

Life Sciences and Medicine

Single-cell protein-DNA interactomics and multiomics tools for deciphering genome regulation

Haiqing Xiong^{1,#,*}, Runyu Wang^{2,#} & Aibin He^{3,*}

¹State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin 300020, China;

²Peking University-Tsinghua University-National Institute of Biological Science Joint Graduate Program, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China;

³Institute of Molecular Medicine, College of Future Technology, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China

#Contributed equally to this work.

*Corresponding authors (emails: xionghaiqing@ihcams.ac.cn (Haiqing Xiong); ahe@pku.edu.cn (Aibin He))

Received 13 October 2022; Revised 22 December 2022; Accepted 9 January 2023; Published online 30 March 2023

Abstract: The emergence of single-cell genomic and transcriptomic sequencing accelerates the development of single-cell epigenomic technologies, providing an unprecedented opportunity for decoding cell fate decisions largely encoded in the epigenome. Recent advances in single-cell multimodality epigenomic technologies facilitate directly interrogating the regulatory relationship between multi-layer molecular information in the same cell. In this review, we discuss recent progress in development of single-cell multimodality epigenomic technologies and applications in elucidating cellular diversifications in development and diseases, with a focus on protein-DNA interactomics and regulatory links between epigenome and transcriptome. Further, we provide perspective on the future direction of single-cell multiomics tool development as well as challenges facing ahead.

Keywords: single-cell protein-DNA interaction, single-cell multiomics, single-cell epigenomics, epigenome, genome regulation

Introduction

Deciphering the epigenetic regulation of dynamic gene expression underlying cell-fate decisions is a long-standing theme in biomedical studies. Insights into the kinship of neighboring cells in space and time have been gained through imaging-based tools such as time-lapse microscopy and genetic fate mapping [1–5]. Emerging single-cell sequencing tools hold great promises to boost our understanding of the cellular heterogeneity and molecular regulation of dynamic cell states during development and diseases. Single-cell mRNA sequencing (scRNA-seq) has revolutionized broad biomedical researches in interpreting cellular heterogeneities and inferring cell differentiation trajectories in pseudotime [6–12]. However, single-cell transcriptomic sequencing datasets yield limited insights into the regulatory principles in cell fate specifications, i.e., how the epigenome is dynamically sculpted in cells diversifying from one another upon sti-

mulation while they share common ancestors. Excitedly, efforts have been made to move beyond the transcriptome to epigenome in deciphering *cis* and *trans* gene regulators. Pioneering single-cell profiling technologies have been realized, ranging from genomics, epigenomics, transcriptomics, to proteomics (Figure 1). Importantly, experimental and computational combination of multimodality molecular in-

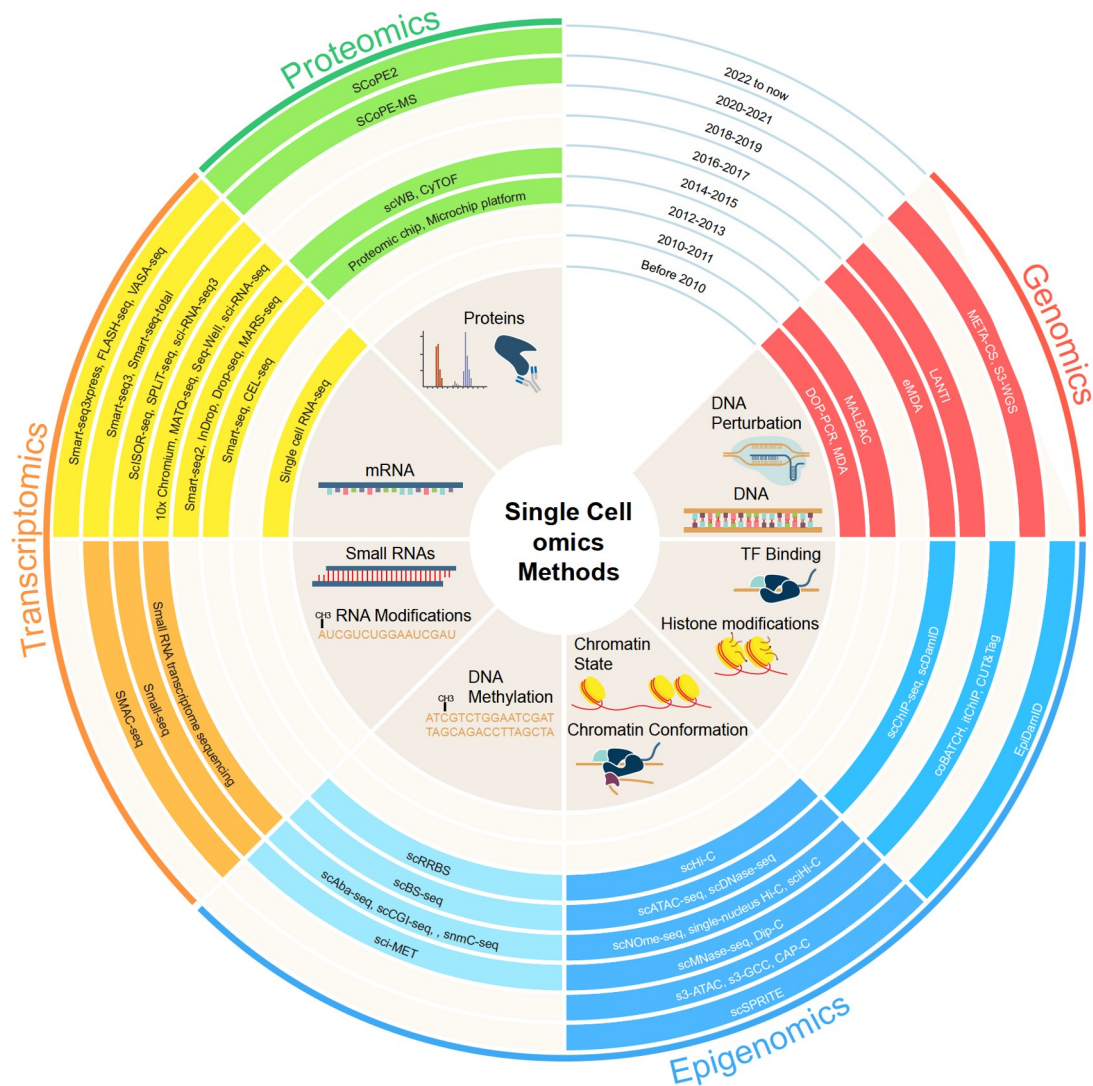


Figure 1 Overview of single-cell Omics methods. Numerous single-cell technologies have been developed to profile genome (red), epigenome (blue), transcriptome (orange), and proteome (green), respectively. The distance from the center of a circle to the edge indicates the year from 2010 to now. The references (PMIDs) for corresponding methods are listed below. Smart-seq3xpress, 35637418. FLASH-seq, 35637419. VASA-seq, 35760914. Smart-seq3, 32518404. Smart-seq-total, 34911763. SciSOR-seq, 30320766. SPLiT-seq, 29545511. Sci-RNA-seq3, 30787437. 10x Chromium, 28091601. MATQ-seq, 31028632. Seq-Well, 28192419. Sci-RNA-seq, 28818938. Smart-seq2, 24056875. InDrop, 26000487. Drop-seq, 26000488. MARS-seq, 24531970. Smart-seq, 22820318. CEL-seq, 22939981. Single cell RNA-seq, 19349980. SMAC-seq, 32042188. Small-seq, 30250291. Small RNA transcriptome sequencing, 27798564. SCoPE2, 33504367. SCoPE-MS, 30343672. SeWB, 24880876. CyTOF, 24584119. Proteomic chip, 22203961. Microchip platform, 23998271. META-CS, 33593904. S3-WGS, 34226710. LIANTI, 28408603. eMDA, 26340991. MALBAC, 23258894. DOP-PCR, 1639399. MDA, 11381035. EpiDamID, 35366395. CoBATCH, 31471188. itChIP, 31481796. CUT&Tag, 31036827. scChIP-seq, 26458175. scDamID, 26257285. scHi-C, 24067610. scATAC-seq, 26083756. scDNase-seq, 26605532. scNome-seq, 28653622. Single-nucleus Hi-C, 28355183. sciHi-C, 28135255. scMNase-seq, 30258225. Dip-C, 30166492. s3-ATAC & s3-GCC, 34226710. CAP-C, 32839564. scSPRITE, 34426703. scRRBS, 25837417. scBS-seq, 25042786. scAba-seq, 27347753. scCGI-seq, 28126923. snmC-seq, 28798132. sci-MET, 29644997.

formation has been forged into single-cell multi-omics measurements. We describe state-of-the-art progress and present future perspective in this flourishing field bridging fundamental biology and biomedicine.

Challenges and solutions in interrogating the epigenome in single cells

Gene activity and cell identity are largely encoded in the multi-layer epigenomic landscape, including transcription factor binding, chromatin accessibility, histone modifications, DNA modifications, and 3-dimensional chromosome conformation in diverse biological processes [13,14]. Genome-wide mapping of epigenetic features in bulk samples provided valuable information in identifying *cis*-regulatory elements, such as promoters and enhancers, and their roles in gene regulation. Nevertheless, averaged signals in bulk assays provide no cell-type specific features in heterogenous tissues.

Since there are only two copies of genomic DNA in a diploid cell, single-cell epigenomic sequencing technologies suffer from more challenges than scRNA-seq dealing with multiple copies of transcripts. To improve the detection efficiency of epigenome in single cells, it is necessary to optimize the experimental protocols, including designing new biochemistry. As an example, Buenrostro *et al.* [15] successfully developed single-cell assay for transposase-accessible chromatin using sequencing (scATAC-seq) based on Tn5 transposase-mediated tagmentation and simultaneous addition of sequencing adaptors for DNA fragments in the library preparation. Single cell reduced representation bisulfite sequencing (scRRBS-seq) enables the detection of approximately 40% of the CpG sites in bulk RRBS by limiting the experimental steps in a single tube before PCR amplification to avoid sample loss [16–18].

The low throughput strategy for achieving single-cell resolution relies on manual pipetting or fluorescence-activated cell sorting (FACS) to sort individual cells or nuclei into multi-well plates with well-specific barcodes [6]. Furthermore, droplet-based microfluidic platforms improve the throughput to thousands of cells in one assay [9,10]. Recently, combinatorial indexing has emerged as a promising strategy for high-throughput single-cell sequencing technologies [19–21]. In addition, droplet-based microfluidic platforms in combination with combinatorial indexing not only greatly facilitate automation but also further increase the throughput and easy adoption in aim for from hundreds of thousands to millions of cells per experiment [22].

Single-cell protein-DNA interactomics

The gene activity is largely gated through chromatin states, including chromatin accessibility, histone modifications, and chromatin protein bindings. Genome-wide protein-DNA interaction measurement—interactomics concerns a root layer of epigenetic regulation. Chromatin immunoprecipitation followed by throughput sequencing (ChIP-seq) is a widely adopted tool for examining genome-wide protein-DNA interactions as well as histone modifications [23,24]. Traditional ChIP-seq bears major limitations such as requirement for a large amount of starting materials, careful validation of ChIP grade antibodies, low signal-to-noise, and low resolution [25]. Overcoming these challenges in the standard ChIP-seq has promoted the rapid development of the field, aiming at improving the signal-to-noise ratio, increasing the resolution, and reducing the input cell number. In the few years, considerable progress has been made to detect protein-DNA

interaction with markedly improved sensitivity using as few as 100 cells including STAR ChIP [26], CUT&RUN [27], nano-ChIP [28], UNINChIP [29], MOWChIP [30], and ChIL-seq [31]. Furthermore, the Bernstein laboratory first reported a prototype of single-cell ChIP-seq procedure using a droplet-based microfluidic platform (scDrop-ChIP) in 2015, and three subpopulations in mouse embryonic stem cells were identified based on single-cell H3K4me2 ChIP-seq data. However, since the experimental protocol is largely built on the traditional ChIP-seq, scDrop-ChIP only generates ~800 unique reads per cell, hampering its potential applications in complex biological samples [32]. Recently, we and Henikoff laboratory have at the same time reported a widely applicable, low-cost and robust strategy, dubbed CoBATCH [33] and CUT&Tag [34], respectively, to profile histone modifications and transcription factor bindings in both low-input cells and single cells based on a fusion protein Protein A-Tn5 (PAT). CoBATCH has reported the unique reads as high as 12,000 per cell. Owing to the targeted tagmentation and enrichment of DNA fragments and obviation of immunoprecipitations, CoBATCH and CUT&Tag profoundly improve the sensitivity and precision of detection with high signal-to-noise. Similarly, by fusing Protein A with micrococcal nuclease (MNase), scChIC-seq [35], uliCUT&RUN [36] and iscChIC-seq [37] have been reported to exhibit high signal-to-noise profiles. iscChIC-seq adopts TdT-tailing-based combinatorial indexing (~11,000 reads per cell for 7800 individual cells). Moreover, improvement for genome-wide measurement of DNA-protein interaction in single cells with higher throughput may be obtained through droplet-based and combinatorial indexing strategies (Figure 2). Two studies couple the scCUT&Tag protocol with the widely available 10x Genomics Chromium platform to explore the epigenetic heterogeneity of mouse brain and dynamic changes in chromatin silencing in patients with brain tumors [38,39]. Although droplet-based scCUT&Tag may be used to identify major subpopulations in heterogenous tissues, it suffers from obnoxious data sparsity (i.e., low coverage) with unique reads per cell ranging from 98 to 453 in various histone modification profiling in mouse brain [38]. Therefore, researchers are confronted by a trade-off between cell throughput and genomic coverage when performing single-cell protein-DNA interactomics profiling. Of note, the split-and-pool

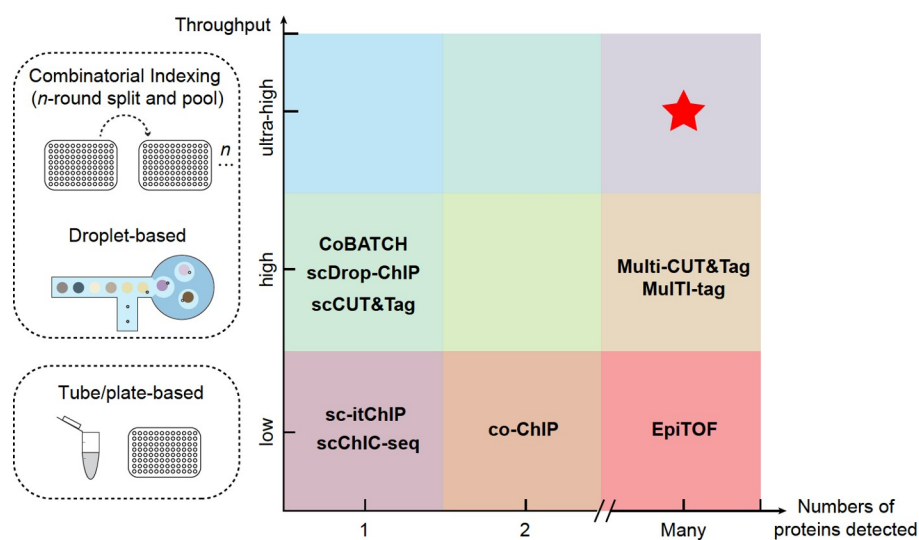


Figure 2 Measurement of histone modifications and transcription factor bindings at single-cell level. The red asterisk highlights that there currently exists a gap for single-cell Omics methods that map many chromatin proteins with ultra-high throughput and precision. CoBATCH, 31471188. sc-itChIP, 31481796. scCUT&Tag, 31036827. scDrop-ChIP, 26458175. scChIC-seq, 30923384. co-ChIP, 27454738. EpiTOF, 29706550. Multi-CUT&Tag, 35098158. MulTI-Tag, 36316484.

based single-cell methods are not compatible with the context with limited cell number, such as early embryo development. As a complementary way, the He laboratory has recently reported a versatile technique, dubbed single-cell itChIP-seq [40]. After cell fixation with formaldehyde, itChIP-seq applied the detergent SDS treatment for relaxing/opening chromatin prior to Tn5 transposition genome wide, enabling the profiling of histone modifications or transcription factor binding with starting materials as few as 50 cells and throughput up to 2000 cells.

With a completely different design, DNA adenine methyltransferase identification (DamID) is also able to capture contacts between proteins of interest (POIs) and chromatin on a genome-wide scale at single-cell resolution [41]. Recently, EpiDamID has been reported to simultaneously measure histone modification and transcriptome in single cells by fusing Dam to POIs, engineered chromatin reader domains, or single-chain variable fragments of antibodies [42]. However, DamID relies on transgenic expression of POIs, largely precluding its potential applications in histone modifications and clinical samples.

It is exciting to witness the emergence of new techniques such as scMulti-CUT&Tag [43], MulTI-Tag [44], nano-CT [45] and NTT-seq [46] for profiling multiple histone modifications at one time in single cells. In scMulti-CUT&Tag, different epitopes on chromatin were recognized by PAT-antibody conjugates with antibody-specific barcodes. They suggested that purification of PAT-barcode-antibody complex may be the key step to avoid swapping of antibody-barcode during targeted tagmentation. While in MulTI-Tag, bar-coded adaptors are firstly covalently conjugated to antibody before preloaded with PAT, and antibody-tethered tagmentation is carried out sequentially to avoid cross contamination between different histone modifications. However, these methods suffer from poor data quality (for example, ~200 unique reads per cell for scMulti-CUT&Tag and ~1000 unique reads per cell for MulTI-Tag) and a limited number of histone modifications profiled at a time. Thus, development of methods for co-occupancy of multiple chromatin proteins with high-sensitivity and high-efficiency is highly demanded.

Together, these pioneer single-cell epigenomic technologies, while still in infancy, offer us the capacity to unbiasedly unlock the regulatory relationship between a myriad of dynamic epigenetic proteins reading out regulatory elements and transcriptional outcome. Our intention is to provide a quick overview for newcomers to the field of single-cell protein-DNA interactomics as well as deeper description with details of the leading technologies to date in light of different biological contexts.

Single-cell multimodality epigenomic technologies

Although single-cell unimodal measures enable identification of heterogeneities in cell types and states, these approaches provide relatively limited values in connecting epigenomic regulation with phenotypic outcome when one layer of molecular information is assayed in single cells at a time. Emerging single-cell multi-Omics present an unprecedented opportunity to explore the cell type-specific gene regulation and intricate relationship between different modalities in tissues and various conditions from a more comprehensive view.

Distinguishing different molecular modalities in single cells in resulting sequencing data is key to realize single-cell multiomics profiling. Hence, a highly effective procedure, compatible to simultaneous measurements of different molecular modalities, is required to capture and discern each of them. In general, varying strategies include: (1) physical separation of molecules. For example, scMT-seq [45] and scTrio-seq

[46] adopted physical separation by micropipette manipulation to separate cytosolic mRNA and nuclear DNA methylome. scTrio-seq2 [47] further improved this approach by using magnetic beads to separate nucleus (DNA) and cytoplasm (mRNA). These methods were applied to unravel the connection between the epigenome and transcriptome in neural lineages and colorectal cancer. (2) Biochemical separation. Namely, although different layers of molecules are derived from the same cell lysates in the multi-omics protocol, they could be distinguished and independently analyzed by virtue of unique biochemical modifications. G&T-seq [48] enabled parallel sequencing of single-cell genome and transcriptome. In brief, as with biotin-labeled oligo-dT primers for mRNA capture and reverse transcription, cDNA and genomic DNA (gDNA) were separated for the respective library preparation after incubation with streptavidin-coupled magnetic beads. (3) Information conversion. Two studies applied similar principles to profile DNA methylation and nucleosome positioning in the same cell. In these cases, through bisulfite conversion, DNA methylation and chromatin accessibility can be further analyzed after the treatment of CpG methylation and GpC methylation, respectively [49,50]. (4) Pre-library split. In the protocol of DR-seq [51], gDNA and cDNA were separated for the following library preparation after reverse transcription and quasilinear whole-genome amplification rather than in the initial step of component separation for individual cells. Similar strategy was also used in the droplet workflow of ISSAAC-seq [52], where pre-library was separated for ATAC and RNA library construction, respectively. Collectively, these strategies are widely applied in most current single-cell multi-omics approaches. However, problems such as the incompatibility of experimental conditions in different single-cell unimodal methods largely affect the data quality. More efforts are still warranted to refine the experimental protocols of multimodal single-cell technologies.

Despite these challenges, great strides have been made in development of multimodal single-cell omics technologies with high throughput (Figure 3, Table 1). For example, single-cell multiomics technologies allowing for joint detection of chromatin accessibility and gene expression in the same cell have been developed to elucidate epigenetic mechanisms of genome regulation. scCAT-seq [53] enables simultaneous profiling of chromatin accessibility and transcriptome through physical separation of nucleus DNA and cytosolic mRNA by mild lysis and centrifuge, notwithstanding in low cell throughputs. In comparison, SNARE-seq [54], sci-CAR [55], Paired-seq [56], and SHARE-seq [57] achieved high-throughput on different platforms. Moving forward, both the Ren and He laboratories have recently described high-throughput platforms for measuring chromatin occupancy along with transcriptome in thousands to hundreds of thousands of single cells in one experiment. CoTECH [58] adopted both modified smart-seq2 and CoBATCH to capture the transcriptome and chromatin protein bindings. Tagmented DNA or mRNA was indexed by different combinations of T5-T7 barcodes or well-specific oligo dT primers. After pre-amplification, amplicons were separated into DNA and RNA partitions. For DNA partition, well-indexed primers were used to introduce the second-round barcodes. For RNA partition, full-length cDNA was tagmented by Tn5 followed by the second-round indexing. The DNA and RNA profiles were linked by corresponding pairs of indexing primers. Paired-Tag [59] adopted four rounds of combinatorial indexing to dramatically scale up the throughput. Similar to histone modifications or transcription factors whose genomic distribution could be captured by Tn5 tagmentation, the abundance of cell surface and intracellular proteins could also be translated into digital sequencing counts. CITE-seq [60] labeled the antibodies with oligonucleotides to translate the information of the surface protein abundance into sequencing-based readout. The antibody oligos and mRNA were captured by oligo-dT primers on the beads in one droplet. The amplified antibody-

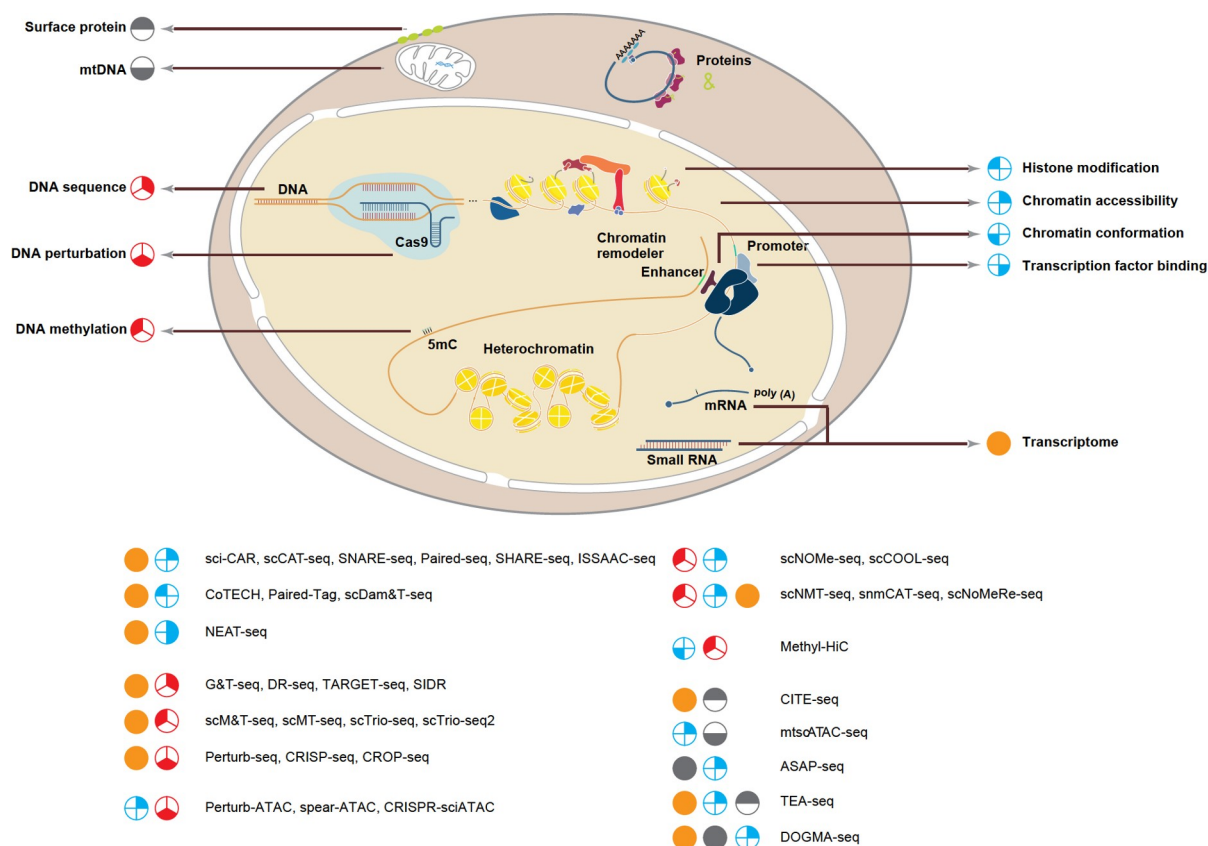


Figure 3 Methods for single-cell multimodal Omics analysis. Top: schematics of modalities in the cell. Different modalities are highlighted by different colored circles. Bottom: different combinations of colored circles stand for the modalities profiled in the corresponding technologies. References for these methods are as listed in Table 1.

derived tags and cDNA were separated by size and sequenced. Built on CITE-seq and mtsaATAC-seq [61], ASAP-seq [62] was developed to profile chromatin accessibility, mtDNA and cell surface/intracellular proteins simultaneously on 10x Genomics Chromium single-cell ATAC Kit platform. Bridge oligos with unique bridge identifiers were used during GEM incubation to ensure that each antibody-oligo complex could be bridged only once. Overall, advances in single-cell multi-Omics, aiming at simultaneously detecting the different modalities including gene expression, chromatin states, chromosome conformation, copy number variation, lineage, cell surface protein, and spatial information in single cells, represent a new frontier for characterization of cell states and cellular dynamics in human health and diseases.

Single-cell multi-omics technologies unleash the power for unbiased characterization of cell types and states

Cellular heterogeneities have been extensively addressed by scRNA-seq analysis. Though highly informative, single-cell transcriptome itself in some cases is insufficient to phenotypically reflect cell types/states. For example, the abundance of cell surface proteins rather than RNA expression is most commonly

Table 1 Methods for single-cell multimodal Omics^{a)}

Methods	Modalities assayed	Strategies for identifying modalities	Throughput	Engineering characteristics	PMID
scG&T-seq	DNA + RNA	I + II	Benchmarking: 130 cells Case study: 7 embryos, 41 iPSCs, HCC38-BL cells (<i>n</i> = 36) and single HCC38 cells (<i>n</i> = 32)	Manually selected or flow cytometry	25915121
DR-seq	DNA + RNA	III + V	Benchmarking: 13 E14 cells, Case study: 21 SK-BR-3 cells Benchmarking: 180 single cells (JURKAT, <i>n</i> = 56; SET2, <i>n</i> = 86; and HSPC, <i>n</i> = 38)	Mouth pipetting	25599178
TARGET-seq	DNA + RNA	III	Case study: 458 HSPCs from 5 MPN patients; 2798 cells from a cohort of 8 patients with MF	Fluorescent activated cell sorting	30765193
SIDR	DNA + RNA	II + III	Benchmarking: 43 cells	Sorting 1 µL of cell suspension (1 cell / µL) into each well of 48-well microplate	29208629
scM&T-seq	DNA methylation + RNA	I + II	Benchmarking & Case study: 76 individual serum ESCs and 16 ESCs grown in "2i" media	FACS after staining	27150361
scMT-seq	DNA methylation + RNA	I	Benchmarking & Case study: 15 sensory neurons from DRG	Micro-capillary pipetting	30498128
scTrio-seq	DNA methylation + RNA	II	Benchmarking: 6 HepG2 + 6 mESCs Case study: 26 hepatocellular carcinoma cells	Mouth pipetting	26902283
scTrio-seq2	DNA methylation + RNA	II	Case study: ~1900 cells from 12 CRC patients	Mouth pipetting	33958790
CoTECH	Protein-DNA interactions + RNA	III	Benchmarking: 1:1 mixed HEK293T and NIH 3T3 cells Case study: 6993 mESCs, 512 AGM cells + 736 YS cells	Two rounds of combinatorial indexing	33958790
Paired-Tag	Protein-DNA interactions + RNA	II + III	Benchmarking: tens of thousands of HeLa cells Case study: ~70,000 nuclei from brain cells	Four rounds of combinatorial indexing	33589836
scDam&T-seq	Protein-DNA interactions + RNA	II	Case study: ~70,000 nuclei from brain cells	FACS	32350457
sciCAR	Chromatin accessibility + RNA	III	Benchmarking: 4825 A549 cells after DEX treatment Case study: 11,296 cells from whole mouse kidneys	Two rounds of combinatorial indexing	30166440
scCAT-seq	Chromatin accessibility + RNA	I	Benchmarking: 176 K562 cells, Hela-S3 and HCT116 Case study: 110 clinically discarded human embryos	Two rounds of combinatorial indexing	30692544
SNARE-seq	Chromatin accessibility + RNA	IV	Benchmarking: 1047 cells from mixtures of BJ, HI, K562 and GM12878	FACS	31611697
SHARE-seq	Chromatin accessibility + RNA	II	Case study: 5081 cells from mouse neonatal cerebral cortex, 10,309 cells from adult mouse cerebral cortex Benchmarking: ~900 GM12878 + ~1300 NIH 3T3 Case study: 34,774 mouse skin cells, 23,278 GM12878 cells	Three rounds of combinatorial indexing	33098772
Paired-seq	Chromatin accessibility + RNA	II + III	Benchmarking: ~2053 nuclei from mixture of NIH3T3, HepG2, HEK293T	Five rounds of combinatorial indexing	31695190
ISSAAC-seq	Chromatin accessibility + RNA	III	Case study: 15,191 nuclei from adult mouse cerebral cortex Benchmarking: mixture of 293T and 3T3 (plate-based), PBMCs and K562 (droplet-based)	FACS or 10x Genomics scATAC platform	36109677
scNOMe-seq	DNA methylation + Nucleosome occupancy	IV	Case study: 10,378 adult mouse cerebral cortex Benchmarking & Case study: 19 GM12878 cells and 11 K562 cells	FACS	28653622

(To be continued on the next page)

(Continued)

Methods	Modalities assayed	Strategies for identifying modalities	Throughput	Engineering characteristics	PMID
scCOOL-seq	DNA methylation + Nucleosome occupancy	IV	Benchmarking: 24 mESCs Case study: 233 cells from mouse preimplantation embryos	Mouth pipetting	28621329
scNMT-seq	DNA methylation + Nucleosome occupancy	II + IV	Benchmarking: 70 mESCs Case study: 43 ES cells that were removed from LIF for 3 days	FACS	29472610
snmCAT-seq	DNA methylation + Nucleosome occupancy	IV + V	Benchmarking: mixture of HI and HEK293T Case study: 4358 nuclei from postmortem human frontal cortex tissue, Benchmarking & Case study: 233 mouse preimplantation embryos from different stages	FACS	35419551
scNoMeRe-seq	DNA methylation + Nucleosome occupancy	II + IV	Benchmarking & Case study: 233 mouse preimplantation embryos from different stages	FACS	33623021
CITE-seq	Cell surface protein + RNA	IV	Benchmarking: mixture of HeLa and 4T1 Case study: 8005 CBMCs	10x Single Cell 3' v1 or v2	28759029
TEA-seq	RNA + Cell surface protein + Chromatin accessibility	IV	Benchmarking & Case study: thousands of PBMCs	10x Genomics scATAC platform	33835024
mtscATAC-seq	Chromatin accessibility + mtDNA	-	Benchmarking: mixture of GM11906 and TF1 Case study: 818 GM11906 cells from MERRF patient	10x Genomics scATAC platform	34083792
ASAP-seq	Chromatin accessibility + Cell surface protein + mtDNA	IV	Benchmarking: mixture of 293T and 3T3, PBMCs Case study: 10,928 cells from bone marrow, 8535 PBMCs after T cell stimulation	10x Genomics Multiome kit	34083792
DOGMA-seq	Chromatin accessibility + RNA + Cell surface protein + mtDNA	IV	Benchmarking: 6139 PBMCs after T cell stimulation	10x Genomics Multiome kit	34083792
NEAT-seq	Chromatin accessibility + RNA + Nuclear protein	IV	Benchmarking: mixture of K562 and mESCs Case study: human CD4 memory T cells	10x Genomics Multiome kit	35501385
Perturb-seq	DNA perturbation + RNA	-	Benchmarking & Case study: 6 experiments analyzing 200,000 cells (BMDCs and K562)	10x Chromium System	27984732
CROP-seq	DNA perturbation + RNA	-	Benchmarking : mixture of 3 KO cell lines of 293T Case study: 5798 Jurkat cells (screen of T-cell receptor activation)	FACS	28099430
CRISP-seq	DNA perturbation + RNA	-	Benchmarking: bone marrow cells from mice Case study: 731 + 1186 + 6749 mice myeloid cells	FACS	27984734
Methy-HiC	DNA methylation + Chromatin conformation	-	Case study: 103 serum mESCs cultured in LIF and 47 2i mESCs cultured in LIF	FACS	31384045
Perturb-ATAC	DNA perturbation + Chromatin accessibility	-	Benchmarking & Case study: 63 genotypes using ~4300 single cells	FACS	30580963
Spear-ATAC	DNA perturbation + Chromatin accessibility	-	Benchmarking: 3045 K562 cells Case study: 32,832 K562 cells representing 128 sgRNA genotypes	10x Genomics scATAC platform	34016988
CRISPR-sciATAC	DNA perturbation + Chromatin accessibility	-	Benchmarking: 1986 cells from mixture of NIH3T3 and HEK293 Case study: 11,104 + 16,676 K562 cells (screening of chromatin modifier and chromatin remodeling complexes)	Two rounds of combinatorial indexing	33927415

a) The summary for single-cell experimental strategies and throughput (represented by cells profiled in the indicated studies) of single-cell technologies as in Figure 3. MPN, myeloproliferative neoplasms. MF, myelofibrosis. DRG, dorsal root ganglion. CRC, colorectal cancer. DEX, dexamethasone. MERRF, myoclonic epilepsy with red ragged fibers. KO, knock out. LIF, leukemia inhibitory factor. I, physical separation. II, biochemical separation. III, split. IV, Conversion. V, computational deconvolution.

used in the identification of CD8⁺ T cell in the immune system [63]. It is necessary to unbiasedly define cell types/states to gain comprehensive understanding of cell fate regulation. The Satija laboratory has reported the weighted nearest neighbor (WNN) algorithm to integrate information of multiple modalities in single cells based on relative importance of each modality [63]. The scNMT-seq is able to jointly detect transcriptome, DNA methylation, and chromatin accessibility in the same cell, and multi-omics factor analysis (MOFA) can be applied to integrate a multi-modal mouse gastrulation atlas [64]. Clearly, the mounting single-cell multi-omics data and the regulatory relation to cell function or phenotype have greatly facilitated the precise annotation of cell types and states.

Single-cell multi-omics for decoding cell lineage potentials

Exploring gene regulatory mechanisms in different biological processes remains challenging. The emergence of genome-scale assays such as single-cell multi-omics are now available for probing gene transcription and its regulation of cell differentiation. For example, SHARE-seq reveals that chromatin accessibility fore-shadows gene expression during lineage differentiation in adult mouse skin. The He laboratory developed CoTECH to interrogate the relationship between the bivalent histone modifications H3K4me3/H3K27me3 and transcriptome in mouse embryonic stem cells and uncovered a context-specific regulatory interplay from naive to primed mESCs. We expect that similar to SHARE-seq, the future study harnessing the knowledge in the kinetics in varying histone modifications during cell differentiation could also open a new avenue to predict prospective cell potential in a given cell populations in development and diseases.

Single-cell epigenomic and multi-omic data analysis

The development and maturation of single-cell epigenomic and multimodality Omic technologies necessitates the emergence of new computational methods to analyze single-cell sequencing data [65]. Single-cell sequencing data analysis is faced with main challenges such as dropout events, high dimensionality, batch effect and detection noise. While single-cell RNA-seq data analytic tools are relatively mature with standardized bioinformatic pipelines, single-cell epigenomic analysis is rather challenging due to the data sparsity thus far. Therefore, it is important to accurately distinguish the real biological signal from noise. To circumvent these limitations, computational approaches have been developed to handle single-cell epigenetic data in the form of sparse matrices. Scasat [66] and SnapATAC [67] binarize raw count matrix and perform cell clustering, which ignores the co-activity and variable peaks among different regulatory elements. In the scRNA-seq analysis, highly variable genes that might be biologically important are identified before dimensional reduction. However, this is not the case for single-cell epigenetic analyses since variable peaks cannot be obtained due to the binary features. Cusanovich *et al.* [21] first took the most accessible features as input to latent semantic indexing (LSI) algorithm, a common workflow in scATAC-seq data processing borrowed from topic modeling. LSI results in high degrees of noise and low reproducibility when running multiple samples. Another topic-based method cisTopic [68] enables the classification of genomic regions into regulatory modules to handle the intrinsic sparsity. Pliner *et al.* [69] introduced Cicero to identify co-

accessible elements from chromatin hubs, linking dynamic *cis*-regulatory elements to their target genes. More recently, ArchR provides an intuitive and user-focused interface for complex downstream analyses [70]. In addition, chromVAR globally predicts enrichment of the TF activity based on scATAC-seq data, and is thus instructive for lineage specification during cell fate decision [71]. Other analytical tools including SnapATAC [67], Signac [72], scATAC-pro [73], and SCALE [74] were also commonly used in scATAC-seq analyses with comparable set of features including doublet removal, cell clustering, identification of differential peaks, and visualization of marker genes. Of note, a sparsity similar to scATAC-seq data is also seen in single-cell profiling of histone modifications and TF binding. Hence, several analytic ideas for scATAC-seq such as dimension reduction and cell clustering may be also applied to the analysis pipeline of single-cell histone modifications and TF bindings profiling. However, single-cell protein-DNA interactomics and scATAC-seq data analyses differ in many aspects (e.g., active and repressive histone modifications exhibit unique features of genomic distribution), implying that existing consensus scATAC-seq analysis pipelines require further optimization when applied to single-cell protein-DNA interactomics dataset. More specialized algorithms including cell and feature selection, dimension reduction, and cell clustering for single-cell histone modification and TF binding analysis are still lacking.

Integrative analysis of single-cell multi-omics data

To make full use of the wealthy information of single-cell multiomics data, Haghverdi *et al.* [75] and Butler *et al.* [76] described two computational approaches, mutual nearest neighbors (MNN) and canonical correlation analysis (CCA), respectively. In brief, these methods integrate different single-cell modalities by identifying a shared structure across datasets, which are aligned for correction of batch effect and further single-cell downstream analyses such as unbiased cell clustering and cell type annotation. Nevertheless, different types of single-cell modality data do not share common features in some cases, leading to an analytic problem when integrating multiple single-cell datasets. For example, the correspondence features between chromatin accessibility and DNA methylome data is unclear [77]. To this end, Welch *et al.* [78] developed an algorithm, named MATCHER, to infer single cell multi-omic profiles from transcriptomic and epigenetic measurements without priori knowledge of cell correspondences through projection of cells from different modalities onto a common pseudotime space. Using this method, the authors analyzed the correlation and dynamics between gene expression and DNA methylation during human iPSC reprogramming.

While the single-cell research field enters the era of single-cell multiomics, integrative computational algorithms are still in their infancy [77]. The following integrative analyses we discussed mainly focus on paired single-cell multiomics datasets, i.e., different modalities are profiled from the same cells.

(1) Reference-mapping. Most of single-cell multi-omics approaches include transcriptomic profiling. Accordingly, utilizing transcriptome as an anchor to link different modality data across different experiments is a choice. Xiong *et al.* [58] established an analytical framework using CoTECH data to investigate the relationship of multiple histone modifications and explore the bivalency dynamics from naive to primed mESCs. A similar work took advantage of the transcriptomic information to reconstruct the developmental trajectory of OGC maturation and analyzed epigenetic regulatory programs of H3K4me1, H3K4me3, and H3K27ac. Instead of applying transcriptome as a reference, scCUT&Tag-pro enabled integrative analyses

ranging from measurements of chromatin accessibility, protein-DNA interactions (six targets), gene expression and surface proteins by harmonizing into a common space based on the protein abundances [79].

(2) Combined effects. Given the contribution of each layer to cellular heterogeneity, single-cell data from all modalities were combined and represented by one graph. For example, Argelaguet *et al.* [80] developed a computational method MOFA to identify a set of factors and discover the source of variation in single-cell multi-Omics data. Using this strategy, the authors performed co-assays of the single-cell transcriptome, chromatin accessibility and DNA methylation during gastrulation in mouse embryos, and found that pluripotent epiblast cells are epigenetically primed for an ectoderm fate, which are earlier than the emergence of cell fate bias of endoderm and mesoderm [81]. Similarly, the computational integration of single-cell multi-modal profiling data such as MNN, scAI, and totalVI may also lead to more accurate characterization of cell states [63,81–83].

Additionally, computational methods are required to impute missing values in some cases [84]. A bunch of bioinformatic tools have been developed for accurate prediction of single-cell DNA methylation [85,86], and epigenomic cell-state dynamics [87]. However, computational approaches may inevitably lead to mis-assignments or inaccuracy. Stuart *et al.* [88] performed CCA using the ATAC and RNA profiles from the same cells separately and compared its inferred pairing with the correct (measured) coupling, and found an accuracy ranging from 36.7%–74.9%. It should be worth noting that such computational errors may also result from the asynchrony between different molecular layers.

Future outlook

In summary, the larger international single-cell Omics community is working together to push the boundary of the field by inventing novel single-cell epigenomics and multi-Omics tools, aiming at unbiasedly defining cell identities and states and exploring the regulatory mechanisms. Despite much progress has been made to reconstruct a comprehensive single-cell atlas to link spatiotemporal dynamics in individual cells in a multicellular organism (Figure 4A and 4B), current single-cell epigenomic and multi-Omics approaches are far from mature due to the data sparsity, limited recovering rate from input cells and relatively low detection efficiency of molecules, not to mention full coverage of single-molecule levels. Therefore, it is critical to distinguish true biological signals in individual cells from potential technical noise. Optimization in molecular biology is required to improve the recovery of transcriptomic and genomic features. For example, the Adey laboratory developed a symmetrical strand sci (“s3”) strategy to overcome the 50% yield limitation of standard Tn5 tagmentation by using a uracil-based adapter switching approach, achieving significant improvements in usable reads per cell [89]. The data quality for other transposase-based single-cell techniques may be also improved via this strategy. Experimentally, single-cell sequencing technologies are moving forward in the following directions: (1) measuring multiple layers instead of single layer in one cell; (2) measuring multiple factors at a time; (3) achieving ultra-high throughput at low cost; (4) compatible with commercial automation platform.

Spatial localization manifests at scales ranging from local (intercellular crosstalk) to global (tissue and organ patterning), strongly impacting on cell function and cell fate decisions [90]. However, most single-cell omics protocols require single-cell dissociation and suspension, resulting in elimination of spatial in-

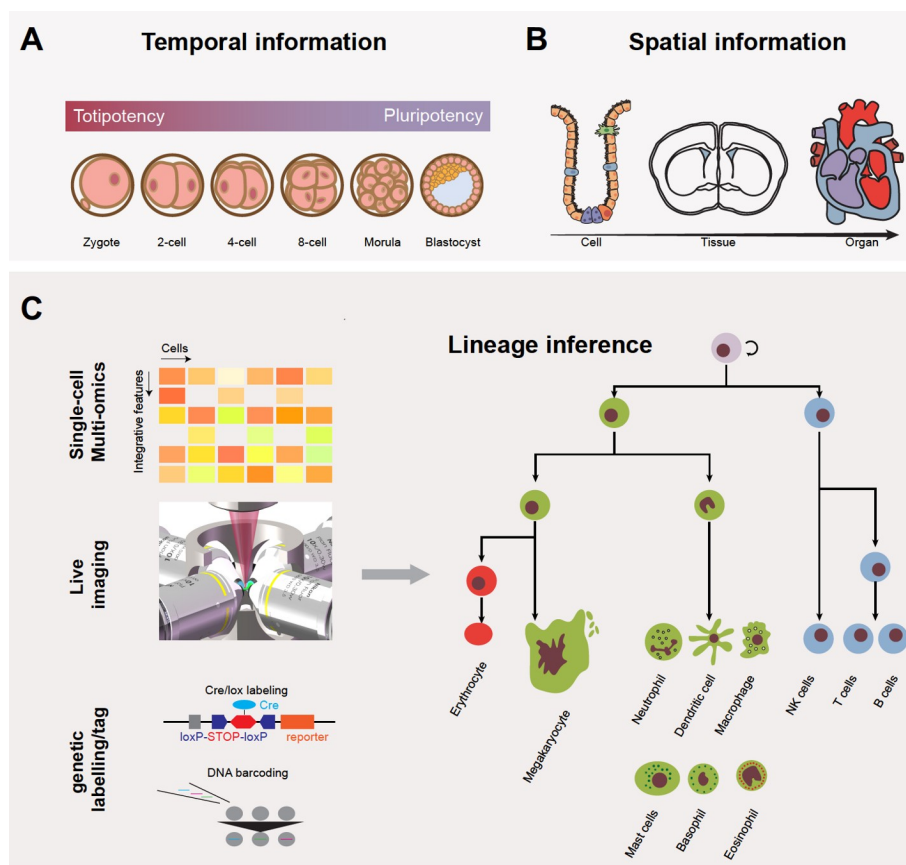


Figure 4 Applications and perspectives for single cell omics technologies. (A), (B) Scalable, robust single-cell multi-omics methods provide insights into mammalian early development and organogenesis from a temporal (A) and spatial (B) view. (C) Combined analysis of cutting-edge methods including single-cell multi-omics, live imaging, and genetic labelling for comprehensively and systematically reconstructing cell lineages in development and diseases.

formation. Spatially-resolved transcriptomic methods have been developed and improved rapidly in recent years based on *in situ* sequencing and single-molecule FISH strategy. For example, technologies including Stereo-seq [91], smFISH [92], MERFISH [93], Slide-seq [94], seqFISH [95], and Tomo-Seq [96] have begun to achieve measures of spatial gene expression at high resolution. Moving beyond spatially resolved transcriptome, the Fan laboratory successfully developed Spatial-CUT&Tag [97] and spatial-ATAC-seq [98] for genome-wide profiling of histone modifications and chromatin accessibility, deepening our understanding of epigenetic regulation in spatial chromatin states. In addition, sciMAP-ATAC successfully characterized the spatially progressive nature in a cerebral ischemia model system [99]. These elegant tools help identify key regulatory elements with specific spatial information. Further efforts would be focused on the extended applications such as tumorigenesis and cancer microenvironment to realize the full potential of spatial epigenome in biomedical researches.

Computational methods have emerged to reconstruct the lineage relationship and infer the underlying dynamics of differentiation in cell fate decisions [100]. Most single-cell pseudotime algorithms are designed for single-cell transcriptomic data analysis, but less is reported on depicting the differentiation potential and developmental trajectory from the perspective in epigenome. Inspired by the ability of chromatin potential

and chromatin velocity to recapitulates differentiation processes [101], we believe that chromatin states captured by single-cell epigenomic and multi-omics technologies are equally important and have the potential to infer cell lineages and predict future cell states. Importantly, several elegant technologies such as live imaging have surfaced and greatly expanded our toolkits to study cell lineage during mammalian organogenesis [102–104]. From our point of view, experimental and computational combination of single-cell multi-omics, live imaging, and genetic lineage tracing will provide an unprecedented opportunity to reveal precise and complete lineage definition and cell identity (Figure 4C).

The field of single-cell sequencing methods is undergoing transformative development from measurements of cellular molecules to large-scale genetic perturbations using the CRISPR-Cas9 system. Early investigations on large-scale cellular function screens were limited to scRNA-seq readout. For example, Perturb-seq, CROP-seq, and CRISP-seq directly linked pooled CRISPR screens with transcriptomic output in thousands to hundreds of thousands of individual cells [105–108]. To facilitate high-throughput functional exploration of epigenetic mechanisms, Perturb-ATAC-seq and droplet-based Spear-ATAC-seq have been developed to characterize the effect of CRISPR/Cas9-based perturbations on chromatin accessibility and examine how epigenomic heterogeneity contributes to transcriptional regulation [109,110]. The integration of single-cell CRISPR-Cas9 functional characterization and epigenomic profiling allows for identification of cell type-specific *cis*-regulatory elements. So far, functional readout through single-cell epigenomic profiling is particularly challenging, and few tools are available for extrapolating cellular functions from large-scale single-cell omics data. In the future, more efforts would be expected to focus on dissecting gene regulatory network by combining modified genome editing toolkits and single-cell multiomics technologies to demystify the roles of *cis*-regulatory elements and *trans* factors in the genome regulation. Breakthrough in technologies would beget the opportunity to associate epigenetic configurations with cell potential and realize precise manipulation of cell fate, with great implications in regenerative medicine, diagnosis and treatment in human diseases [111].

Funding

This work was supported by the National Key Research and Development Program of China (2021YFA1100101 and 2017YFA0103402), the National Natural Science Foundation of China (32025015 and 31771607), and the Peking-Tsinghua Center for Life Sciences.

Author contributions

A.H. conceived the structure of the manuscript; A.H., H.X., and R.W. wrote the paper; H.X. and R.W. prepared the figures and tables.

Conflict of interest

The authors declare no conflict of interest.

References

- 1 Lemaire P. Unfolding a chordate developmental program, one cell at a time: invariant cell lineages, short-range inductions and evolutionary plasticity in ascidians. *Dev Biol* 2009; **332**: 48–60.
- 2 Sulston JE, Schierenberg E, White JG, *et al.* The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev*

- Biol* 1983; **100**: 64–119.
- 3 Kretzschmar K, Watt FM. Lineage tracing. *Cell* 2012; **148**: 33–45.
 - 4 Gómez-Villafuertes R, Paniagua-Herranz L, Gascon S, *et al.* Live imaging followed by single cell tracking to monitor cell biology and the lineage progression of multiple neural populations. *J Vis Exp* 2017; <https://doi.org/10.3791/56291>.
 - 5 Zovein AC, Hofmann JJ, Lynch M, *et al.* Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell* 2008; **3**: 625–636.
 - 6 Tang F, Barbacioru C, Wang Y, *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 2009; **6**: 377–382.
 - 7 Picelli S, Faridani OR, Björklund ÅK, *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* 2014; **9**: 171–181.
 - 8 Hashimshony T, Wagner F, Sher N, *et al.* CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep* 2012; **2**: 666–673.
 - 9 Macosko EZ, Basu A, Satija R, *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 2015; **161**: 1202–1214.
 - 10 Klein AM, Mazutis L, Akartuna I, *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* 2015; **161**: 1187–1201.
 - 11 Jaitin DA, Kenigsberg E, Keren-Shaul H, *et al.* Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* 2014; **343**: 776–779.
 - 12 Sheng K, Cao W, Niu Y, *et al.* Effective detection of variation in single-cell transcriptomes using MATQ-seq. *Nat Methods* 2017; **14**: 267–270.
 - 13 Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 2003; **33**: 245–254.
 - 14 Handy DE, Castro R, Loscalzo J. Epigenetic modifications. *Circulation* 2011; **123**: 2145–2156.
 - 15 Buenrostro JD, Wu B, Litzenburger UM, *et al.* Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 2015; **523**: 486–490.
 - 16 Meissner A. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res* 2005; **33**: 5868–5877.
 - 17 Guo H, Zhu P, Wu X, *et al.* Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res* 2013; **23**: 2126–2135.
 - 18 Guo H, Zhu P, Guo F, *et al.* Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nat Protoc* 2015; **10**: 645–659.
 - 19 Cao J, Packer JS, Ramani V, *et al.* Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* 2017; **357**: 661–667.
 - 20 Rosenberg AB, Roco CM, Muscat RA, *et al.* Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* 2018; **360**: 176–182.
 - 21 Cusanovich DA, Hill AJ, Aghamirzaie D, *et al.* A single-cell atlas of *in vivo* mammalian chromatin accessibility. *Cell* 2018; **174**: 1309–1324.
 - 22 Lareau CA, Duarte FM, Chew JG, *et al.* Droplet-based combinatorial indexing for massive-scale single-cell chromatin accessibility. *Nat Biotechnol* 2019; **37**: 916–924.
 - 23 Johnson DS, Mortazavi A, Myers RM, *et al.* Genome-wide mapping of *in vivo* protein-DNA interactions. *Science* 2007; **316**: 1497–1502.
 - 24 Barski A, Cuddapah S, Cui K, *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* 2007; **129**: 823–837.
 - 25 Park PJ. ChIP-seq: Advantages and challenges of a maturing technology. *Nat Rev Genet* 2009; **10**: 669–680.
 - 26 Zheng H, Huang B, Zhang B, *et al.* Resetting epigenetic memory by reprogramming of histone modifications in mammals. *Mol Cell* 2016; **63**: 1066–1079.

- 27 Skene PJ, Henikoff JG, Henikoff S. Targeted *in situ* genome-wide profiling with high efficiency for low cell numbers. *Nat Protoc* 2018; **13**: 1006–1019.
- 28 Adli M, Bernstein BE. Whole-genome chromatin profiling from limited numbers of cells using nano-ChIP-seq. *Nat Protoc* 2011; **6**: 1656–1668.
- 29 Brind'Amour J, Liu S, Hudson M, *et al.* An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat Commun* 2015; **6**: 6033.
- 30 Cao Z, Chen C, He B, *et al.* A microfluidic device for epigenomic profiling using 100 cells. *Nat Methods* 2015; **12**: 959–962.
- 31 Harada A, Maehara K, Handa T, *et al.* A chromatin integration labelling method enables epigenomic profiling with lower input. *Nat Cell Biol* 2019; **21**: 287–296.
- 32 Rotem A, Ram O, Shores N, *et al.* Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol* 2015; **33**: 1165–1172.
- 33 Wang Q, Xiong H, Ai S, *et al.* CoBATCH for high-throughput single-cell epigenomic profiling. *Mol Cell* 2019; **76**: 206–216.
- 34 Kaya-Okur HS, Wu SJ, Codomo CA, *et al.* CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* 2019; **10**: 1930.
- 35 Ku WL, Nakamura K, Gao W, *et al.* Single-cell chromatin immunocleavage sequencing (scChIC-seq) to profile histone modification. *Nat Methods* 2019; **16**: 323–325.
- 36 Hainer SJ, Bošković A, McCannell KN, *et al.* Profiling of pluripotency factors in single cells and early embryos. *Cell* 2019; **177**: 1319–1329.
- 37 Ku WL, Pan L, Cao Y, *et al.* Profiling single-cell histone modifications using indexing chromatin immunocleavage sequencing. *Genome Res* 2021; **31**: 1831–1842.
- 38 Bartosovic M, Kabbe M, Castelo-Branco G. Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. *Nat Biotechnol* 2021; **39**: 825–835.
- 39 Wu SJ, Furlan SN, Mihalas AB, *et al.* Single-cell CUT&Tag analysis of chromatin modifications in differentiation and tumor progression. *Nat Biotechnol* 2021; **39**: 819–824.
- 40 Ai S, Xiong H, Li CC, *et al.* Profiling chromatin states using single-cell itChIP-seq. *Nat Cell Biol* 2019; **21**: 1164–1172.
- 41 Kind J, Pagie L, de Vries SS, *et al.* Genome-wide maps of nuclear lamina interactions in single human cells. *Cell* 2015; **163**: 134–147.
- 42 Rang FJ, de Luca KL, de Vries SS, *et al.* Single-cell profiling of transcriptome and histone modifications with EpiDamID. *Mol Cell* 2022; **82**: 1956–1970.
- 43 Gopalan S, Wang Y, Harper NW, *et al.* Simultaneous profiling of multiple chromatin proteins in the same cells. *Mol Cell* 2021; **81**: 4736–4746.
- 44 Meers MP, Llagas G, Janssens DH, *et al.* Multifactorial profiling of epigenetic landscapes at single-cell resolution using Multi-Tag. *Nat Biotechnol* 2022; <https://doi.org/10.1038/s41587-022-01522-9>.
- 45 Hu Y, Huang K, An Q, *et al.* Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol* 2016; **17**: 88.
- 46 Hou Y, Guo H, Cao C, *et al.* Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res* 2016; **26**: 304–319.
- 47 Bian S, Hou Y, Zhou X, *et al.* Single-cell multiomics sequencing and analyses of human colorectal cancer. *Science* 2018; **362**: 1060–1063.
- 48 Macaulay IC, Haerty W, Kumar P, *et al.* G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods* 2015; **12**: 519–522.
- 49 Pott S. Simultaneous measurement of chromatin accessibility, DNA methylation, and nucleosome phasing in single cells. *eLife* 2017; **6**: e23203.

- 50 Li L, Guo F, Gao Y, *et al.* Single-cell multi-omics sequencing of human early embryos. *Nat Cell Biol* 2018; **20**: 847–858.
- 51 Dey SS, Kester L, Spanjaard B, *et al.* Integrated genome and transcriptome sequencing of the same cell. *Nat Biotechnol* 2015; **33**: 285–289.
- 52 Xu W, Yang W, Zhang Y, *et al.* ISSAAC-seq enables sensitive and flexible multimodal profiling of chromatin accessibility and gene expression in single cells. *Nat Methods* 2022; **19**: 1243–1249.
- 53 Liu L, Liu C, Quintero A, *et al.* Deconvolution of single-cell multi-omics layers reveals regulatory heterogeneity. *Nat Commun* 2019; **10**: 470.
- 54 Chen S, Lake BB, Zhang K. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nat Biotechnol* 2019; **37**: 1452–1457.
- 55 Cao J, Cusanovich DA, Ramani V, *et al.* Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science* 2018; **361**: 1380–1385.
- 56 Zhu C, Yu M, Huang H, *et al.* An ultra high-throughput method for single-cell joint analysis of open chromatin and transcriptome. *Nat Struct Mol Biol* 2019; **26**: 1063–1070.
- 57 Ma S, Zhang B, LaFave LM, *et al.* Chromatin potential identified by shared single-cell profiling of RNA and chromatin. *Cell* 2020; **183**: 1103–1116.
- 58 Xiong H, Luo Y, Wang Q, *et al.* Single-cell joint detection of chromatin occupancy and transcriptome enables higher-dimensional epigenomic reconstructions. *Nat Methods* 2021; **18**: 652–660.
- 59 Zhu C, Zhang Y, Li YE, *et al.* Joint profiling of histone modifications and transcriptome in single cells from mouse brain. *Nat Methods* 2021; **18**: 283–292.
- 60 Stoeckius M, Hafemeister C, Stephenson W, *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* 2017; **14**: 865–868.
- 61 Lareau CA, Ludwig LS, Muus C, *et al.* Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. *Nat Biotechnol* 2021; **39**: 451–461.
- 62 Mimitou EP, Lareau CA, Chen KY, *et al.* Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells. *Nat Biotechnol* 2021; **39**: 1246–1258.
- 63 Hao Y, Hao S, Andersen-Nissen E, *et al.* Integrated analysis of multimodal single-cell data. *Cell* 2021; **184**: 3573–3587.
- 64 Clark SJ, Argelaguet R, Kapourani CA, *et al.* scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat Commun* 2018; **9**: 781.
- 65 Efremova M, Teichmann SA. Computational methods for single-cell omics across modalities. *Nat Methods* 2020; **17**: 14–17.
- 66 Baker SM, Rogerson C, Hayes A, *et al.* Classifying cells with Scasat, a single-cell ATAC-seq analysis tool. *Nucleic Acids Res* 2019; **47**: e10.
- 67 Fang R, Preissl S, Li Y, *et al.* Comprehensive analysis of single cell ATAC-seq data with SnapATAC. *Nat Commun* 2021; **12**: 1337.
- 68 Bravo González-Blas C, Minnoye L, Papanokrati D, *et al.* cisTopic: cis-regulatory topic modeling on single-cell ATAC-seq data. *Nat Methods* 2019; **16**: 397–400.
- 69 Pliner HA, Packer JS, McFaline-Figueroa JL, *et al.* Cicero predicts cis-regulatory DNA interactions from single-cell chromatin accessibility data. *Mol Cell* 2018; **71**: 858–871.
- 70 Granja JM, Corces MR, Pierce SE, *et al.* ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet* 2021; **53**: 403–411.
- 71 Schep AN, Wu B, Buenrostro JD, *et al.* chromVAR: Inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods* 2017; **14**: 975–978.
- 72 Stuart T, Srivastava A, Madad S, *et al.* Single-cell chromatin state analysis with Signac. *Nat Methods* 2021; **18**: 1333–1341.

- 73 Yu W, Uzun Y, Zhu Q, *et al.* scATAC-pro: A comprehensive workbench for single-cell chromatin accessibility sequencing data. *Genome Biol* 2020; **21**: 94.
- 74 Xiong L, Xu K, Tian K, *et al.* SCALE method for single-cell ATAC-seq analysis via latent feature extraction. *Nat Commun* 2019; **10**: 4576.
- 75 Haghverdi L, Lun ATL, Morgan MD, *et al.* Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol* 2018; **36**: 421–427.
- 76 Butler A, Hoffman P, Smibert P, *et al.* Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018; **36**: 411–420.
- 77 Stuart T, Satija R. Integrative single-cell analysis. *Nat Rev Genet* 2019; **20**: 257–272.
- 78 Welch JD, Hartemink AJ, Prins JF. MATCHER: Manifold alignment reveals correspondence between single cell transcriptome and epigenome dynamics. *Genome Biol* 2017; **18**: 138.
- 79 Zhang B, Srivastava A, Mimitou E, *et al.* Characterizing cellular heterogeneity in chromatin state with scCUT&Tag-pro. *Nat Biotechnol* 2022; **40**: 1220–1230.
- 80 Argelaguet R, Velten B, Arnol D, *et al.* Multi-Omics factor analysis—A framework for unsupervised integration of multi-omics data sets. *Mol Syst Biol* 2018; **14**: e8124.
- 81 Argelaguet R, Clark SJ, Mohammed H, *et al.* Multi-Omics profiling of mouse gastrulation at single-cell resolution. *Nature* 2019; **576**: 487–491.
- 82 Jin S, Zhang L, Nie Q. scAI: An unsupervised approach for the integrative analysis of parallel single-cell transcriptomic and epigenomic profiles. *Genome Biol* 2020; **21**: 25.
- 83 Gayoso A, Steier Z, Lopez R, *et al.* Joint probabilistic modeling of single-cell multi-Omic data with totalVI. *Nat Methods* 2021; **18**: 272–282.
- 84 Ernst J, Kellis M. Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. *Nat Biotechnol* 2015; **33**: 364–376.
- 85 Angermueller C, Lee HJ, Reik W, *et al.* DeepCpG: Accurate prediction of single-cell DNA methylation states using deep learning. *Genome Biol* 2017; **18**: 67.
- 86 Kapourani CA, Sanguinetti G. Melissa: Bayesian clustering and imputation of single-cell methylomes. *Genome Biol* 2019; **20**: 61.
- 87 Farlik M, Sheffield NC, Nuzzo A, *et al.* Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. *Cell Rep* 2015; **10**: 1386–1397.
- 88 Stuart T, Butler A, Hoffman P, *et al.* Comprehensive integration of single-cell data. *Cell* 2019; **177**: 1888–1902.
- 89 Mulqueen RM, Pokholok D, O’Connell BL, *et al.* High-content single-cell combinatorial indexing. *Nat Biotechnol* 2021; **39**: 1574–1580.
- 90 Longo SK, Guo MG, Ji AL, *et al.* Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat Rev Genet* 2021; **22**: 627–644.
- 91 Wei X, Fu S, Li H, *et al.* Single-cell Stereo-seq reveals induced progenitor cells involved in axolotl brain regeneration. *Science* 2022; **377**: eabp9444.
- 92 Femino AM, Fay FS, Fogarty K, *et al.* Visualization of single RNA transcripts *in situ*. *Science* 1998; **280**: 585–590.
- 93 Chen KH, Boettiger AN, Moffitt JR, *et al.* Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* 2015; **348**: aaa6090.
- 94 Rodriques SG, Stickels RR, Goeva A, *et al.* Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* 2019; **363**: 1463–1467.
- 95 Shah S, Lubeck E, Zhou W, *et al.* *In situ* transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron* 2016; **92**: 342–357.
- 96 Kruse F, Junker JP, van Oudenaarden A, *et al.* Tomo-seq: A method to obtain genome-wide expression data with spatial resolution. *Methods Cell Biol* 2016; **135**: 299–307.
- 97 Deng Y, Bartosovic M, Kukanja P, *et al.* Spatial-CUT&Tag: Spatially resolved chromatin modification profiling at the

- cellular level. *Science* 2022; **375**: 681–686.
- 98 Deng Y, Bartosovic M, Ma S, *et al.* Spatial profiling of chromatin accessibility in mouse and human tissues. *Nature* 2022; **609**: 375–383.
- 99 Thornton CA, Mulqueen RM, Torkenczy KA, *et al.* Spatially mapped single-cell chromatin accessibility. *Nat Commun* 2021; **12**: 1274.
- 100 Sagar, Grün D. Deciphering cell fate decision by integrated single-cell sequencing analysis. *Annu Rev Biomed Data Sci* 2020; **3**: 1–22.
- 101 Tedesco M, Giannese F, Lazarević D, *et al.* Chromatin velocity reveals epigenetic dynamics by single-cell profiling of heterochromatin and euchromatin. *Nat Biotechnol* 2022; **40**: 235–244.
- 102 Yue Y, Zong W, Li X, *et al.* Long-term, in toto live imaging of cardiomyocyte behaviour during mouse ventricle chamber formation at single-cell resolution. *Nat Cell Biol* 2020; **22**: 332–340.
- 103 Keller PJ, Schmidt AD, Santella A, *et al.* Fast, high-contrast imaging of animal development with scanned light sheet-based structured-illumination microscopy. *Nat Methods* 2010; **7**: 637–642.
- 104 Keller PJ, Schmidt AD, Wittbrodt J, *et al.* Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science* 2008; **322**: 1065–1069.
- 105 Dixit A, Parnas O, Li B, *et al.* Perturb-Seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* 2016; **167**: 1853–1866.
- 106 Adamson B, Norman TM, Jost M, *et al.* A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell* 2016; **167**: 1867–1882.
- 107 Datlinger P, Rendeiro AF, Schmidl C, *et al.* Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods* 2017; **14**: 297–301.
- 108 Jaitin DA, Weiner A, Yofe I, *et al.* Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. *Cell* 2016; **167**: 1883–1896.
- 109 Rubin AJ, Parker KR, Satpathy AT, *et al.* Coupled single-cell CRISPR screening and epigenomic profiling reveals causal gene regulatory networks. *Cell* 2019; **176**: 361–376.
- 110 Pierce SE, Granja JM, Greenleaf WJ. High-throughput single-cell chromatin accessibility CRISPR screens enable unbiased identification of regulatory networks in cancer. *Nat Commun* 2021; **12**: 2969.
- 111 Shema E, Bernstein BE, Buenrostro JD. Single-cell and single-molecule epigenomics to uncover genome regulation at unprecedented resolution. *Nat Genet* 2019; **51**: 19–25.