L., and Noble, M. (2006). Cancer stem cells. N Engl J Med 355, 1253–1261.
Jorstad, N.L., Close, J., Johansen, N., Yanny, A.M., Barkan, E.R., Travaglini, K.J., Bertagnolli, D., Campos, J., Casper, T., Crichton, K., et al. (2023a). Transcriptomic cytoarchitecture reveals principles of human neocortex organization. Science 382, eadf6812.
Jorstad, N.L., Song, J.H.T., Exposito-Alonso, D., Suresh, H., Castro-Pacheco, N., Krienen, F.M., Yanny, A.M., Close, J., Gelfand, E., Long, B., et al. (2023b). Comparative transcriptomics reveals human-specific cortical features. Science 382, eade9516.
Jovic, D., Liang, X., Zeng, H., Lin, L., Xu, F., and Luo, Y. (2022). Single-cell RNA sequencing technologies and applications: a brief overview. Clin Transl Med 12, e694.

Junker, J.P., Noël, E.S., Guryev, V., Peterson, K.A., Shah, G., Huisken, J., McMahon, A. P., Berezikov, E., Bakkers, J., and van Oudenaarden, A. (2014). Genome-wide RNA tomography in the zebrafish embryo. Cell 159, 662–675.
Kalucka, J., de Rooij, L.P.M.H., Goveia, J., Rohlenova, K., Dumas, S.J., Meta, E., Conchinha, N.V., Taverna, F., Teuwen, L.A., Veys, K., et al. (2020). Single-cell transcriptome atlas of murine endothelial cells. Cell 180, 764–779.e20.
Kang, B., Camps, J., Fan, B., Jiang, H., Ibrahim, M.M., Hu, X., Qin, S., Kirchhoff, D., Chiang, D.Y., Wang, S., et al. (2022). Parallel single-cell and bulk transcriptome analyses reveal key features of the gastric tumor microenvironment. Genome Biol 23, 265.
Kang, K., Meng, Q., Shats, I., Umbach, D.M., Li, M., Li, Y., Li, X., Li, L., and Przytycka, T.M. (2019). CDSeq: A novel complete deconvolution method for dissecting heterogeneous samples using gene expression data. PLoS Comput Biol 15, e1007510.
Kang, Y., and Pantel, K. (2013). Tumor cell dissemination: emerging biological insights from animal models and cancer patients. Cancer Cell 23, 573–581.
Kanton, S., Boyle, M.J., He, Z., Santel, M., Weigert, A., Sanchís-Calleja, F., Guijarro, P., Sidow, L., Fleck, J.S., Han, D., et al. (2019). Organoid single-cell genomic atlas uncovers human-specific features of brain development. Nature 574, 418–422.
Karimi, E., Yu, M.W., Maritan, S.M., Perus, L.J.M., Rezanejad, M., Sorin, M., Dankner, M., Fallah, P., Doré, S., Zuo, D., et al. (2023). Single-cell spatial immune landscapes of primary and metastatic brain tumours. Nature 614, 555–563.
Karras, P., Bordeu, I., Pozniak, J., Nowosad, A., Pazzi, C., Van Raemdonck, N., Landeloos, E., Van Herck, Y., Pedri, D., Bervoets, G., et al. (2022). A cellular hierarchy in melanoma uncouples growth and metastasis. Nature 610, 190–198.
Kaur, R.P., Ludhiadch, A., and Munshi, A. (2019). Single-cell genomics: technology and applications. In: Single-Cell Omics. New York: Academic Press. 179–197.
Kaya-Okur, H.S., Wu, S.J., Codomo, C.A., Pledger, E.S., Bryson, T.D., Henikoff, J.G., Ahmad, K., and Henikoff, S. (2019). CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 10, 1930.
Ke, R., Mignardi, M., Pacureanu, A., Svedlund, J., Botling, J., Wählby, C., and Nilsson, M. (2013). *In situ* sequencing for RNA analysis in preserved tissue and cells. Nat Methods 10, 857–860.
Kelly, T.K., Liu, Y., Lay, F.D., Liang, G., Berman, B.P., and Jones, P.A. (2012). Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. Genome Res 22, 2497–2506.
Kelsey, G., Stegle, O., and Reik, W. (2017). Single-cell epigenomics: Recording the past and predicting the future. Science 358, 69–75.
Kieffer, Y., Hocine, H.R., Gentric, G., Pelon, F., Bernard, C., Bourachot, B., Lameiras, S., Albergante, L., Bonneau, C., Guyard, A., et al. (2020). Single-cell analysis reveals fibroblast clusters linked to immunotherapy resistance in cancer. Cancer Discov 10, 1330–1351.
Kim, C., Gao, R., Sei, E., Brandt, R., Hartman, J., Hatschek, T., Crosetto, N., Foukakis, T., and Navin, N.E. (2018). Chemoresistance evolution in triple-negative breast cancer delineated by single-cell sequencing. Cell 173, 879–893.e13.
Kim, C.N., Shin, D., Wang, A., and Nowakowski, T.J. (2023a). Spatiotemporal molecular dynamics of the developing human thalamus. Science 382, eadf9941.
Kim, D., Kobayashi, T., Voisin, B., Jo, J.H., Sakamoto, K., Jin, S.P., Kelly, M., Pasieka, H.B., Naff, J.L., Meyerle, J.H., et al. (2020). Targeted therapy guided by single-cell transcriptomic analysis in drug-induced hypersensitivity syndrome: A case report. Nat Med 26, 236–243.
Kim, S.H., Marinov, G.K., Bagdatli, S.T., Higashino, S.I., Shipony, Z., Kundaje, A., and Greenleaf, W.J. (2023b). Simultaneous single-cell profiling of the transcriptome and accessible chromatin using SHARE-seq. Methods Mol Biol 2611, 187–230.
King, H.W., Wells, K.L., Shipony, Z., Kathiria, A.S., Wagar, L.E., Lareau, C., Orban, N., Capasso, R., Davis, M.M., Steinmetz, L.M., et al. (2021). Integrated single-cell transcriptomics and epigenomics reveals strong germinal center-associated etiology of autoimmune risk loci. Sci Immunol 6, eabh3768.
Kishi, J.Y., Lapan, S.W., Beliveau, B.J., West, E.R., Zhu, A., Sasaki, H.M., Saka, S.K., Wang, Y., Cepko, C.L., and Yin, P. (2019). SABER amplifies FISH: enhanced multiplexed imaging of RNA and DNA in cells and tissues. Nat Methods 16, 533–544.
Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161, 1187–1201.
Kleshchevnikov, V., Shmatko, A., Dann, E., Aivazidis, A., King, H.W., Li, T., Elmentaite, R., Lomakin, A., Kedlian, V., Gayoso, A., et al. (2022). Cell2location maps fine-grained cell types in spatial transcriptomics. Nat Biotechnol 40, 661–671.
Koboldt, D.C., Steinberg, K.M., Larson, D.E., Wilson, R.K., and Mardis, E.R. (2013). The next-generation sequencing revolution and its impact on genomics. Cell 155, 27–38.
Kolodziejczyk, A.A., Kim, J.K., Svensson, V., Marioni, J.C., and Teichmann, S.A. (2015). The technology and biology of single-cell RNA sequencing. Mol Cell 58, 610–620.
Komatsu, J., Cico, A., Poncin, R., Le Bohec, M., Morf, J., Lipin, S., Graindorge, A., Eckert, H., Saffarian, A., Cathaly, L., et al. (2023). RevGel-seq: Instrument-free single-cell RNA sequencing using a reversible hydrogel for cell-specific barcoding. Sci Rep 13, 4866.
Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16, 1289–1296.
Kouno, T., Moody, J., Kwon, A.T.J., Shibayama, Y., Kato, S., Huang, Y., Böttcher, M., Motakis, E., Mendez, M., Severin, J., et al. (2019). C1 CAGE detects transcription start sites and enhancer activity at single-cell resolution. Nat Commun 10, 360.
Kozlov, A., Alves, J.M., Stamatakis, A., and Posada, D. (2022). CellPhy: Accurate and fast probabilistic inference of single-cell phylogenies from scDNA-seq data. Genome Biol 23, 37.
Krishnamurty, A.T., Shyer, J.A., Thai, M., Gandham, V., Buechler, M.B., Yang, Y.A., Pradhan, R.N., Wang, A.W., Sanchez, P.L., Qu, Y., et al. (2022). LRRC15+ myofibroblasts dictate the stromal setpoint to suppress tumour immunity. Nature 611, 148–154.
Ku, W.L., Nakamura, K., Gao, W., Cui, K., Hu, G., Tang, Q., Ni, B., and Zhao, K. (2019). Single-cell chromatin immunocleavage sequencing (scChIC-seq) to profile histone modification. Nat Methods 16, 323–325.
Ku, W.L., Pan, L., Cao, Y., Gao, W., and Zhao, K. (2021). Profiling single-cell histone modifications using indexing chromatin immunocleavage sequencing. Genome Res 31, 1831–1842.
Kuchina, A., Brettner, L.M., Paleologu, L., Roco, C.M., Rosenberg, A.B., Carignano, A., Kibler, R., Hirano, M., DePaolo, R.W., and Seelig, G. (2021). Microbial single-cell RNA sequencing by split-pool barcoding. Science 371, eaba5257.
Lähnemann, D., Köster, J., Szczurek, E., McCarthy, D.J., Hicks, S.C., Robinson, M.D., Vallejos, C.A., Campbell, K.R., Beerenwinkel, N., Mahfouz, A., et al. (2020). Eleven grand challenges in single-cell data science. Genome Biol 21, 31.
Lai, B., Gao, W., Cui, K., Xie, W., Tang, Q., Jin, W., Hu, G., Ni, B., and Zhao, K. (2018). Principles of nucleosome organization revealed by single-cell micrococcal nuclease sequencing. Nature 562, 281–285.
Lake, B.B., Chen, S., Hoshi, M., Plongthongkum, N., Salamon, D., Knoten, A., Vijayan, A., Venkatesh, R., Kim, E.H., Gao, D., et al. (2019). A single-nucleus RNA-sequencing pipeline to decipher the molecular anatomy and pathophysiology of human kidneys. Nat Commun 10, 2832.
Lake, B.B., Chen, S., Sos, B.C., Fan, J., Kaeser, G.E., Yung, Y.C., Duong, T.E., Gao, D., Chun, J., Kharchenko, P.V., et al. (2018). Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. Nat Biotechnol 36, 70–80.
Lake, B.B., Menon, R., Winfree, S., Hu, Q., Melo Ferreira, R., Kalhor, K., Barwinska, D., Otto, E.A., Ferkowicz, M., Diep, D., et al. (2023). An atlas of healthy and injured cell states and niches in the human kidney. Nature 619, 585–594.
Lareau, C.A., Duarte, F.M., Chew, J.G., Kartha, V.K., Burkett, Z.D., Kohlway, A.S., Pokholok, D., Aryee, M.J., Steemers, F.J., Lebofsky, R., et al. (2019). Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility. Nat Biotechnol 37, 916–924.
Lareau, C.A., Liu, V., Muus, C., Praktiknjo, S.D., Nitsch, L., Kautz, P., Sandor, K., Yin, Y., Gutierrez, J.C., Pelka, K., et al. (2023). Mitochondrial single-cell ATAC-seq for high-throughput multi-omic detection of mitochondrial genotypes and chromatin accessibility. Nat Protoc 18, 1416–1440.
Lareau, C.A., Ludwig, L.S., Muus, C., Gohil, S.H., Zhao, T., Chiang, Z., Pelka, K., Verboon, J.M., Luo, W., Christian, E., et al. (2021a). Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. Nat Biotechnol 39, 451–461.
Lareau, C.A., Parker, K.R., and Satpathy, A.T. (2021b). Charting the tumor antigen maps drawn by single-cell genomics. Cancer Cell 39, 1553–1557.
Larsson, L., Frisén, J., and Lundeberg, J. (2021). Spatially resolved transcriptomics adds a new dimension to genomics. Nat Methods 18, 15–18.
Lawson, D.A., Bhakta, N.R., Kessenbrock, K., Prummel, K.D., Yu, Y., Takai, K., Zhou, A., Eyob, H., Balakrishnan, S., Wang, C.Y., et al. (2015). Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. Nature 526, 131–135.
Leader, A.M., Grout, J.A., Maier, B.B., Nabet, B.Y., Park, M.D., Tabachnikova, A., Chang, C., Walker, L., Lansky, A., Le Berichel, J., et al. (2021). Single-cell analysis of human non-small cell lung cancer lesions refines tumor classification and patient stratification. Cancer Cell 39, 1594–1609.e12.
Lee, B.R., Dalley, R., Miller, J.A., Chartrand, T., Close, J., Mann, R., Mukora, A., Ng, L., Alfiler, L., Baker, K., et al. (2023). Signature morphoelectric properties of diverse GABAergic interneurons in the human neocortex. Science 382, eadf6484.
Lee, D.S., Luo, C., Zhou, J., Chandran, S., Rivkin, A., Bartlett, A., Nery, J.R., Fitzpatrick, C., O'Connor, C., Dixon, J.R., et al. (2019). Simultaneous profiling of 3D genome structure and DNA methylation in single human cells. Nat Methods 16,

999–1006.
Lee, J.H., Daugharthy, E.R., Scheiman, J., Kalhor, R., Ferrante, T.C., Terry, R., Turczyk, B.M., Yang, J.L., Lee, H.S., Aach, J., et al. (2015). Fluorescent *in situ* sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. Nat Protoc 10, 442–458.
Lee, J.H., Daugharthy, E.R., Scheiman, J., Kalhor, R., Yang, J.L., Ferrante, T.C., Terry, R., Jeanty, S.S.F., Li, C., Amamoto, R., et al. (2014). Highly multiplexed subcellular RNA sequencing *in situ*. Science 343, 1360–1363.
Lee, K.S., Pereira, F.C., Palatinszky, M., Behrendt, L., Alcolombri, U., Berry, D., Wagner, M., and Stocker, R. (2021). Optofluidic Raman-activated cell sorting for targeted genome retrieval or cultivation of microbial cells with specific functions. Nat Protoc 16, 634–676.
Leighton, J., Hu, M., Sei, E., Meric-Bernstam, F., and Navin, N.E. (2023). Reconstructing mutational lineages in breast cancer by multi-patient-targeted single-cell DNA sequencing. Cell Genomics 3, 100215.
Lein, E., Borm, L.E., and Linnarsson, S. (2017). The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. Science 358, 64–69.
Li, B., Li, T., Liu, J.S., and Liu, X.S. (2020). Computational deconvolution of tumor-infiltrating immune components with bulk tumor gene expression data. Methods Mol Biol 2120, 249–262.
Li, G., Liu, Y., Zhang, Y., Kubo, N., Yu, M., Fang, R., Kellis, M., and Ren, B. (2019). Joint profiling of DNA methylation and chromatin architecture in single cells. Nat Methods 16, 991–993.
Li, H., Janssens, J., De Waegeneer, M., Kolluru, S.S., Davie, K., Gardeux, V., Saelens, W., David, F.P.A., Brbić, M., Spanier, K., et al. (2022a). Fly Cell Atlas: a single-nucleus transcriptomic atlas of the adult fruit fly. Science 375, eabk2432.
Li, H., Ma, T., Hao, M., Guo, W., Gu, J., Zhang, X., and Wei, L. (2023a). Decoding functional cell-cell communication events by multi-view graph learning on spatial transcriptomics. Brief Bioinf 24, bbad359.
Li, H., Sun, Y., Hong, H., Huang, X., Tao, H., Huang, Q., Wang, L., Xu, K., Gan, J., Chen, H., et al. (2022b). Inferring transcription factor regulatory networks from single-cell ATAC-seq data based on graph neural networks. Nat Mach Intell 4, 389–400.
Li, J., Wang, J., Zhang, P., Wang, R., Mei, Y., Sun, Z., Fei, L., Jiang, M., Ma, L., E, W., et al. (2022c). Deep learning of cross-species single-cell landscapes identifies conserved regulatory programs underlying cell types. Nat Genet 54, 1711–1720.
Li, J., Zhang, Z., Song, R., Li, Y., and Du, Q. (2024). SCFormer: spectral coordinate transformer for cross-domain few-shot hyperspectral image classification. IEEE Trans Image Process 33, 840–855.
Li, J., Zheng, Y., Yan, P., Song, M., Wang, S., Sun, L., Liu, Z., Ma, S., Izpisua Belmonte, J.C., Chan, P., et al. (2021a). A single-cell transcriptomic atlas of primate pancreatic islet aging. Natl Sci Rev 8, nwaa127.
Li, L., Dong, J., Yan, L., Yong, J., Liu, X., Hu, Y., Fan, X., Wu, X., Guo, H., Wang, X., et al. (2017). Single-cell RNA-Seq analysis maps development of human germline cells and gonadal niche interactions. Cell Stem Cell 20, 858–873.e4.
Li, L., Guo, F., Gao, Y., Ren, Y., Yuan, P., Yan, L., Li, R., Lian, Y., Li, J., Hu, B., et al. (2018). Single-cell multi-omics sequencing of human early embryos. Nat Cell Biol 20, 847–858.
Li, R., Wang, Z., Wang, J.W., and Li, L. (2023b). Combining single-cell RNA sequencing with spatial transcriptome analysis reveals dynamic molecular maps of cambium differentiation in the primary and secondary growth of trees. Plant Commun 4, 100665.
Li, W., Calder, R.B., Mar, J.C., and Vijg, J. (2015). Single-cell transcriptogenomics reveals transcriptional exclusion of ENU-mutated alleles. Mutat Res 772, 55–62.
Li, W., Lu, J., Lu, P., Gao, Y., Bai, Y., Chen, K., Su, X., Li, M., Liu, J., Chen, Y., et al. (2023c). scNanoHi-C: a single-cell long-read concatemer sequencing method to reveal high-order chromatin structures within individual cells. Nat Methods 20, 1493–1505.
Li, W.V., and Li, J.J. (2018). An accurate and robust imputation method scImpute for single-cell RNA-seq data. Nat Commun 9, 997.
Li, X., and Wang, C.Y. (2021). From bulk, single-cell to spatial RNA sequencing. Int J Oral Sci 13, 36.
Li, Y., Huang, Z., Zhang, Z., Wang, Q., Li, F., Wang, S., Ji, X., Shu, S., Fang, X., and Jiang, L. (2023d). FIPRESCI: Droplet microfluidics based combinatorial indexing for massive-scale 5′-end single-cell RNA sequencing. Genome Biol 24, 70.
Li, Y., Li, H., Xie, Y., Chen, S., Qin, R., Dong, H., Yu, Y., Wang, J., Qian, X., and Qin, W. (2021b). An integrated strategy for mass spectrometry-based multiomics analysis of single cells. Anal Chem 93, 14059–14067.
Li, Y., Li, Z., Wang, C., Yang, M., He, Z., Wang, F., Zhang, Y., Li, R., Gong, Y., Wang, B., et al. (2023e). Spatiotemporal transcriptome atlas reveals the regional specification of the developing human brain. Cell 186, 5892–5909.e22.
Li, Y.E., Preissl, S., Miller, M., Johnson, N.D., Wang, Z., Jiao, H., Zhu, C., Wang, Z., Xie, Y., Poirion, O., et al. (2023f). A comparative atlas of single-cell chromatin accessibility in the human brain. Science 382, eadf7044.
Liao, S., Heng, Y., Liu, W., Xiang, J., Ma, Y., Chen, L., Feng, X., Jia, D., Liang, D., Huang, C., et al. (2023). Integrated spatial transcriptomic and proteomic analysis of fresh frozen tissue based on Stereo-seq. bioRxiv, 2023.04.28.538364.
Liao, Y., Ma, L., Guo, Q., Weigao, E., Fang, X., Yang, L., Ruan, F., Wang, J., Zhang, P., Sun, Z., et al. (2022). Cell landscape of larval and adult *Xenopus laevis* at single-cell resolution. Nat Commun 13, 4306.
Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293.
Lim, J., Chin, V., Fairfax, K., Moutinho, C., Suan, D., Ji, H., and Powell, J.E. (2023). Transitioning single-cell genomics into the clinic. Nat Rev Genet 24, 573–584.
Lin, D., Shen, L., Luo, M., Zhang, K., Li, J., Yang, Q., Zhu, F., Zhou, D., Zheng, S., Chen, Y., et al. (2021a). Circulating tumor cells: biology and clinical significance. Sig Transduct Target Ther 6, 404.
Lin, T.Y., Gerber, T., Taniguchi-Sugiura, Y., Murawala, P., Hermann, S., Grosser, L., Shibata, E., Treutlein, B., and Tanaka, E.M. (2021b). Fibroblast dedifferentiation as a determinant of successful regeneration. Dev Cell 56, 1541–1551.e6.
Lin, Y., Ghazanfar, S., Wang, K.Y.X., Gagnon-Bartsch, J.A., Lo, K.K., Su, X., Han, Z.G., Ormerod, J.T., Speed, T.P., Yang, P., et al. (2019). scMerge leverages factor analysis, stable expression, and pseudoreplication to merge multiple single-cell RNA-seq datasets. Proc Natl Acad Sci USA 116, 9775–9784.
Lindeboom, R.G.H., Regev, A., and Teichmann, S.A. (2021). Towards a human cell atlas: taking notes from the past. Trends Genet 37, 625–630.
Linderman, G.C., Zhao, J., Roulis, M., Bielecki, P., Flavell, R.A., Nadler, B., and Kluger, Y. (2022). Zero-preserving imputation of single-cell RNA-seq data. Nat Commun 13, 192.
Ling, A.L., Solomon, I.H., Landivar, A.M., Nakashima, H., Woods, J.K., Santos, A., Masud, N., Fell, G., Mo, X., Yilmaz, A.S., et al. (2023). Clinical trial links oncolytic immunoactivation to survival in glioblastoma. Nature 623, 157–166.
Liscovitch-Brauer, N., Montalbano, A., Deng, J., Méndez-Mancilla, A., Wessels, H.H., Moss, N.G., Kung, C.Y., Sookdeo, A., Guo, X., Geller, E., et al. (2021). Profiling the genetic determinants of chromatin accessibility with scalable single-cell CRISPR screens. Nat Biotechnol 39, 1270–1277.
Litviňuková, M., Talavera-López, C., Maatz, H., Reichart, D., Worth, C.L., Lindberg, E.L., Kanda, M., Polanski, K., Heinig, M., Lee, M., et al. (2020). Cells of the adult human heart. Nature 588, 466–472.
Liu, B., Hu, X., Feng, K., Gao, R., Xue, Z., Zhang, S., Zhang, Y., Corse, E., Hu, Y., Han, W., et al. (2022a). Temporal single-cell tracing reveals clonal revival and expansion of precursor exhausted T cells during anti-PD-1 therapy in lung cancer. Nat Cancer 3, 108–121.
Liu, B., Zhang, Y., Wang, D., Hu, X., and Zhang, Z. (2022b). Single-cell meta-analyses reveal responses of tumor-reactive CXCL13^{+} T cells to immune-checkpoint blockade. Nat Cancer 3, 1123–1136.
Liu, C., Li, X., Huang, Q., Zhang, M., Lei, T., Wang, F., Zou, W., Huang, R., Hu, X., Wang, C., et al. (2023a). Single-cell RNA-sequencing reveals radiochemotherapy-induced innate immune activation and MHC-II upregulation in cervical cancer. Sig Transduct Target Ther 8, 44.
Liu, F., Wang, Y., Gu, H., and Wang, X. (2023b). Technologies and applications of single-cell DNA methylation sequencing. Theranostics 13, 2439–2454.
Liu, H., Zhou, J., Tian, W., Luo, C., Bartlett, A., Aldridge, A., Lucero, J., Osteen, J.K., Nery, J.R., Chen, H., et al. (2021). DNA methylation atlas of the mouse brain at single-cell resolution. Nature 598, 120–128.
Liu, J., Gao, C., Sodicoff, J., Kozareva, V., Macosko, E.Z., and Welch, J.D. (2020a). Jointly defining cell types from multiple single-cell datasets using LIGER. Nat Protoc 15, 3632–3662.
Liu, L., Liu, C., Quintero, A., Wu, L., Yuan, Y., Wang, M., Cheng, M., Leng, L., Xu, L., Dong, G., et al. (2019). Deconvolution of single-cell multi-omics layers reveals regulatory heterogeneity. Nat Commun 10, 470.
Liu, X., Mao, D., Song, Y., Zhu, L., Isak, A.N., Lu, C., Deng, G., Chen, F., Sun, F., Yang, Y., et al. (2022c). Computer-aided design of reversible hybridization chain reaction (CAD-HCR) enables multiplexed single-cell spatial proteomics imaging. Sci Adv 8, eabk0133.
Liu, X., Shen, Q., and Zhang, S. (2023c). Cross-species cell-type assignment from single-cell RNA-seq data by a heterogeneous graph neural network. Genome Res 33, 96–111.
Liu, X., Song, J., Zhang, H., Liu, X., Zuo, F., Zhao, Y., Zhao, Y., Yin, X., Guo, X., Wu, X., et al. (2023d). Immune checkpoint HLA-E: CD94-NKG2A mediates evasion of circulating tumor cells from NK cell surveillance. Cancer Cell 41, 272–287.e9.
Liu, Y., DiStasio, M., Su, G., Asashima, H., Enninful, A., Qin, X., Deng, Y., Nam, J., Gao, F., Bordignon, P., et al. (2023e). High-plex protein and whole transcriptome co-mapping at cellular resolution with spatial CITE-seq. Nat Biotechnol 41, 1405–1409.
Liu, Y., Wang, H., Taylor, M., Cook, C., Martínez-Berdeja, A., North, J.P., Harirchian,

P., Hailer, A.A., Zhao, Z., Ghadially, R., et al. (2022d). Classification of human chronic inflammatory skin disease based on single-cell immune profiling. Sci Immunol 7, eabl9165.
Liu, Y., Yang, M., Deng, Y., Su, G., Enninful, A., Guo, C.C., Tebaldi, T., Zhang, D., Kim, D., Bai, Z., et al. (2020b). High-spatial-resolution multi-omics sequencing via deterministic barcoding in tissue. Cell 183, 1665–1681.e18.
Liu, Z., Chen, Y., Xia, Q., Liu, M., Xu, H., Chi, Y., Deng, Y., and Xing, D. (2023f). Linking genome structures to functions by simultaneous single-cell Hi-C and RNA-seq. Science 380, 1070–1076.
Llorens-Bobadilla, E., Zamboni, M., Marklund, M., Bhalla, N., Chen, X., Hartman, J., Frisén, J., and Ståhl, P.L. (2023). Solid-phase capture and profiling of open chromatin by spatial ATAC. Nat Biotechnol 41, 1085–1088.
Lochs, S.J.A., van der Weide, R.H., de Luca, K.L., Korthout, T., van Beek, R.E., Kimura, H., and Kind, J. (2024). Combinatorial single-cell profiling of major chromatin types with MAbID. Nat Methods 21, 72–82.
Lopez, R., Regier, J., Cole, M.B., Jordan, M.I., and Yosef, N. (2018). Deep generative modeling for single-cell transcriptomics. Nat Methods 15, 1053–1058.
Lotfollahi, M., Wolf, F.A., and Theis, F.J. (2019). scGen predicts single-cell perturbation responses. Nat Methods 16, 715–721.
Lovatt, D., Ruble, B.K., Lee, J., Dueck, H., Kim, T.K., Fisher, S., Francis, C., Spaethling, J.M., Wolf, J.A., Grady, M.S., et al. (2014). Transcriptome in vivo analysis (TIVA) of spatially defined single cells in live tissue. Nat Methods 11, 190–196.
Lu, Q., Liu, J., Zhao, S., Gomez Castro, M.F., Laurent-Rolle, M., Dong, J., Ran, X., Damani-Yokota, P., Tang, H., Karakousi, T., et al. (2021). SARS-CoV-2 exacerbates proinflammatory responses in myeloid cells through C-type lectin receptors and Tweety family member 2. Immunity 54, 1304–1319.e9.
Lu, T., Ang, C.E., and Zhuang, X. (2022). Spatially resolved epigenomic profiling of single cells in complex tissues. Cell 185, 4448–4464.e17.
Lu, T.C., Brbić, M., Park, Y.J., Jackson, T., Chen, J., Kolluru, S.S., Qi, Y., Katheder, N. S., Cai, X.T., Lee, S., et al. (2023). Aging Fly Cell Atlas identifies exhaustive aging features at cellular resolution. Science 380, eadg0934.
Lu, Y., Zhao, Q., Liao, J.Y., Song, E., Xia, Q., Pan, J., Li, Y., Li, J., Zhou, B., Ye, Y., et al. (2020). Complement signals determine opposite effects of B cells in chemotherapy-induced immunity. Cell 180, 1081–1097.e24.
Lubeck, E., and Cai, L. (2012). Single-cell systems biology by super-resolution imaging and combinatorial labeling. Nat Methods 9, 743–748.
Luecken, M.D., Büttner, M., Chaichoompu, K., Danese, A., Interlandi, M., Mueller, M. F., Strobl, D.C., Zappia, L., Dugas, M., Colomé-Tatché, M., et al. (2022). Benchmarking atlas-level data integration in single-cell genomics. Nat Methods 19, 41–50.
Luo, C., Keown, C.L., Kurihara, L., Zhou, J., He, Y., Li, J., Castanon, R., Lucero, J., Nery, J.R., Sandoval, J.P., et al. (2017). Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. Science 357, 600–604.
Luo, C., Liu, H., Xie, F., Armand, E.J., Siletti, K., Bakken, T.E., Fang, R., Doyle, W.I., Stuart, T., Hodge, R.D., et al. (2022). Single nucleus multi-omics identifies human cortical cell regulatory genome diversity. Cell Genomics 2, 100107.
Luo, C., Rivkin, A., Zhou, J., Sandoval, J.P., Kurihara, L., Lucero, J., Castanon, R., Nery, J.R., Pinto-Duarte, A., Bui, B., et al. (2018). Robust single-cell DNA methylome profiling with snmC-seq2. Nat Commun 9, 3824.
Luo, Z., Wen, X., Zhao, W., Nguyen, T.C., Wan, X., Richard, J.L.C., and Zhong, S. (2023). Joint profiling of multiplex chromatin interactions, gene expression, and RNA-chromatin associations in single cells of the human brain. bioRxiv, 2023.06.28.546457.
Luoma, A.M., Suo, S., Williams, H.L., Sharova, T., Sullivan, K., Manos, M., Bowling, P., Hodi, F.S., Rahma, O., Sullivan, R.J., et al. (2020). Molecular pathways of colon inflammation induced by cancer immunotherapy. Cell 182, 655–671.e22.
Luquette, L.J., Bohrson, C.L., Sherman, M.A., and Park, P.J. (2019). Identification of somatic mutations in single cell DNA-seq using a spatial model of allelic imbalance. Nat Commun 10, 3908.
Lv, J., Wei, Y., Yin, J.H., Chen, Y.P., Zhou, G.Q., Wei, C., Liang, X.Y., Zhang, Y., Zhang, C.J., He, S.W., et al. (2023). The tumor immune microenvironment of nasopharyngeal carcinoma after gemcitabine plus cisplatin treatment. Nat Med 29, 1424–1436.
Ma, A., Wang, C., Chang, Y., Brennan, F.H., McDermaid, A., Liu, B., Zhang, C., Popovich, P.G., and Ma, Q. (2020a). IRIS3: Integrated cell-type-specific regulon inference server from single-cell RNA-Seq. Nucleic Acids Res 48, W275–W286.
Ma, C., Yang, C., Peng, A., Sun, T., Ji, X., Mi, J., Wei, L., Shen, S., and Feng, Q. (2023a). Pan-cancer spatially resolved single-cell analysis reveals the crosstalk between cancer-associated fibroblasts and tumor microenvironment. Mol Cancer 22, 170.
Ma, P., Amemiya, H.M., He, L.L., Gandhi, S.J., Nicol, R., Bhattacharyya, R.P., Smillie, C.S., and Hung, D.T. (2023b). Bacterial droplet-based single-cell RNA-seq reveals antibiotic-associated heterogeneous cellular states. Cell 186, 877–891.e14.
Ma, Q., and Xu, D. (2022). Deep learning shapes single-cell data analysis. Nat Rev Mol Cell Biol 23, 303–304.
Ma, S., Sun, S., Li, J., Fan, Y., Qu, J., Sun, L., Wang, S., Zhang, Y., Yang, S., Liu, Z., et al. (2021). Single-cell transcriptomic atlas of primate cardiopulmonary aging. Cell Res 31, 415–432.
Ma, S., Zhang, B., LaFave, L.M., Earl, A.S., Chiang, Z., Hu, Y., Ding, J., Brack, A., Kartha, V.K., Tay, T., et al. (2020b). Chromatin potential identified by shared single-cell profiling of RNA and chromatin. Cell 183, 1103–1116.e20.
Ma, Y., and Zhou, X. (2022). Spatially informed cell-type deconvolution for spatial transcriptomics. Nat Biotechnol 40, 1349–1359.
Macarron, R., Banks, M.N., Bojanic, D., Burns, D.J., Cirovic, D.A., Garyantes, T., Green, D.V.S., Hertzberg, R.P., Janzen, W.P., Paslay, J.W., et al. (2011). Impact of high-throughput screening in biomedical research. Nat Rev Drug Discov 10, 188–195.
Macaulay, I.C., Haerty, W., Kumar, P., Li, Y.I., Hu, T.X., Teng, M.J., Goolam, M., Saurat, N., Coupland, P., Shirley, L.M., et al. (2015). G&T-seq: Parallel sequencing of single-cell genomes and transcriptomes. Nat Methods 12, 519–522.
Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 161, 1202–1214.
Maden, S.K., Kwon, S.H., Huuki-Myers, L.A., Collado-Torres, L., Hicks, S.C., and Maynard, K.R. (2023). Challenges and opportunities to computationally deconvolve heterogeneous tissue with varying cell sizes using single-cell RNA-sequencing datasets. Genome Biol 24, 288.
Magen, A., Hamon, P., Fiaschi, N., Soong, B.Y., Park, M.D., Mattiuz, R., Humblin, E., Troncoso, L., D'souza, D., Dawson, T., et al. (2023). Intratumoral dendritic cell-CD4$^+$ T helper cell niches enable CD8$^+$ T cell differentiation following PD-1 blockade in hepatocellular carcinoma. Nat Med 29, 1389–1399.
Mah, J.L., and Dunn, C.W. (2024). Cell type evolution reconstruction across species through cell phylogenies of single-cell RNA sequencing data. Nat Ecol Evol 8, 325–338.
Mahat, D.B., Tippens, N.D., Martin-Rufino, J.D., Waterton, S.K., Fu, J., Blatt, S.E., and Sharp, P.A. (2024). Single-cell nascent RNA sequencing unveils coordinated global transcription. Nature 631, 216–223.
Maier, B., Leader, A.M., Chen, S.T., Tung, N., Chang, C., LeBerichel, J., Chudnovskiy, A., Maskey, S., Walker, L., Finnigan, J.P., et al. (2020). A conserved dendritic-cell regulatory program limits antitumour immunity. Nature 580, 257–262.
Markham, A.F. (1993). The polymerase chain reaction: a tool for molecular medicine. BMJ 306, 441–446.
Markodimitraki, C.M., Rang, F.J., Rooijers, K., de Vries, S.S., Chialastri, A., de Luca, K. L., Lochs, S.J.A., Mooijman, D., Dey, S.S., and Kind, J. (2020). Simultaneous quantification of protein–DNA interactions and transcriptomes in single cells with scDam&T-seq. Nat Protoc 15, 1922–1953.
Martin, J.C., Chang, C., Boschetti, G., Ungaro, R., Giri, M., Grout, J.A., Gettler, K., Chuang, L., Nayar, S., Greenstein, A.J., et al. (2019). Single-cell analysis of Crohn's disease lesions identifies a pathogenic cellular module associated with resistance to anti-TNF therapy. Cell 178, 1493–1508.e20.
Martinelli, A.L., and Rapsomaniki, M.A. (2022). ATHENA: Analysis of tumor heterogeneity from spatial omics measurements. Bioinformatics 38, 3151–3153.
Martins, F., Sofiya, L., Sykiotis, G.P., Lamine, F., Maillard, M., Fraga, M., Shabafrouz, K., Ribi, C., Cairoli, A., Guex-Crosier, Y., et al. (2019). Adverse effects of immune-checkpoint inhibitors: epidemiology, management and surveillance. Nat Rev Clin Oncol 16, 563–580.
Marx, V. (2021). Method of the year: spatially resolved transcriptomics. Nat Methods 18, 9–14.
McGinnis, C.S., Patterson, D.M., Winkler, J., Conrad, D.N., Hein, M.Y., Srivastava, V., Hu, J.L., Murrow, L.M., Weissman, J.S., Werb, Z., et al. (2019). MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. Nat Methods 16, 619–626.
McKellar, D.W., Mantri, M., Hinchman, M.M., Parker, J.S.L., Sethupathy, P., Cosgrove, B.D., and De Vlaminck, I. (2023). Spatial mapping of the total transcriptome by *in situ* polyadenylation. Nat Biotechnol 41, 513–520.
McNulty, R., Sritharan, D., Pahng, S.H., Meisch, J.P., Liu, S., Brennan, M.A., Saxer, G., Hormoz, S., and Rosenthal, A.Z. (2023). Probe-based bacterial single-cell RNA sequencing predicts toxin regulation. Nat Microbiol 8, 934–945.
Meers, M.P., Llagas, G., Janssens, D.H., Codomo, C.A., and Henikoff, S. (2023). Multifactorial profiling of epigenetic landscapes at single-cell resolution using MulTI-Tag. Nat Biotechnol 41, 708–716.
Menden, K., Marouf, M., Oller, S., Dalmia, A., Magruder, D.S., Kloiber, K., Heutink, P., and Bonn, S. (2020). Deep learning-based cell composition analysis from tissue expression profiles. Sci Adv 6, eaba2619.
Meric-Bernstam, F., Larkin, J., Tabernero, J., and Bonini, C. (2021). Enhancing anti-tumour efficacy with immunotherapy combinations. Lancet 397, 1010–1022.
Merritt, C.R., Ong, G.T., Church, S.E., Barker, K., Danaher, P., Geiss, G., Hoang, M.,

Jung, J., Liang, Y., McKay-Fleisch, J., et al. (2020). Multiplex digital spatial profiling of proteins and RNA in fixed tissue. Nat Biotechnol 38, 586–599.
Metzker, M.L. (2010). Sequencing technologies—the next generation. Nat Rev Genet 11, 31–46.
Mezger, A., Klemm, S., Mann, I., Brower, K., Mir, A., Bostick, M., Farmer, A., Fordyce, P., Linnarsson, S., and Greenleaf, W. (2018). High-throughput chromatin accessibility profiling at single-cell resolution. Nat Commun 9, 3647.
Mimitou, E.P., Cheng, A., Montalbano, A., Hao, S., Stoeckius, M., Legut, M., Roush, T., Herrera, A., Papalexi, E., Ouyang, Z., et al. (2019). Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. Nat Methods 16, 409–412.
Minoura, K., Abe, K., Nam, H., Nishikawa, H., and Shimamura, T. (2021). A mixture-of-experts deep generative model for integrated analysis of single-cell multiomics data. Cell Rep Methods 1, 100071.
Mitra, R.D., Butty, V.L., Shendure, J., Williams, B.R., Housman, D.E., and Church, G. M. (2003). Digital genotyping and haplotyping with polymerase colonies. Proc Natl Acad Sci USA 100, 5926–5931.
Mitra, R.D., and Church, G.M. (1999). *In situ* localized amplification and contact replication of many individual DNA molecules. Nucleic Acids Res 27, 34e–34.
Mittnenzweig, M., Mayshar, Y., Cheng, S., Ben-Yair, R., Hadas, R., Rais, Y., Chomsky, E., Reines, N., Uzonyi, A., Lumerman, L., et al. (2021). A single-embryo, single-cell time-resolved model for mouse gastrulation. Cell 184, 2825–2842.e22.
Moffitt, J.R., Hao, J., Wang, G., Chen, K.H., Babcock, H.P., and Zhuang, X. (2016). High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence *in situ* hybridization. Proc Natl Acad Sci USA 113, 11046–11051.
Moffitt, J.R., and Zhuang, X. (2016). RNA imaging with multiplexed error-robust fluorescence *in situ* hybridization (MERFISH). Methods Enzymol 572, 1–49.
Monaco, G., Lee, B., Xu, W., Mustafah, S., Hwang, Y.Y., Carré, C., Burdin, N., Visan, L., Ceccarelli, M., Poidinger, M., et al. (2019). RNA-Seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types. Cell Rep 26, 1627–1640.e7.
Mongia, A., Sengupta, D., and Majumdar, A. (2019). McImpute: matrix completion based imputation for single cell RNA-seq data. Front Genet 10, 9.
Morris, J.A., Caragine, C., Daniloski, Z., Domingo, J., Barry, T., Lu, L., Davis, K., Ziosi, M., Glinos, D.A., Hao, S., et al. (2023). Discovery of target genes and pathways at GWAS loci by pooled single-cell CRISPR screens. Science 380, eadh7699.
Moses, L., and Pachter, L. (2022). Museum of spatial transcriptomics. Nat Methods 19, 534–546.
Mukherjee, P.K., Nguyen, Q.T., Li, J., Zhao, S., Christensen, S.M., West, G.A., Chandra, J., Gordon, I.O., Lin, S., Wang, J., et al. (2023). Stricturing Crohn's disease single-cell RNA sequencing reveals fibroblast heterogeneity and inter-cellular interactions. Gastroenterology 165, 1180–1196.
Mulder, K., Patel, A.A., Kong, W.T., Piot, C., Halitzki, E., Dunsmore, G., Khalilnezhad, S., Irac, S.E., Dubuisson, A., Chevrier, M., et al. (2021). Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease. Immunity 54, 1883–1900.e5.
Mulqueen, R.M., Pokholok, D., O'Connell, B.L., Thornton, C.A., Zhang, F., O'Roak, B. J., Link, J., Yardımcı, G.G., Sears, R.C., Steemers, F.J., et al. (2021). High-content single-cell combinatorial indexing. Nat Biotechnol 39, 1574–1580.
Nadappuram, B.P., Cadinu, P., Barik, A., Ainscough, A.J., Devine, M.J., Kang, M., Gonzalez-Garcia, J., Kittler, J.T., Willison, K.R., Vilar, R., et al. (2019). Nanoscale tweezers for single-cell biopsies. Nat Nanotech 14, 80–88.
Nagano, T., Lubling, Y., Stevens, T.J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E.D., Tanay, A., and Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature 502, 59–64.
Nathan, A., Asgari, S., Ishigaki, K., Valencia, C., Amariuta, T., Luo, Y., Beynor, J.I., Baglaenko, Y., Suliman, S., Price, A.L., et al. (2022). Single-cell eQTL models reveal dynamic T cell state dependence of disease loci. Nature 606, 120–128.
Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn, M., and Alizadeh, A.A. (2015). Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 12, 453–457.
Newman, A.M., Steen, C.B., Liu, C.L., Gentles, A.J., Chaudhuri, A.A., Scherer, F., Khodadoust, M.S., Esfahani, M.S., Luca, B.A., Steiner, D., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol 37, 773–782.
Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32, 268–274.
Nichterwitz, S., Chen, G., Aguila Benitez, J., Yilmaz, M., Storvall, H., Cao, M., Sandberg, R., Deng, Q., and Hedlund, E. (2016). Laser capture microscopy coupled with Smart-seq2 for precise spatial transcriptomic profiling. Nat Commun 7, 12139.
Nishino, M., Ramaiya, N.H., Hatabu, H., and Hodi, F.S. (2017). Monitoring immune-checkpoint blockade: response evaluation and biomarker development. Nat Rev Clin Oncol 14, 655–668.
Niu, M., Cao, W., Wang, Y., Zhu, Q., Luo, J., Wang, B., Zheng, H., Weitz, D.A., and Zong, C. (2023). Droplet-based transcriptome profiling of individual synapses. Nat Biotechnol 41, 1332–1344.
Nolan, E., Lindeman, G.J., and Visvader, J.E. (2023). Deciphering breast cancer: from biology to the clinic. Cell 186, 1708–1728.
O'Connor, O.M., Alnahhas, R.N., Lugagne, J.B., and Dunlop, M.J. (2022). DeLTA 2.0: a deep learning pipeline for quantifying single-cell spatial and temporal dynamics. PLoS Comput Biol 18, e1009797.
Oliveira, G., Stromhaug, K., Klaeger, S., Kula, T., Frederick, D.T., Le, P.M., Forman, J., Huang, T., Li, S., Zhang, W., et al. (2021). Phenotype, specificity and avidity of antitumour CD8$^+$ T cells in melanoma. Nature 596, 119–125.
Packer, J.S., Zhu, Q., Huynh, C., Sivaramakrishnan, P., Preston, E., Dueck, H., Stefanik, D., Tan, K., Trapnell, C., Kim, J., et al. (2019). A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. Science 365, 1265.
Palla, G., Fischer, D.S., Regev, A., and Theis, F.J. (2022a). Spatial components of molecular tissue biology. Nat Biotechnol 40, 308–318.
Palla, G., Spitzer, H., Klein, M., Fischer, D., Schaar, A.C., Kuemmerle, L.B., Rybakov, S., Ibarra, I.L., Holmberg, O., Virshup, I., et al. (2022b). Squidpy: a scalable framework for spatial omics analysis. Nat Methods 19, 171–178.
Pan, J., Li, Y., Lin, Z., Lan, Q., Chen, H., Zhai, M., Sui, S., Zhang, G., Cheng, Y., Tang, Y., et al. (2023). Spatiotemporal transcriptome atlas of human embryos after gastrulation. bioRxiv, 2023.2004.2022.537909.
Pan, L., Ku, W.L., Tang, Q., Cao, Y., and Zhao, K. (2022). scPCOR-seq enables co-profiling of chromatin occupancy and RNAs in single cells. Commun Biol 5, 678.
Pang, M., and Tegnér, J. (2020). Multitask learning for transformers with application to large-scale single-cell transcriptomes. bioRxiv, 2020.2002.2005.935239.
Papalexi, E., Mimitou, E.P., Butler, A.W., Foster, S., Bracken, B., Mauck Iii, W.M., Wessels, H.H., Hao, Y., Yeung, B.Z., Smibert, P., et al. (2021). Characterizing the molecular regulation of inhibitory immune checkpoints with multimodal single-cell screens. Nat Genet 53, 322–331.
Patruno, L., Maspero, D., Craighero, F., Angaroni, F., Antoniotti, M., and Graudenzi, A. (2021). A review of computational strategies for denoising and imputation of single-cell transcriptomic data. Brief Bioinf 22, bbaa222.
Pellegrino, M., Sciambi, A., Treusch, S., Durruthy-Durruthy, R., Gokhale, K., Jacob, J., Chen, T.X., Geis, J.A., Oldham, W., Matthews, J., et al. (2018). High-throughput single-cell DNA sequencing of acute myeloid leukemia tumors with droplet microfluidics. Genome Res 28, 1345–1352.
Peng, G., Suo, S., Chen, J., Chen, W., Liu, C., Yu, F., Wang, R., Chen, S., Sun, N., Cui, G., et al. (2016). Spatial transcriptome for the molecular annotation of lineage fates and cell identity in mid-gastrula mouse embryo. Dev Cell 36, 681–697.
Peng, G., Suo, S., Cui, G., Yu, F., Wang, R., Chen, J., Chen, S., Liu, Z., Chen, G., Qian, Y., et al. (2019a). Molecular architecture of lineage allocation and tissue organization in early mouse embryo. Nature 572, 528–532.
Peng, S., Petersen, J.L., Bellone, R.R., Kalbfleisch, T., Kingsley, N.B., Barber, A.M., Cappelletti, E., Giulotto, E., and Finno, C.J. (2021). Decoding the equine genome: lessons from ENCODE. Genes 12, 1707.
Peng, X.L., Moffitt, R.A., Torphy, R.J., Volmar, K.E., and Yeh, J.J. (2019b). *De novo* compartment deconvolution and weight estimation of tumor samples using DECODER. Nat Commun 10, 4729.
Penkov, D., Zubkova, E., and Parfyonova, Y. (2023). Tn5 DNA transposase in multi-omics research. Methods Protoc 6, 24.
Pensold, D., and Zimmer-Bensch, G. (2020). Methods for single-cell isolation and preparation. In: Yu, B., Zhang, J., Zeng, Y., Li, L., and Wang, X., eds. Single-cell Sequencing and Methylation. Advances in Experimental Medicine and Biology. Singapore: Springer. 7–27.
Perez, R.K., Gordon, M.G., Subramaniam, M., Kim, M.C., Hartoularos, G.C., Targ, S., Sun, Y., Ogorodnikov, A., Bueno, R., Lu, A., et al. (2022). Single-cell RNA-seq reveals cell type-specific molecular and genetic associations to lupus. Science 376, eabf1970.
Picelli, S., Björklund, Å.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods 10, 1096–1098.
Pierce, S.E., Granja, J.M., and Greenleaf, W.J. (2021). High-throughput single-cell chromatin accessibility CRISPR screens enable unbiased identification of regulatory networks in cancer. Nat Commun 12, 2969.
Pijuan-Sala, B., Griffiths, J.A., Guibentif, C., Hiscock, T.W., Jawaid, W., Calero-Nieto, F.J., Mulas, C., Ibarra-Soria, X., Tyser, R.C.V., Ho, D.L.L., et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. Nature 566, 490–495.
Pisetsky, D.S. (2023). Pathogenesis of autoimmune disease. Nat Rev Nephrol 19, 509–524.

Player, A.N., Shen, L.P., Kenny, D., Antao, V.P., and Kolberg, J.A. (2001). Single-copy gene detection using branched DNA (bDNA) *in situ* hybridization. J Histochem Cytochem 49, 603–611.
Pliner, H.A., Shendure, J., and Trapnell, C. (2019). Supervised classification enables rapid annotation of cell atlases. Nat Methods 16, 983–986.
Poirion, O.B., Zhu, X., Ching, T., and Garmire, L. (2016). Single-cell transcriptomics bioinformatics and computational challenges. Front Genet 7, 163.
Polański, K., Young, M.D., Miao, Z., Meyer, K.B., Teichmann, S.A., Park, J.E., and Berger, B. (2020). BBKNN: fast batch alignment of single cell transcriptomes. Bioinformatics 36, 964–965.
Pott, S. (2017). Simultaneous measurement of chromatin accessibility, DNA methylation, and nucleosome phasing in single cells. eLife 6, e23203.
Prasetyanti, P.R., and Medema, J.P. (2017). Intra-tumor heterogeneity from a cancer stem cell perspective. Mol Cancer 16, 41.
Preissl, S., Gaulton, K.J., and Ren, B. (2023). Characterizing *cis*-regulatory elements using single-cell epigenomics. Nat Rev Genet 24, 21–43.
Qian, J., Liao, J., Liu, Z., Chi, Y., Fang, Y., Zheng, Y., Shao, X., Liu, B., Cui, Y., Guo, W., et al. (2023). Reconstruction of the cell pseudo-space from single-cell RNA sequencing data with scSpace. Nat Commun 14, 2484.
Qiu, C., Martin, B.K., Welsh, I.C., Daza, R.M., Le, T.M., Huang, X., Nichols, E.K., Taylor, M.L., Fulton, O., O'Day, D.R., et al. (2024). A single-cell time-lapse of mouse prenatal development from gastrula to birth. Nature 626, 1084–1093.
Qiu, Q., Hu, P., Qiu, X., Govek, K.W., Cámara, P.G., and Wu, H. (2020). Massively parallel and time-resolved RNA sequencing in single cells with scNT-seq. Nat Methods 17, 991–1001.
Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nat Methods 14, 979–982.
Qu, J., Yang, F., Zhu, T., Wang, Y., Fang, W., Ding, Y., Zhao, X., Qi, X., Xie, Q., Chen, M., et al. (2022). A reference single-cell regulomic and transcriptomic map of cynomolgus monkeys. Nat Commun 13, 4069.
Qu, Y., Liao, Z., Wang, X., Zhang, J., and Liu, C. (2019). EFLDO sensitizes liver cancer cells to TNFSF10-induced apoptosis in a p53-dependent manner. Mol Med Report 19, 3799–3806.
Quon, G., Haider, S., Deshwar, A.G., Cui, A., Boutros, P.C., and Morris, Q. (2013). Computational purification of individual tumor gene expression profiles leads to significant improvements in prognostic prediction. Genome Med 5, 29.
Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nat Methods 5, 877–879.
Ramsköld, D., Luo, S., Wang, Y.C., Li, R., Deng, Q., Faridani, O.R., Daniels, G.A., Khrebtukova, I., Loring, J.F., Laurent, L.C., et al. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 30, 777–782.
Rang, F.J., de Luca, K.L., de Vries, S.S., Valdes-Quezada, C., Boele, E., Nguyen, P.D., Guerreiro, I., Sato, Y., Kimura, H., Bakkers, J., et al. (2022). Single-cell profiling of transcriptome and histone modifications with EpiDamID. Mol Cell 82, 1956–1970.e14.
Rappez, L., Stadler, M., Triana, S., Gathungu, R.M., Ovchinnikova, K., Phapale, P., Heikenwalder, M., and Alexandrov, T. (2021). SpaceM reveals metabolic states of single cells. Nat Methods 18, 799–805.
Rashid, S., Shah, S., Bar-Joseph, Z., Pandya, R., and Janet, K. (2021). Dhaka: variational autoencoder for unmasking tumor heterogeneity from single cell genomic data. Bioinformatics 37, 1535–1543.
Redit, C., Cha, S., and Ai, N. (2023). Single-cell proteomics: challenges and prospects. Nature methods 20, 317–318.
Reed, B.D., Meyer, M.J., Abramzon, V., Ad PhD, O., Ad, O., Adcock, P., Ahmad, F.R., Alppay, G., Ball, J.A., Beach, J., et al. (2022). Real-time dynamic single-molecule protein sequencing on an integrated semiconductor device. Science 378, 186–192.
Regev, A., Teichmann, S., Rozenblatt-Rosen, O., Stubbington, M., Ardlie, K., Amit, I., Arlotta, P., Bader, G., Benoist, C., and Biton, M. (2018). The human cell atlas white paper. arXiv,181005192.
Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., et al. (2017). The Human Cell Atlas. eLife 6, e27041.
Reimegård, J., Tarbier, M., Danielsson, M., Schuster, J., Baskaran, S., Panagiotou, S., Dahl, N., Friedländer, M.R., and Gallant, C.J. (2021). A combined approach for single-cell mRNA and intracellular protein expression analysis. Commun Biol 4, 624.
Ren, X., Wen, W., Fan, X., Hou, W., Su, B., Cai, P., Li, J., Liu, Y., Tang, F., Zhang, F., et al. (2021). COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. Cell 184, 1895–1913.e19.
Reuter, J.A., Spacek, D.V., and Snyder, M.P. (2015). High-throughput sequencing technologies. Mol Cell 58, 586–597.
Reza, K.K., Dey, S., Wuethrich, A., Jing Wang, A., Behren, A., Antaw, F., Wang, Y., Sina, A.A.I., and Trau, M. (2021). *In situ* single cell proteomics reveals circulating tumor cell heterogeneity during treatment. ACS Nano 15, 11231–11243.
Ring, A., Nguyen-Sträuli, B.D., Wicki, A., and Aceto, N. (2023). Biology, vulnerabilities and clinical applications of circulating tumour cells. Nat Rev Cancer 23, 95–111.
Risso, D., Perraudeau, F., Gribkova, S., Dudoit, S., and Vert, J.P. (2018). A general and flexible method for signal extraction from single-cell RNA-seq data. Nat Commun 9, 284.
Rodriguez-Meira, A., Buck, G., Clark, S.A., Povinelli, B.J., Alcolea, V., Louka, E., McGowan, S., Hamblin, A., Sousos, N., Barkas, N., et al. (2019). Unravelling intratumoral heterogeneity through high-sensitivity single-cell mutational analysis and parallel RNA sequencing. Mol Cell 73, 1292–1305.e8.
Rodriques, S.G., Stickels, R.R., Goeva, A., Martin, C.A., Murray, E., Vanderburg, C.R., Welch, J., Chen, L.M., Chen, F., and Macosko, E.Z. (2019). Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. Science 363, 1463–1467.
Rood, J.E., Maartens, A., Hupalowska, A., Teichmann, S.A., and Regev, A. (2022). Impact of the Human Cell Atlas on medicine. Nat Med 28, 2486–2496.
Roohani, Y., Huang, K., and Leskovec, J. (2024). Predicting transcriptional outcomes of novel multigene perturbations with GEARS. Nat Biotechnol 42, 927–935.
Rosen, Y., Brbić, M., Roohani, Y., Swanson, K., Li, Z., and Leskovec, J. (2024). Toward universal cell embeddings: integrating single-cell RNA-seq datasets across species with SATURN. Nat Methods 21, 1492–1500.
Rosen, Y., Roohani, Y., Agarwal, A., Samotorčan, L., Quake, S.R., and Leskovec, J. (2023). Universal cell embeddings: a foundation model for cell biology. bioRxiv, 2023.2011.2028.568918.
Rosenberg, A.B., Roco, C.M., Muscat, R.A., Kuchina, A., Sample, P., Yao, Z., Graybuck, L.T., Peeler, D.J., Mukherjee, S., Chen, W., et al. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science 360, 176–182.
Rotem, A., Ram, O., Shoresh, N., Sperling, R.A., Goren, A., Weitz, D.A., and Bernstein, B.E. (2015). Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat Biotechnol 33, 1165–1172.
Rouhanifard, S.H., Mellis, I.A., Dunagin, M., Bayatpour, S., Jiang, C.L., Dardani, I., Symmons, O., Emert, B., Torre, E., Cote, A., et al. (2019). ClampFISH detects individual nucleic acid molecules using click chemistry-based amplification. Nat Biotechnol 37, 84–89.
Rousseeuw, P.J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20, 53–65.
Rubin, A.J., Parker, K.R., Satpathy, A.T., Qi, Y., Wu, B., Ong, A.J., Mumbach, M.R., Ji, A.L., Kim, D.S., Cho, S.W., et al. (2019). Coupled single-cell CRISPR screening and epigenomic profiling reveals causal gene regulatory networks. Cell 176, 361–376.e17.
Ruohan, W., Yuwei, Z., Mengbo, W., Xikang, F., Jianping, W., and Shuai Cheng, L. (2022). Resolving single-cell copy number profiling for large datasets. Brief Bioinf 23, bbac264.
Russell, A.J.C., Weir, J.A., Nadaf, N.M., Shabet, M., Kumar, V., Kambhampati, S., Raichur, R., Marrero, G.J., Liu, S., Balderrama, K.S., et al. (2024). Slide-tags enables single-nucleus barcoding for multimodal spatial genomics. Nature 625, 101–109.
Ryu, Y., Han, G.H., Jung, E., and Hwang, D. (2023). Integration of single-cell RNA-Seq datasets: a review of computational methods. Molecules Cells 46, 106–119.
Salek, M., Li, N., Chou, H.P., Saini, K., Jovic, A., Jacobs, K.B., Johnson, C., Lu, V., Lee, E.J., Chang, C., et al. (2023). COSMOS: a platform for real-time morphology-based, label-free cell sorting using deep learning. Commun Biol 6, 971.
Salmen, F., De Jonghe, J., Kaminski, T.S., Alemany, A., Parada, G.E., Verity-Legg, J., Yanagida, A., Kohler, T.N., Battich, N., van den Brekel, F., et al. (2022). High-throughput total RNA sequencing in single cells using VASA-seq. Nat Biotechnol 40, 1780–1793.
Sampath Kumar, A., Tian, L., Bolondi, A., Hernández, A.A., Stickels, R., Kretzmer, H., Murray, E., Wittler, L., Walther, M., Barakat, G., et al. (2023). Spatiotemporal transcriptomic maps of whole mouse embryos at the onset of organogenesis. Nat Genet 55, 1176–1185.
Sankaran, V.G., Weissman, J.S., and Zon, L.I. (2022). Cellular barcoding to decipher clonal dynamics in disease. Science 378, eabm5874.
Sasagawa, Y., Nikaido, I., Hayashi, T., Danno, H., Uno, K.D., Imai, T., and Ueda, H.R. (2013). Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. Genome Biol 14, R31.
Sashittal, P., Zhang, H., Iacobuzio-Donahue, C.A., and Raphael, B.J. (2023). ConDoR: tumor phylogeny inference with a copy-number constrained mutation loss model. Genome Biol 24, 272.
Sasse, A., Ng, B., Spiro, A.E., Tasaki, S., Bennett, D.A., Gaiteri, C., De Jager, P.L., Chikina, M., and Mostafavi, S. (2023). Benchmarking of deep neural networks for predicting personal gene expression from DNA sequence highlights shortcomings.

Nat Genet 55, 2060–2064.
Satpathy, A.T., Granja, J.M., Yost, K.E., Qi, Y., Meschi, F., McDermott, G.P., Olsen, B. N., Mumbach, M.R., Pierce, S.E., Corces, M.R., et al. (2019). Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. Nat Biotechnol 37, 925–936.
Schede, H.H., Schneider, C.G., Stergiadou, J., Borm, L.E., Ranjak, A., Yamawaki, T.M., David, F.P.A., Lönnerberg, P., Tosches, M.A., Codeluppi, S., et al. (2021). Spatial tissue profiling by imaging-free molecular tomography. Nat Biotechnol 39, 968–977.
Schott, M., León-Periñán, D., Splendiani, E., Strenger, L., Licha, J.R., Pentimalli, T.M., Schallenberg, S., Alles, J., Samut Tagliaferro, S., Boltengagen, A., et al. (2024). Open-ST: high-resolution spatial transcriptomics in 3D. Cell 187, 3953–3972.e26.
Schreibing, F., and Kramann, R. (2022). Mapping the human kidney using single-cell genomics. Nat Rev Nephrol 18, 347–360.
Schulz, D., Zanotelli, V.R.T., Fischer, J.R., Schapiro, D., Engler, S., Lun, X.K., Jackson, H.W., and Bodenmiller, B. (2018). Simultaneous multiplexed imaging of mRNA and proteins with subcellular resolution in breast cancer tissue samples by mass cytometry. Cell Syst 6, 25–36.e5.
Schwessinger, R., Deasy, J., Woodruff, R.T., Young, S., and Branson, K.M. (2023). Single-cell gene expression prediction from DNA sequence at large contexts. bioRxiv, 2023.2007.2026.550634.
Sender, R., Weiss, Y., Navon, Y., Milo, I., Azulay, N., Keren, L., Fuchs, S., Ben-Zvi, D., Noor, E., and Milo, R. (2023). The total mass, number, and distribution of immune cells in the human body. Proc Natl Acad Sci USA 120, e2308511120.
Seninge, L., Anastopoulos, I., Ding, H., and Stuart, J. (2021). VEGA is an interpretable generative model for inferring biological network activity in single-cell transcriptomics. Nat Commun 12, 5684.
Shah, S., Lubeck, E., Schwarzkopf, M., He, T.F., Greenbaum, A., Sohn, C.H., Lignell, A., Choi, H.M.T., Gradinaru, V., Pierce, N.A., et al. (2016). Single-molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding and clearing. Development 143, 2862–2867.
Shalek, A.K., Satija, R., Shuga, J., Trombetta, J.J., Gennert, D., Lu, D., Chen, P., Gertner, R.S., Gaublomme, J.T., Yosef, N., et al. (2014). Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature 510, 363–369.
Shao, Y., Wang, X.B., Zhang, J.J., Li, M.L., Wu, S.S., Ma, X.Y., Wang, X., Zhao, H.F., Li, Y., Zhu, H.H., et al. (2020). Genome and single-cell RNA-sequencing of the earthworm *Eisenia andrei* identifies cellular mechanisms underlying regeneration. Nat Commun 11, 2656.
Sharma, P., Goswami, S., Raychaudhuri, D., Siddiqui, B.A., Singh, P., Nagarajan, A., Liu, J., Subudhi, S.K., Poon, C., Gant, K.L., et al. (2023). Immune checkpoint therapy—current perspectives and future directions. Cell 186, 1652–1669.
Shekhar, K., Lapan, S.W., Whitney, I.E., Tran, N.M., Macosko, E.Z., Kowalczyk, M., Adiconis, X., Levin, J.Z., Nemesh, J., Goldman, M., et al. (2016). Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. Cell 166, 1308–1323.e30.
Shen-Orr, S.S., Tibshirani, R., Khatri, P., Bodian, D.L., Staedtler, F., Perry, N.M., Hastie, T., Sarwal, M.M., Davis, M.M., and Butte, A.J. (2010). Cell type-specific gene expression differences in complex tissues. Nat Methods 7, 287–289.
Shen, H., Liu, J., Hu, J., Shen, X., Zhang, C., Wu, D., Feng, M., Yang, M., Li, Y., Yang, Y., et al. (2023). Generative pretraining from large-scale transcriptomes for single-cell deciphering. iScience 26, 106536.
Shen, H., Shen, X., Feng, M., Wu, D., Zhang, C., Yang, Y., Yang, M., Hu, J., Liu, J., Wang, W., et al. (2022). A universal approach for integrating super large-scale single-cell transcriptomes by exploring gene rankings. Brief Bioinf 23, bbab573.
Sheng, K., Cao, W., Niu, Y., Deng, Q., and Zong, C. (2017). Effective detection of variation in single-cell transcriptomes using MATQ-seq. Nat Methods 14, 267–270.
Shi, H., He, Y., Zhou, Y., Huang, J., Maher, K., Wang, B., Tang, Z., Luo, S., Tan, P., Wu, M., et al. (2023). Spatial atlas of the mouse central nervous system at molecular resolution. Nature 622, 552–561.
Shrestha, B. (2020). Single-Cell Metabolomics by Mass Spectrometry. Methods Mol Biol 2064, 1–8.
Sibai, M., Altuntaş, E., Süzek, B.ı.E., Şahin, B., Parlayan, C., Öztürk, G., Baykal, A.T.ı., and Demircan, T. (2020). Comparison of protein expression profile of limb regeneration between neotenic and metamorphic axolotl. Biochem Biophys Res Commun 522, 428–434.
Sikkema, L., Ramírez-Suástegui, C., Strobl, D.C., Gillett, T.E., Zappia, L., Madissoon, E., Markov, N.S., Zaragosi, L.E., Ji, Y., Ansari, M., et al. (2023). An integrated cell atlas of the lung in health and disease. Nat Med 29, 1563–1577.
Siletti, K., Hodge, R., Albiach, A.M., Hu, L., Lee, K.W., Lönnerberg, P., Bakken, T., Ding, S.L., Clark, M., Casper, T., et al. (2022). Transcriptomic diversity of cell types across the adult human brain. bioRxiv, 2022.2010.2012.511898.
Smallwood, S.A., Lee, H.J., Angermueller, C., Krueger, F., Saadeh, H., Peat, J., Andrews, S.R., Stegle, O., Reik, W., and Kelsey, G. (2014). Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Methods 11, 817–820.
Smith, R.J., Zhang, H., Hu, S.S., Yung, T., Francis, R., Lee, L., Onaitis, M.W., Dirks, P. B., Zang, C., and Kim, T.H. (2022). Single-cell chromatin profiling of the primitive gut tube reveals regulatory dynamics underlying lineage fate decisions. Nat Commun 13, 2965.
Snyder, M.P., Lin, S., Posgai, A., Atkinson, M., Regev, A., Rood, J., Rozenblatt-Rosen, O., Gaffney, L., Hupalowska, A., Satija, R., et al. (2019). The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. Nature 574, 187–192.
Sokolova, K., Theesfeld, C.L., Wong, A.K., Zhang, Z., Dolinski, K., and Troyanskaya, O. G. (2023). Atlas of primary cell-type-specific sequence models of gene expression and variant effects. Cell Rep Methods 3, 100580.
Song, L., Zhang, Z., Grasfeder, L.L., Boyle, A.P., Giresi, P.G., Lee, B.K., Sheffield, N.C., Gräf, S., Huss, M., Keefe, D., et al. (2011). Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. Genome Res 21, 1757–1767.
Song, Q., and Su, J. (2021). DSTG: Deconvoluting spatial transcriptomics data through graph-based artificial intelligence. Brief Bioinf 22, bbaa414.
Song, Q., Su, J., and Zhang, W. (2021). scGCN is a graph convolutional networks algorithm for knowledge transfer in single cell omics. Nat Commun 12, 3826.
Song, T., Dai, H., Wang, S., Wang, G., Zhang, X., Zhang, Y., and Jiao, L. (2022). TransCluster: a cell-type identification method for single-cell RNA-Seq data using deep learning based on transformer. Front Genet 13, 1038919.
Song, Y., Miao, Z., Brazma, A., and Papatheodorou, I. (2023). Benchmarking strategies for cross-species integration of single-cell RNA sequencing data. Nat Commun 14, 6495.
Sos, B.C., Fung, H.L., Gao, D.R., Osothprarop, T.F., Kia, A., He, M.M., and Zhang, K. (2016). Characterization of chromatin accessibility with a transposome hypersensitive sites sequencing (THS-seq) assay. Genome Biol 17, 20.
Srivatsan, S.R., McFaline-Figueroa, J.L., Ramani, V., Saunders, L., Cao, J., Packer, J., Pliner, H.A., Jackson, D.L., Daza, R.M., Christiansen, L., et al. (2020). Massively multiplex chemical transcriptomics at single-cell resolution. Science 367, 45–51.
Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science 353, 78–82.
Stewart, B.J., Ferdinand, J.R., and Clatworthy, M.R. (2020). Using single-cell technologies to map the human immune system—implications for nephrology. Nat Rev Nephrol 16, 112–128.
Stickels, R.R., Murray, E., Kumar, P., Li, J., Marshall, J.L., Di Bella, D.J., Arlotta, P., Macosko, E.Z., and Chen, F. (2021). Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. Nat Biotechnol 39, 313–319.
Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P. K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. Nat Methods 14, 865–868.
Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B.Z., Mauck Iii, W.M., Smibert, P., and Satija, R. (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. Genome Biol 19, 224.
Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genomics 19, 477.
Stuart, T., Hao, S., Zhang, B., Mekerishvili, L., Landau, D.A., Maniatis, S., Satija, R., and Raimondi, I. (2022). Nanobody-tethered transposition allows for multifactorial chromatin profiling at single-cell resolution. bioRxiv, 2022.03.08.483436.
Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. Nat Methods 18, 1333–1341.
Stubbington, M.J.T., Rozenblatt-Rosen, O., Regev, A., and Teichmann, S.A. (2017). Single-cell transcriptomics to explore the immune system in health and disease. Science 358, 58–63.
Sui, X., Lo, J.A., Luo, S., He, Y., Tang, Z., Lin, Z., Zhou, Y., Wang, W.X., Liu, J., and Wang, X. (2024). Scalable spatial single-cell transcriptomics and translatomics in 3D thick tissue blocks. bioRxiv, 2024.2008.2005.606553.
Sun, T., Wu, R., and Ming, L. (2019). The role of m^{6}A RNA methylation in cancer. Biomed Pharmacother 112, 108613.
Sun, Z., Tang, Y., Zhang, Y., Fang, Y., Jia, J., Zeng, W., and Fang, D. (2021a). Joint single-cell multiomic analysis in Wnt3a induced asymmetric stem cell division. Nat Commun 12, 5941.
Sun, Z., Vaisvila, R., Hussong, L.M., Yan, B., Baum, C., Saleh, L., Samaranayake, M., Guan, S., Dai, N., Corrêa Jr., I.R., et al. (2021b). Nondestructive enzymatic deamination enables single-molecule long-read amplicon sequencing for the determination of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Genome Res 31, 291–300.
Suo, C., Dann, E., Goh, I., Jardine, L., Kleshchevnikov, V., Park, J.E., Botting, R.A., Stephenson, E., Engelbert, J., Tuong, Z.K., et al. (2022). Mapping the developing

human immune system across organs. Science 376, eabo0510.
Suvilesh, K.N., Nussbaum, Y.I., Radhakrishnan, V., Manjunath, Y., Avella, D.M., Staveley-O'Carroll, K.F., Kimchi, E.T., Chaudhuri, A.A., Shyu, C.R., Li, G., et al. (2022). Tumorigenic circulating tumor cells from xenograft mouse models of non-metastatic NSCLC patients reveal distinct single cell heterogeneity and drug responses. Mol Cancer 21, 73.
Svensson, V., da-Veiga-Beltrame, E., and Pachter, L. (2020). A curated database reveals trends in single-cell transcriptomics. Database 2020, baaa073.
Tabula Muris, C., Overall, c., Logistical, c., Organ, c., processing, Library, p., sequencing, Computational data, a., Cell type, a., Writing, g., et al. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature 562, 367–372.
Takei, Y., Yun, J., Zheng, S., Ollikainen, N., Pierson, N., White, J., Shah, S., Thomassie, J., Suo, S., Eng, C.H.L., et al. (2021). Integrated spatial genomics reveals global architecture of single nuclei. Nature 590, 344–350.
Takeshima, H., and Ushijima, T. (2019). Accumulation of genetic and epigenetic alterations in normal cells and cancer risk. npj Precis Onc 3, 7.
Tan, L., Ma, W., Wu, H., Zheng, Y., Xing, D., Chen, R., Li, X., Daley, N., Deisseroth, K., and Xie, X.S. (2021). Changes in genome architecture and transcriptional dynamics progress independently of sensory experience during post-natal brain development. Cell 184, 741–758.e17.
Tan, L., Xing, D., Chang, C.H., Li, H., and Xie, X.S. (2018). Three-dimensional genome structures of single diploid human cells. Science 361, 924–928.
Tan, S.J., Li, Q., and Lim, C.T. (2010). Manipulation and isolation of single cells and nuclei. Methods Cell Biol 98, 79–96.
Tanay, A., and Regev, A. (2017). Scaling single-cell genomics from phenomenology to mechanism. Nature 541, 331–338.
Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods 6, 377–382.
Tang, F., Li, J., Qi, L., Liu, D., Bo, Y., Qin, S., Miao, Y., Yu, K., Hou, W., Li, J., et al. (2023a). A pan-cancer single-cell panorama of human natural killer cells. Cell 186, 4235–4251.e20.
Tang, L. (2021). Multiomics sequencing goes spatial. Nat Methods 18, 31.
Tang, W., Bertaux, F., Thomas, P., Stefanelli, C., Saint, M., Marguerat, S., Shahrezaei, V., and Kelso, J. (2020). bayNorm: Bayesian gene expression recovery, imputation and normalization for single-cell RNA-sequencing data. Bioinformatics 36, 1174–1181.
Tang, W., Wen, H., Liu, R., Ding, J., Jin, W., Xie, Y., Liu, H., and Tang, J. (2023b). Single-Cell Multimodal Prediction via Transformers. arXiv, 2303.00233.
Tang, X., Chen, J., Zhang, X., Liu, X., Xie, Z., Wei, K., Qiu, J., Ma, W., Lin, C., and Ke, R. (2023c). Improved *in situ* sequencing for high-resolution targeted spatial transcriptomic analysis in tissue sections. J Genet Genomics 50, 652–660.
Tarashansky, A.J., Musser, J.M., Khariton, M., Li, P., Arendt, D., Quake, S.R., and Wang, B. (2021). Mapping single-cell atlases throughout Metazoa unravels cell type evolution. eLife 10, 10.7554–66747.
Tedesco, M., Giannese, F., Lazarević, D., Giansanti, V., Rosano, D., Monzani, S., Catalano, I., Grassi, E., Zanella, E.R., Botrugno, O.A., et al. (2022). Chromatin velocity reveals epigenetic dynamics by single-cell profiling of heterochromatin and euchromatin. Nat Biotechnol 40, 235–244.
The Tabula Sapiens Consortium, Jones, R.C., Karkanias, J., Krasnow, M.A., Pisco, A. O., Quake, S.R., Salzman, J., Yosef, N., Bulthaup, B., Brown, P., et al. (2022). The Tabula Sapiens: a multiple-organ, single-cell transcriptomic atlas of humans. Science 376, eabl4896.
Theodoris, C.V., Xiao, L., Chopra, A., Chaffin, M.D., Al Sayed, Z.R., Hill, M.C., Mantineo, H., Brydon, E.M., Zeng, Z., Liu, X.S., et al. (2023). Transfer learning enables predictions in network biology. Nature 618, 616–624.
Tian, F., Zhou, F., Li, X., Ma, W., Wu, H., Yang, M., Chapman, A.R., Lee, D.F., Tan, L., Xing, D., et al. (2020). Genomic architecture of cells in tissues (GeACT): study of human mid-gestation fetus. bioRxiv, 2020.2004.2012.038000.
Tian, L., Chen, F., and Macosko, E.Z. (2023a). The expanding vistas of spatial transcriptomics. Nat Biotechnol 41, 773–782.
Tian, W., Zhou, J., Bartlett, A., Zeng, Q., Liu, H., Castanon, R.G., Kenworthy, M., Altshul, J., Valadon, C., Aldridge, A., et al. (2023b). Single-cell DNA methylation and 3D genome architecture in the human brain. Science 382, eadf5357.
Tran, D., Nguyen, H., Tran, B., La Vecchia, C., Luu, H.N., and Nguyen, T. (2021). Fast and precise single-cell data analysis using a hierarchical autoencoder. Nat Commun 12, 1029.
Tran, H.T.N., Ang, K.S., Chevrier, M., Zhang, X., Lee, N.Y.S., Goh, M., and Chen, J. (2020). A benchmark of batch-effect correction methods for single-cell RNA sequencing data. Genome Biol 21, 12.
Tsompana, M., and Buck, M.J. (2014). Chromatin accessibility: a window into the genome. Epigenet Chromatin 7, 33.
Tyser, R.C.V., Mahammadov, E., Nakanoh, S., Vallier, L., Scialdone, A., and Srinivas, S. (2021). Single-cell transcriptomic characterization of a gastrulating human embryo. Nature 600, 285–289.
Vallejo, A.F., Harvey, K., Wang, T., Wise, K., Butler, L.M., Polo, J., Plummer, J., Swarbrick, A., and Martelotto, L.G. (2022). snPATHO-seq: unlocking the FFPE archives for single nucleus RNA profiling. bioRxiv, 2022.2008.2023.505054.
Van de Sande, B., Lee, J.S., Mutasa-Gottgens, E., Naughton, B., Bacon, W., Manning, J., Wang, Y., Pollard, J., Mendez, M., Hill, J., et al. (2023). Applications of single-cell RNA sequencing in drug discovery and development. Nat Rev Drug Discov 22, 496–520.
van Dijk, D., Sharma, R., Nainys, J., Yim, K., Kathail, P., Carr, A.J., Burdziak, C., Moon, K.R., Chaffer, C.L., Pattabiraman, D., et al. (2018). Recovering gene interactions from single-cell data using data diffusion. Cell 174, 716–729.e27.
Vandereyken, K., Sifrim, A., Thienpont, B., and Voet, T. (2023). Methods and applications for single-cell and spatial multi-omics. Nat Rev Genet 24, 494–515.
VanInsberghe, M., van den Berg, J., Andersson-Rolf, A., Clevers, H., and van Oudenaarden, A. (2021). Single-cell Ribo-seq reveals cell cycle-dependent translational pausing. Nature 597, 561–565.
Varrone, M., Tavernari, D., Santamaria-Martínez, A., Walsh, L.A., and Ciriello, G. (2024). CellCharter reveals spatial cell niches associated with tissue remodeling and cell plasticity. Nat Genet 56, 74–84.
Velmeshev, D., Perez, Y., Yan, Z., Valencia, J.E., Castaneda-Castellanos, D.R., Wang, L., Schirmer, L., Mayer, S., Wick, B., Wang, S., et al. (2023). Single-cell analysis of prenatal and postnatal human cortical development. Science 382, eadf0834.
Vento-Tormo, R., Efremova, M., Botting, R.A., Turco, M.Y., Vento-Tormo, M., Meyer, K.B., Park, J.E., Stephenson, E., POLAŃSKI, K., Goncalves, A., et al. (2018). Single-cell reconstruction of the early maternal-fetal interface in humans. Nature 563, 347–353.
Vickovic, S., Eraslan, G., Salmén, F., Klughammer, J., Stenbeck, L., Schapiro, D., Äijö, T., Bonneau, R., Bergenstråhle, L., Navarro, J.F., et al. (2019). High-definition spatial transcriptomics for *in situ* tissue profiling. Nat Methods 16, 987–990.
Vu, T.N., Nguyen, H.N., Calza, S., Kalari, K.R., Wang, L., Pawitan, Y., and Berger, B. (2019). Cell-level somatic mutation detection from single-cell RNA sequencing. Bioinformatics 35, 4679–4687.
Wagner, A., Wang, C., Fessler, J., DeTomaso, D., Avila-Pacheco, J., Kaminski, J., Zaghouani, S., Christian, E., Thakore, P., Schellhaass, B., et al. (2021). Metabolic modeling of single Th17 cells reveals regulators of autoimmunity. Cell 184, 4168–4185.e21.
Wagner, F., Yan, Y., and Yanai, I. (2018). K-nearest neighbor smoothing for high-throughput single-cell RNA-Seq data. bioRxiv, 217737.
Walker, A., and Parkhill, J. (2008). Single-cell genomics. Nat Rev Microbiol 6, 176–177.
Wan, X., Xiao, J., Tam, S.S.T., Cai, M., Sugimura, R., Wang, Y., Wan, X., Lin, Z., Wu, A.R., and Yang, C. (2023). Integrating spatial and single-cell transcriptomics data using deep generative models with SpatialScope. Nat Commun 14, 7848.
Wang, B., Lin, A.E., Yuan, J., Novak, K.E., Koch, M.D., Wingreen, N.S., Adamson, B., and Gitai, Z. (2023a). Single-cell massively-parallel multiplexed microbial sequencing (M3-seq) identifies rare bacterial populations and profiles phage infection. Nat Microbiol 8, 1846–1862.
Wang, F., Wang, Q., Mohanty, V., Liang, S., Dou, J., Han, J., Minussi, D.C., Gao, R., Ding, L., Navin, N., et al. (2021a). MEDALT: Single-cell copy number lineage tracing enabling gene discovery. Genome Biol 22, 70.
Wang, G., Moffitt, J.R., and Zhuang, X. (2018a). Multiplexed imaging of high-density libraries of RNAs with MERFISH and expansion microscopy. Sci Rep 8, 4847.
Wang, J., Chen, L., Chen, Z., and Zhang, W. (2015). RNA-seq based transcriptomic analysis of single bacterial cells. Integr Biol 7, 1466–1476.
Wang, J., Sun, H., Jiang, M., Li, J., Zhang, P., Chen, H., Mei, Y., Fei, L., Lai, S., Han, X., et al. (2021b). Tracing cell-type evolution by cross-species comparison of cell atlases. Cell Rep 34, 108803.
Wang, Q., Xiong, H., Ai, S., Yu, X., Liu, Y., Zhang, J., and He, A. (2019a). CoBATCH for high-throughput single-cell epigenomic profiling. Mol Cell 76, 206–216.e7.
Wang, R., and Jiang, Y. (2022). Copy number variation detection by single-cell DNA sequencing with SCOPE. Methods Mol Biol 2493, 279– 288.
Wang, R., Lin, D.Y., and Jiang, Y. (2020a). SCOPE: a normalization and copy-number estimation method for single-cell DNA sequencing. Cell Syst 10, 445–452.e6.
Wang, R., Zhang, P., Wang, J., Ma, L., E, W., Suo, S., Jiang, M., Li, J., Chen, H., Sun, H., et al. (2023b). Construction of a cross-species cell landscape at single-cell level. Nucleic Acids Res 51, 501–516.
Wang, S., Zheng, Y., Li, J., Yu, Y., Zhang, W., Song, M., Liu, Z., Min, Z., Hu, H., Jing, Y., et al. (2020b). Single-cell transcriptomic atlas of primate ovarian aging. Cell 180, 585–600.e19.
Wang, X., Allen, W.E., Wright, M.A., Sylwestrak, E.L., Samusik, N., Vesuna, S., Evans, K., Liu, C., Ramakrishnan, C., Liu, J., et al. (2018b). Three-dimensional intact-tissue sequencing of single-cell transcriptional states. Science 361, eaat5691.
Wang, X., Park, J., Susztak, K., Zhang, N.R., and Li, M. (2019b). Bulk tissue cell type

deconvolution with multi-subject single-cell expression reference. Nat Commun 10, 380.
Wang, Y., Eddison, M., Fleishman, G., Weigert, M., Xu, S., Wang, T., Rokicki, K., Goina, C., Henry, F.E., Lemire, A.L., et al. (2021c). EASI-FISH for thick tissue defines lateral hypothalamus spatio-molecular organization. Cell 184, 6361–6377.e24.
Wang, Y., Fan, J.L., Melms, J.C., Amin, A.D., Georgis, Y., Barrera, I., Ho, P., Tagore, S., Abril-Rodríguez, G., He, S., et al. (2023c). Multimodal single-cell and whole-genome sequencing of small, frozen clinical specimens. Nat Genet 55, 19–25.
Wang, Y., Yuan, P., Yan, Z., Yang, M., Huo, Y., Nie, Y., Zhu, X., Qiao, J., and Yan, L. (2021d). Single-cell multiomics sequencing reveals the functional regulatory landscape of early embryos. Nat Commun 12, 1247.
Weber, E.W., Maus, M.V., and Mackall, C.L. (2020). The emerging landscape of immune cell therapies. Cell 181, 46–62.
Weber, L.L., Sashittal, P., and El-Kebir, M. (2021). doubletD: Detecting doublets in single-cell DNA sequencing data. Bioinformatics 37, i214–i221.
Weber, L.L., Zhang, C., Ochoa, I., El-Kebir, M., and Przytycka, T.M. (2023). Phertilizer: growing a clonal tree from ultra-low coverage single-cell DNA sequencing of tumors. PLoS Comput Biol 19, e1011544.
Wei, X., Fu, S., Li, H., Liu, Y., Wang, S., Feng, W., Yang, Y., Liu, X., Zeng, Y.Y., Cheng, M., et al. (2022). Single-cell Stereo-seq reveals induced progenitor cells involved in axolotl brain regeneration. Science 377, eabp9444.
Weinstein, J.A., Regev, A., and Zhang, F. (2019). DNA microscopy: optics-free spatio-genetic imaging by a stand-alone chemical reaction. Cell 178, 229–241.e16.
Welch, J., Kozareva, V., Ferreira, A., Vanderburg, C., Martin, C., and Macosko, E. (2018). Integrative inference of brain cell similarities and differences from single-cell genomics. bioRxiv, 459891.
Wen, L., and Tang, F. (2022). Recent advances in single-cell sequencing technologies. Precision Clin Med 5, pbac002.
Wen, N., Zhao, Z., Fan, B., Chen, D., Men, D., Wang, J., and Chen, J. (2016). Development of droplet microfluidics enabling high-throughput single-cell analysis. Molecules 21, 881.
Weninger, A., and Arlotta, P. (2023). A family portrait of human brain cells. Science 382, 168–169.
Westermann, A.J., and Vogel, J. (2021). Cross-species RNA-seq for deciphering host-microbe interactions. Nat Rev Genet 22, 361–378.
Wherry, E.J. (2011). T cell exhaustion. Nat Immunol 12, 492–499.
Winkler, E.A., Kim, C.N., Ross, J.M., Garcia, J.H., Gil, E., Oh, I., Chen, L.Q., Wu, D., Catapano, J.S., Raygor, K., et al. (2022). A single-cell atlas of the normal and malformed human brain vasculature. Science 375, eabi7377.
Winkler, I., Tolkachov, A., Lammers, F., Lacour, P., Daugelaite, K., Schneider, N., Koch, M.L., Panten, J., Grünschläger, F., Poth, T., et al. (2024). The cycling and aging mouse female reproductive tract at single-cell resolution. Cell 187, 981–998.e25.
Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 19, 15.
Wood, L.D., Parsons, D.W., Jones, S., Lin, J., Sjöblom, T., Leary, R.J., Shen, D., Boca, S. M., Barber, T., Ptak, J., et al. (2007). The genomic landscapes of human breast and colorectal cancers. Science 318, 1108–1113.
Wu, C., Simonetti, M., Rossell, C., Mignardi, M., Mirzazadeh, R., Annaratone, L., Marchiò, C., Sapino, A., Bienko, M., Crosetto, N., et al. (2018). RollFISH achieves robust quantification of single-molecule RNA biomarkers in paraffin-embedded tumor tissue samples. Commun Biol 1, 209.
Wu, H., Zhang, J., Jian, F., Chen, J.P., Zheng, Y., Tan, L., and Sunney Xie, X. (2024a). Simultaneous single-cell three-dimensional genome and gene expression profiling uncovers dynamic enhancer connectivity underlying olfactory receptor choice. Nat Methods 21, 974–982.
Wu, S.J., Furlan, S.N., Mihalas, A.B., Kaya-Okur, H.S., Feroze, A.H., Emerson, S.N., Zheng, Y., Carson, K., Cimino, P.J., Keene, C.D., et al. (2021). Single-cell CUT&Tag analysis of chromatin modifications in differentiation and tumor progression. Nat Biotechnol 39, 819–824.
Wu, T., Cheng, A.Y., Zhang, Y., Xu, J., Wu, J., Wen, L., Li, X., Liu, B., Dou, X., Wang, P., et al. (2024b). KARR-seq reveals cellular higher-order RNA structures and RNA-RNA interactions. Nat Biotechnol 42, 1909–1920.
Xia, C., Fan, J., Emanuel, G., Hao, J., and Zhuang, X. (2019). Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. Proc Natl Acad Sci USA 116, 19490–19499.
Xia, K., Sun, H.X., Li, J., Li, J., Zhao, Y., Chen, L., Qin, C., Chen, R., Chen, Z., Liu, G., et al. (2022). The single-cell stereo-seq reveals region-specific cell subtypes and transcriptome profiling in Arabidopsis leaves. Dev Cell 57, 1299–1310.e4.
Xiao, Y., Wang, J., Li, J., Zhang, P., Li, J., Zhou, Y., Zhou, Q., Chen, M., Sheng, X., Liu, Z., et al. (2023). An analytical framework for decoding cell type-specific genetic variation of gene regulation. Nat Commun 14, 3884.
Xing, Q.R., Farran, C.A.E., Zeng, Y.Y., Yi, Y., Warrier, T., Gautam, P., Collins, J.J., Xu, J., Dröge, P., Koh, C.G., et al. (2020). Parallel bimodal single-cell sequencing of transcriptome and chromatin accessibility. Genome Res 30, 1027–1039.
Xiong, H., Luo, Y., Wang, Q., Yu, X., and He, A. (2021). Single-cell joint detection of chromatin occupancy and transcriptome enables higher-dimensional epigenomic reconstructions. Nat Methods 18, 652–660.
Xiong, L., Chen, T., and Kellis, M. (2023). scCLIP: multi-modal single-cell contrastive learning integration pre-training. In: NeurIPS 2023 AI for Science Workshop. Available from URL: https://api.semanticscholar.org/CorpusID:268891994.
Xiong, L., Xu, K., Tian, K., Shao, Y., Tang, L., Gao, G., Zhang, M., Jiang, T., and Zhang, Q.C. (2019). SCALE method for single-cell ATAC-seq analysis via latent feature extraction. Nat Commun 10, 4576.
Xiong, Z., Xu, K., Lin, Z., Kong, F., Wang, Q., Quan, Y., Sha, Q., Li, F., Zou, Z., Liu, L., et al. (2022). Ultrasensitive Ribo-seq reveals translational landscapes during mammalian oocyte-to-embryo transition and pre-implantation development. Nat Cell Biol 24, 968–980.
Xu, C., Jin, X., Wei, S., Wang, P., Luo, M., Xu, Z., Yang, W., Cai, Y., Xiao, L., Lin, X., et al. (2022a). DeepST: Identifying spatial domains in spatial transcriptomics by deep learning. Nucleic Acids Res 50, e131.
Xu, C., Lopez, R., Mehlman, E., Regier, J., Jordan, M.I., and Yosef, N. (2021a). Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. Mol Syst Biol 17, e9620.
Xu, C., Prete, M., Webb, S., Jardine, L., Stewart, B.J., Hoo, R., He, P., Meyer, K.B., and Teichmann, S.A. (2023a). Automatic cell-type harmonization and integration across Human Cell Atlas datasets. Cell 186, 5876–5891.e20.
Xu, J., Zhang, A., Liu, F., Chen, L., and Zhang, X. (2023b). CIForm as a Transformer-based model for cell-type annotation of large-scale single-cell RNA-seq data. Brief Bioinf 24, bbad195.
Xu, J., Zhang, A., Liu, F., Zhang, X., and Birol, I. (2023c). STGRNS: An interpretable transformer-based method for inferring gene regulatory networks from single-cell transcriptomic data. Bioinformatics 39, btad165.
Xu, W., Wen, Y., Liang, Y., Xu, Q., Wang, X., Jin, W., and Chen, X. (2021b). A plate-based single-cell ATAC-seq workflow for fast and robust profiling of chromatin accessibility. Nat Protoc 16, 4084–4107.
Xu, W., Yang, W., Zhang, Y., Chen, Y., Hong, N., Zhang, Q., Wang, X., Hu, Y., Song, K., Jin, W., et al. (2022b). ISSAAC-seq enables sensitive and flexible multimodal profiling of chromatin accessibility and gene expression in single cells. Nat Methods 19, 1243–1249.
Xu, Y., Zhang, T., Zhou, Q., Hu, M., Qi, Y., Xue, Y., Nie, Y., Wang, L., Bao, Z., and Shi, W. (2023d). A single-cell transcriptome atlas profiles early organogenesis in human embryos. Nat Cell Biol 25, 604–615.
Xu, Z., Wang, Y., Sheng, K., Rosenthal, R., Liu, N., Hua, X., Zhang, T., Chen, J., Song, M., Lv, Y., et al. (2023e). Droplet-based high-throughput single microbe RNA sequencing by smRandom-seq. Nat Commun 14, 5130.
Xu, Z., Zhang, T., Chen, H., Zhu, Y., Lv, Y., Zhang, S., Chen, J., Chen, H., Yang, L., Jiang, W., et al. (2023f). High-throughput single nucleus total RNA sequencing of formalin-fixed paraffin-embedded tissues by snRandom-seq. Nat Commun 14, 2734.
Xue, R., Zhang, Q., Cao, Q., Kong, R., Xiang, X., Liu, H., Feng, M., Wang, F., Cheng, J., Li, Z., et al. (2022). Liver tumour immune microenvironment subtypes and neutrophil heterogeneity. Nature 612, 141–147.
Yamada, S., Nagafuchi, Y., Wang, M., Ota, M., Hatano, H., Takeshima, Y., Okubo, M., Kobayashi, S., Sugimori, Y., Masahiro, N., et al. (2023). Immunomics analysis of rheumatoid arthritis identified precursor dendritic cells as a key cell subset of treatment resistance. Ann Rheum Dis 82, 809–819.
Yamagishi, M., Kuze, Y., Kobayashi, S., Nakashima, M., Morishima, S., Kawamata, T., Makiyama, J., Suzuki, K., Seki, M., Abe, K., et al. (2024). Mechanisms of action and resistance in histone methylation-targeted therapy. Nature 627, 221–228.
Yan, R., Gu, C., You, D., Huang, Z., Qian, J., Yang, Q., Cheng, X., Zhang, L., Wang, H., Wang, P., et al. (2021). Decoding dynamic epigenetic landscapes in human oocytes using single-cell multi-omics sequencing. Cell Stem Cell 28, 1641–1656.e7.
Yang, D., Wang, L.L., Dong, T.T., Shen, Y.H., Guo, X.S., Liu, C.Y., Liu, J., Zhang, P., Li, J., and Sun, Y.P. (2015). Progranulin promotes colorectal cancer proliferation and angiogenesis through TNFR2/Akt and ERK signaling pathways. Am J Cancer Res 5, 3085–3097.
Yang, F., Wang, W., Wang, F., Fang, Y., Tang, D., Huang, J., Lu, H., and Yao, J. (2022). scBERT as a large-scale pretrained deep language model for cell type annotation of single-cell RNA-seq data. Nat Mach Intell 4, 852–866.
Yang, S., Liu, C., Jiang, M., Liu, X., Geng, L., Zhang, Y., Sun, S., Wang, K., Yin, J., Ma, S., et al. (2024). A single-nucleus transcriptomic atlas of primate liver aging uncovers the pro-senescence role of SREBP2 in hepatocytes. Protein Cell 15, 98–120.
Yang, W., Wang, P., Luo, M., Cai, Y., Xu, C., Xue, G., Jin, X., Cheng, R., Que, J., Pang, F., et al. (2023a). DeepCCI: a deep learning framework for identifying cell-cell interactions from single-cell RNA sequencing data. Bioinformatics 39, btad596.

Yang, X., Liu, G., Feng, G., Bu, D., Wang, P., Jiang, J., Chen, S., Yang, Q., Zhang, Y., Man, Z., et al. (2023b). GeneCompass: deciphering universal gene regulatory mechanisms with knowledge-informed cross-species foundation model. bioRxiv, 2023.2009.2026.559542.
Yang, Y., Li, G., Qian, H., Wilhelmsen, K.C., Shen, Y., and Li, Y. (2021). SMNN: Batch effect correction for single-cell RNA-seq data via supervised mutual nearest neighbor detection. Brief Bioinf 22, bbaa097.
Yang, Z., Li, C., Fan, Z., Liu, H., Zhang, X., Cai, Z., Xu, L., Luo, J., Huang, Y., He, L., et al. (2017). Single-cell sequencing reveals variants in ARID1A, GPRC5A and MLL2 driving self-renewal of human bladder cancer stem cells. Eur Urology 71, 8–12.
Yao, H., Gao, C.C., Zhang, D., Xu, J., Song, G., Fan, X., Liang, D.B., Chen, Y.S., Li, Q., Guo, Y., et al. (2023). scm^{6}A-seq reveals single-cell landscapes of the dynamic m^{6}A during oocyte maturation and early embryonic development. Nat Commun 14, 315.
Yazar, S., Alquicira-Hernandez, J., Wing, K., Senabouth, A., Gordon, M.G., Andersen, S., Lu, Q., Rowson, A., Taylor, T.R.P., Clarke, L., et al. (2022). Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. Science 376, eabf3041.
Ye, F., Wang, J., Li, J., Mei, Y., and Guo, G. (2024a). Mapping cell atlases at the single-cell level. Adv Sci 11, e2305449.
Ye, F., Zhang, G., E., W., Chen, H., Yu, C., Yang, L., Fu, Y., Li, J., Fu, S., Sun, Z., et al. (2022). Construction of the axolotl cell landscape using combinatorial hybridization sequencing at single-cell resolution. Nat Commun 13, 4228.
Ye, F., Zhang, S., Fu, Y., Yang, L., Zhang, G., Wu, Y., Pan, J., Chen, H., Wang, X., Ma, L., et al. (2024b). Fast and flexible profiling of chromatin accessibility and total RNA expression in single nuclei using Microwell-seq3. Cell Discov 10, 33.
Ye, X., Huang, X., Fu, X., Zhang, X., Lin, R., Zhang, W., Zhang, J., and Lu, Y. (2023). Myeloid-like tumor hybrid cells in bone marrow promote progression of prostate cancer bone metastasis. J Hematol Oncol 16, 46.
Yeung, J., Florescu, M., Zeller, P., de Barbanson, B.A., Wellenstein, M.D., and van Oudenaarden, A. (2023). scChIX-seq infers dynamic relationships between histone modifications in single cells. Nat Biotechnol 41, 813–823.
Yin, K., Zhao, M., Lin, L., Chen, Y., Huang, S., Zhu, C., Liang, X., Lin, F., Wei, H., Zeng, H., et al. (2022). Well-paired-seq: a size-exclusion and locally quasi-static hydrodynamic microwell chip for single-cell RNA-Seq. Small Methods 6, e2200341.
Yin, Y., Jiang, Y., Lam, K.W.G., Berletch, J.B., Disteche, C.M., Noble, W.S., Steemers, F. J., Camerini-Otero, R.D., Adey, A.C., and Shendure, J. (2019). High-throughput single-cell sequencing with linear amplification. Mol Cell 76, 676–690.e10.
Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P.W., Levine, D.A., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 4, 2612.
Yost, K.E., Satpathy, A.T., Wells, D.K., Qi, Y., Wang, C., Kageyama, R., McNamara, K. L., Granja, J.M., Sarin, K.Y., Brown, R.A., et al. (2019). Clonal replacement of tumor-specific T cells following PD-1 blockade. Nat Med 25, 1251–1259.
You, Y., Fu, Y., Li, L., Zhang, Z., Jia, S., Lu, S., Ren, W., Liu, Y., Xu, Y., Liu, X., et al. (2024). Systematic comparison of sequencing-based spatial transcriptomic methods. Nat Methods 21, 1743–1754.
Yu, H., Ai, N., Peng, P., Ke, Y., Chen, X., Li, Y., Zhao, T., Jiang, S., Liu, J., and Jiang, L. (2021a). Comprehensive identification of fetal *cis*-regulatory elements in the human genome by single-cell multi-omics analysis. bioRxiv, 2021.2011.2002.466852.
Yu, L., Wang, X., Mu, Q., Tam, S.S.T., Loi, D.S.C., Chan, A.K.Y., Poon, W.S., Ng, H.K., Chan, D.T.M., Wang, J., et al. (2023a). scONE-seq: a single-cell multi-omics method enables simultaneous dissection of phenotype and genotype heterogeneity from frozen tumors. Sci Adv 9, eabp8901.
Yu, M., Abnousi, A., Zhang, Y., Li, G., Lee, L., Chen, Z., Fang, R., Lagler, T.M., Yang, Y., Wen, J., et al. (2021b). SnapHiC: a computational pipeline to identify chromatin loops from single-cell Hi-C data. Nat Methods 18, 1056–1059.
Yu, X., Liu, Z., and Sun, X. (2023b). Single-cell and spatial multi-omics in the plant sciences: technical advances, applications, and perspectives. Plant Commun 4, 100508.
Yu, Z., and Du, F. (2022). AMC: accurate mutation clustering from single-cell DNA sequencing data. Bioinformatics 38, 1732–1734.
Yu, Z., Du, F., and Song, L. (2022). SCClone: accurate clustering of tumor single-cell DNA sequencing data. Front Genet 13, 823941.
Yu, Z., Liu, F., Shi, F., and Du, F. (2023c). rcCAE: a convolutional autoencoder method for detecting intra-tumor heterogeneity and single-cell copy number alterations. Brief Bioinf 24, bbad108.
Yuan, H., and Kelley, D.R. (2022). scBasset: Sequence-based modeling of single-cell ATAC-seq using convolutional neural networks. Nat Methods 19, 1088–1096.
Yuan, Y., and Bar-Joseph, Z. (2020). GCNG: graph convolutional networks for inferring gene interaction from spatial transcriptomics data. Genome Biol 21, 300.
Yuan, Z., Zhou, Q., Cai, L., Pan, L., Sun, W., Qumu, S., Yu, S., Feng, J., Zhao, H., Zheng, Y., et al. (2021). SEAM is a spatial single nuclear metabolomics method for dissecting tissue microenvironment. Nat Methods 18, 1223–1232.
Zachariadis, V., Cheng, H., Andrews, N., and Enge, M. (2020). A highly scalable method for joint whole-genome sequencing and gene-expression profiling of single cells. Mol Cell 80, 541–553.e5.
Zeisel, A., Hochgerner, H., Lönnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Häring, M., Braun, E., Borm, L.E., La Manno, G., et al. (2018). Molecular architecture of the mouse nervous system. Cell 174, 999–1014.e22.
Zeng, H. (2022). What is a cell type and how to define it? Cell 185, 2739–2755.
Zeng, H., Huang, J., Ren, J., Wang, C.K., Tang, Z., Zhou, H., Zhou, Y., Shi, H., Aditham, A., Sui, X., et al. (2023a). Spatially resolved single-cell translatomics at molecular resolution. Science 380, eadd3067.
Zeng, H., Huang, J., Zhou, H., Meilandt, W.J., Dejanovic, B., Zhou, Y., Bohlen, C.J., Lee, S.H., Ren, J., Liu, A., et al. (2023b). Integrative *in situ* mapping of single-cell transcriptional states and tissue histopathology in a mouse model of Alzheimer's disease. Nat Neurosci 26, 430–446.
Zeng, X., Jiang, S., Zhong, Z., Yang, X., Chen, Q., Li, J., Zhu, Z., Song, J., and Yang, C. (2024). DIRECT: digital microfluidics for isolation-free shared library construction of single-cell DNA methylome and transcriptome. Small Methods 8, e2301075.
Zeng, Z., Li, Y., Li, Y., and Luo, Y. (2022). Statistical and machine learning methods for spatially resolved transcriptomics data analysis. Genome Biol 23, 83.
Zhai, J., Guo, J., Wan, H., Qi, L., Liu, L., Xiao, Z., Yan, L., Schmitz, D.A., Xu, Y., Yu, D., et al. (2022). Primate gastrulation and early organogenesis at single-cell resolution. Nature 612, 732–738.
Zhang, A.W., O'Flanagan, C., Chavez, E.A., Lim, J.L.P., Ceglia, N., McPherson, A., Wiens, M., Walters, P., Chan, T., Hewitson, B., et al. (2019a). Probabilistic cell-type assignment of single-cell RNA-seq for tumor microenvironment profiling. Nat Methods 16, 1007–1015.
Zhang, B., He, P., Lawrence, J.E., Wang, S., Tuck, E., Williams, B.A., Roberts, K., Kleshchevnikov, V., Mamanova, L., Bolt, L., et al. (2023a). A human embryonic limb cell atlas resolved in space and time. bioRxiv, 2022.2004.2027.489800.
Zhang, B., Srivastava, A., Mimitou, E., Stuart, T., Raimondi, I., Hao, Y., Smibert, P., and Satija, R. (2022a). Characterizing cellular heterogeneity in chromatin state with scCUT&Tag-pro. Nat Biotechnol 40, 1220–1230.
Zhang, F., Wei, K., Slowikowski, K., Fonseka, C.Y., Rao, D.A., Kelly, S., Goodman, S. M., Tabechian, D., Hughes, L.B., Salomon-Escoto, K., et al. (2019b). Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. Nat Immunol 20, 928–942.
Zhang, H., Li, J., Ren, J., Sun, S., Ma, S., Zhang, W., Yu, Y., Cai, Y., Yan, K., Li, W., et al. (2021a). Single-nucleus transcriptomic landscape of primate hippocampal aging. Protein Cell 12, 695–716.
Zhang, K., Hocker, J.D., Miller, M., Hou, X., Chiou, J., Poirion, O.B., Qiu, Y., Li, Y.E., Gaulton, K.J., Wang, A., et al. (2021b). A single-cell atlas of chromatin accessibility in the human genome. Cell 184, 5985–6001.e19.
Zhang, L., Li, Z., Skrzypczynska, K.M., Fang, Q., Zhang, W., O'Brien, S.A., He, Y., Wang, L., Zhang, Q., Kim, A., et al. (2020a). Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. Cell 181, 442–459.e29.
Zhang, L., Yu, X., Zheng, L., Zhang, Y., Li, Y., Fang, Q., Gao, R., Kang, B., Zhang, Q., Huang, J.Y., et al. (2018). Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. Nature 564, 268–272.
Zhang, L., and Zhang, S. (2018). PBLR: an accurate single cell RNA-seq data imputation tool considering cell heterogeneity and prior expression level of dropouts. bioRxiv, 379883.
Zhang, M., Eichhorn, S.W., Zingg, B., Yao, Z., Cotter, K., Zeng, H., Dong, H., and Zhuang, X. (2021c). Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. Nature 598, 137–143.
Zhang, M., Yang, H., Wan, L., Wang, Z., Wang, H., Ge, C., Liu, Y., Hao, Y., Zhang, D., Shi, G., et al. (2020b). Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma. J Hepatol 73, 1118–1130.
Zhang, Q., He, Y., Luo, N., Patel, S.J., Han, Y., Gao, R., Modak, M., Carotta, S., Haslinger, C., Kind, D., et al. (2019c). Landscape and dynamics of single immune cells in hepatocellular carcinoma. Cell 179, 829–845.e20.
Zhang, S., Fang, W., Zhou, S., Zhu, D., Chen, R., Gao, X., Li, Z., Fu, Y., Zhang, Y., Yang, F., et al. (2023b). Single cell transcriptomic analyses implicate an immunosuppressive tumor microenvironment in pancreatic cancer liver metastasis. Nat Commun 14, 5123.
Zhang, T., Zhang, Z., Li, L., Dong, B., Wang, G., and Zhang, D. (2023c). GTAD: A graph-based approach for cell spatial composition inference from integrated scRNA-seq and ST-seq data. Brief Bioinf 25, bbad469.
Zhang, W., Zhang, S., Yan, P., Ren, J., Song, M., Li, J., Lei, J., Pan, H., Wang, S., Ma, X., et al. (2020c). A single-cell transcriptomic landscape of primate arterial aging. Nat Commun 11, 2202.
Zhang, W., Zhang, X., Liu, Q., Wei, L., Qiao, X., Gao, R., Liu, Z., and Wang, X.

(2023d). Deconer: a comprehensive and systematic evaluation toolkit for reference-based cell type deconvolution algorithms using gene expression data. bioRxiv, 2023.2012.2024.573278.
Zhang, X., Hu, C., Huang, C., Wei, Y., Li, X., Hu, M., Li, H., Wu, J., Czajkowsky, D.M., Guo, Y., et al. (2022b). Robust acquisition of spatial transcriptional programs in tissues with immunofluorescence-guided laser capture microdissection. Front Cell Dev Biol 10, 853188.
Zhang, X., Li, T., Liu, F., Chen, Y., Yao, J., Li, Z., Huang, Y., and Wang, J. (2019d). Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-Seq systems. Mol Cell 73, 130–142.e5.
Zhang, Y., Chen, H., Mo, H., Hu, X., Gao, R., Zhao, Y., Liu, B., Niu, L., Sun, X., Yu, X., et al. (2021d). Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. Cancer Cell 39, 1578–1593.e8.
Zhao, S., Zhang, J., and Nie, Z. (2023). Large-scale cell representation learning via divide-and-conquer contrastive learning. arXiv, 2306.04371.
Zhao, T., Chiang, Z.D., Morriss, J.W., LaFave, L.M., Murray, E.M., Del Priore, I., Meli, K., Lareau, C.A., Nadaf, N.M., Li, J., et al. (2022). Spatial genomics enables multi-modal study of clonal heterogeneity in tissues. Nature 601, 85–91.
Zheng, C., Fass, J.N., Shih, Y.P., Gunderson, A.J., Sanjuan Silva, N., Huang, H., Bernard, B.M., Rajamanickam, V., Slagel, J., Bifulco, C.B., et al. (2022). Transcriptomic profiles of neoantigen-reactive T cells in human gastrointestinal cancers. Cancer Cell 40, 410–423.e7.
Zheng, C., Zheng, L., Yoo, J.K., Guo, H., Zhang, Y., Guo, X., Kang, B., Hu, R., Huang, J. Y., Zhang, Q., et al. (2017a). Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. Cell 169, 1342–1356.e16.
Zheng, G.X.Y., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S. B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017b). Massively parallel digital transcriptional profiling of single cells. Nat Commun 8, 14049.
Zheng, L., Qin, S., Si, W., Wang, A., Xing, B., Gao, R., Ren, X., Wang, L., Wu, X., Zhang, J., et al. (2021). Pan-cancer single-cell landscape of tumor-infiltrating T cells. Science 374, abe6474.
Zheng, W., Min, W., Wang, S., and Boeva, V. (2023). TsImpute: An accurate two-step imputation method for single-cell RNA-seq data. Bioinformatics 39, btad731.
Zhong, Y., Wan, Y.W., Pang, K., Chow, L.M., and Liu, Z. (2013). Digital sorting of complex tissues for cell type-specific gene expression profiles. BMC Bioinf 14, 89.
Zhou, F., Wang, R., Yuan, P., Ren, Y., Mao, Y., Li, R., Lian, Y., Li, J., Wen, L., Yan, L., et al. (2019). Reconstituting the transcriptome and DNA methylome landscapes of human implantation. Nature 572, 660–664.
Zhou, J., Theesfeld, C.L., Yao, K., Chen, K.M., Wong, A.K., and Troyanskaya, O.G. (2018). Deep learning sequence-based ab initio prediction of variant effects on expression and disease risk. Nat Genet 50, 1171–1179.
Zhou, T., Zhang, R., Jia, D., Doty, R.T., Munday, A.D., Gao, D., Xin, L., Abkowitz, J.L., Duan, Z., and Ma, J. (2023). Concurrent profiling of multiscale 3D genome organization and gene expression in single mammalian cells. bioRxiv, 2023.07.20.549578.
Zhou, W., Yan, Y., Guo, Q., Ji, H., Wang, H., Xu, T., Makabel, B., Pilarsky, C., He, G., Yu, X., et al. (2021). Microfluidics applications for high-throughput single cell sequencing. J Nanobiotechnol 19, 312.
Zhou, Y., Bian, S., Zhou, X., Cui, Y., Wang, W., Wen, L., Guo, L., Fu, W., and Tang, F. (2020). Single-cell multiomics sequencing reveals prevalent genomic alterations in tumor stromal cells of human colorectal cancer. Cancer Cell 38, 818–828.e5.
Zhu, B., Chen, S., Bai, Y., Chen, H., Liao, G., Mukherjee, N., Vazquez, G., McIlwain, D. R., Tzankov, A., Lee, I.T., et al. (2023). Robust single-cell matching and multimodal analysis using shared and distinct features. Nat Methods 20, 304–315.
Zhu, C., Yu, M., Huang, H., Juric, I., Abnousi, A., Hu, R., Lucero, J., Behrens, M.M., Hu, M., and Ren, B. (2019). An ultra high-throughput method for single-cell joint analysis of open chromatin and transcriptome. Nat Struct Mol Biol 26, 1063–1070.
Zhu, C., Zhang, Y., Li, Y.E., Lucero, J., Behrens, M.M., and Ren, B. (2021). Joint profiling of histone modifications and transcriptome in single cells from mouse brain. Nat Methods 18, 283–292.
Zhu, H., Galdos, F.X., Lee, D., Waliany, S., Huang, Y.V., Ryan, J., Dang, K., Neal, J.W., Wakelee, H.A., Reddy, S.A., et al. (2022). Identification of pathogenic immune cell subsets associated with checkpoint inhibitor-induced myocarditis. Circulation 146, 316–335.
Zhu, J., and Sabatti, C. (2020). Integrative spatial single-cell analysis with graph-based feature learning. bioRxiv, 2020.2008.2012.248971.
Zhu, Y.Y., Machleder, E.M., Chenchik, A., Li, R., and Siebert, P.D. (2001). Reverse transcriptase template switching: a SMART™ approach for full-length cDNA library construction. BioTechniques 30, 892–897.
Zong, C., Lu, S., Chapman, A.R., and Xie, X.S. (2012). Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. Science 338, 1622–1626.
Zou, B., Zhang, T., Zhou, R., Jiang, X., Yang, H., Jin, X., and Bai, Y. (2021). deepMNN: deep learning-based single-cell RNA sequencing data batch correction using mutual nearest neighbors. Front Genet 12, 708981.
Zou, Z., Zhang, C., Wang, Q., Hou, Z., Xiong, Z., Kong, F., Wang, Q., Song, J., Liu, B., Liu, B., et al. (2022). Translatome and transcriptome co-profiling reveals a role of TPRXs in human zygotic genome activation. Science 378, abo7923.