COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

Exploring long non-coding RNA networks from single cell omics data

Xue Zhao [1], Yangming Lan [1], Dijun Chen *

State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, Nanjing 210023, China

ARTICLE INFO

ABSTRACT

Single-cell omics technologies provide an unprecedented opportunity to decipher molecular mechanisms underlying various biological processes in a cellular heterogeneity manner. The emergence of such techniques promotes the exploration of lncRNAs, which are known to be tissue- and cell-specific noncoding transcripts involving the regulation of multiple important cellular processes. In this review, we introduce the advancement of lncRNA studies which benefit from single-cell omics data analysis. We discuss the expression heterogeneity of lncRNAs, their cell-type specificity and associated gene regulatory networks (GRNs) from a single-cell perspective. We also summarized the state-of-the-art single-cell omics resources and tools for the construction of single-cell GRNs (scGRNs) that could be potentially used for lncRNA functional study. Finally, we highlight the challenges and prospective for scGRN exploration in lncRNA biology.

Contents

1. Introduction

Long noncoding RNAs (lncRNAs) are a group of noncoding RNAs (ncRNAs) with length more than 200 nucleotides, which generally located at intergenic regions or orientated in an antisense direction of protein-coding genes (Fig. 1). Although lncRNAs do not encode for proteins, they are proven to play a vital role in numerous important biological phenomena such as chromatin remodelling, genome architecture, RNA stabilization and transcription regulation [1,2]. They can function through multiple ways including but not limited to: **(1)** action in *cis* to recruit proteins or protein complexes to specific loci [3], **(2)** scaffolding of nuclear or cytoplasmic complexes [4], **(3)** work as mediators of RNA dependent DNA methylation [5], and **(4)** pairing with other RNAs to trigger post-transcriptional regulation [6]. The function and type of lncRNA have been extensively summarized by many reviews ([7–16]).

---

* Corresponding author.
  E-mail address: dijunchen@nju.edu.cn (D. Chen).
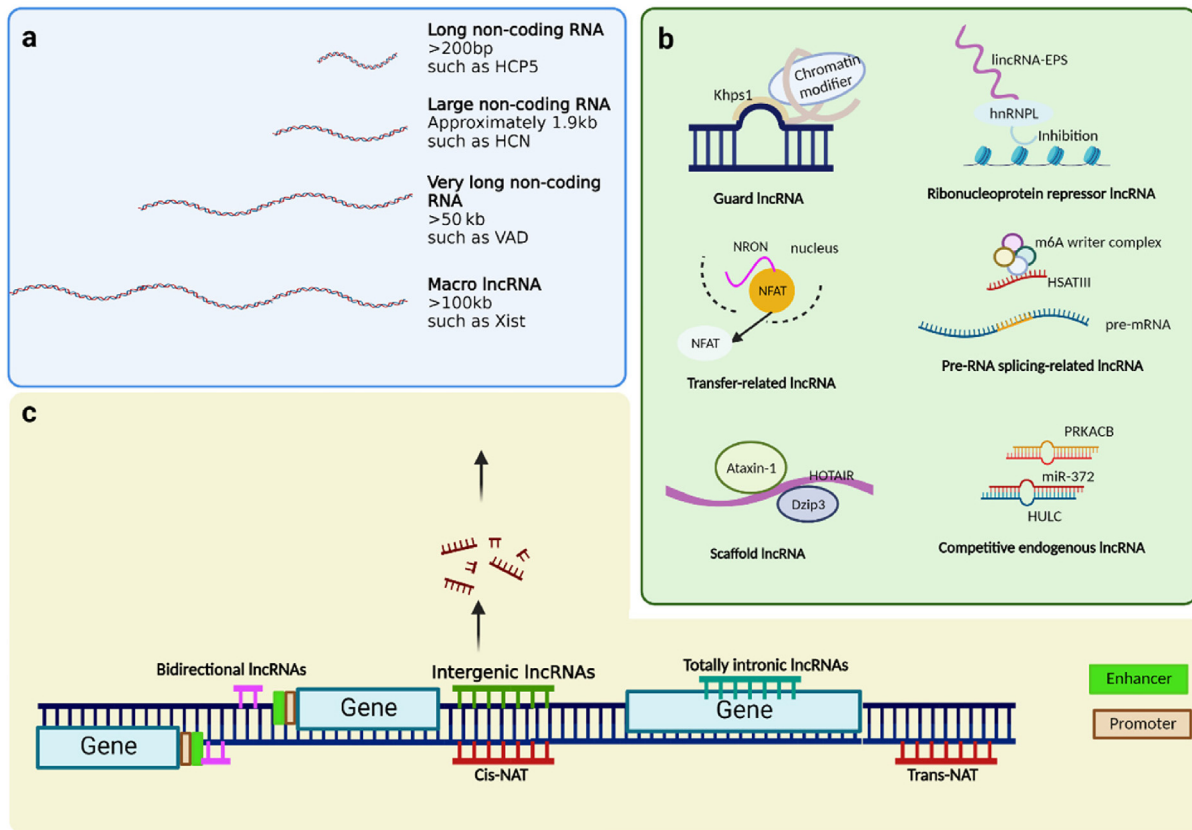 [1] These authors contributed equally to this work.

Fig. 1. The different taxonomies of lncRNAs. (a) LncRNAs are categorized by length. (b) LncRNAs are categorized by function. (c) LncRNAs are categorized by location.

Considering the large number of lncRNAs annotated in one organism and their hypothetically important regulatory role, the functional analysis of the majority of lncRNAs is still insufficient. One of the obstacles is that lncRNAs are generally expressed at relatively lower level than mRNAs and exhibit poor sequence conservation [17]. Various transcriptome-wide studies have shown that lncRNAs generally express in a tissue- or cell-specific manner or under a certain condition [18]. This specificity has been used as evidence since expression profiles of lncRNA may provide important information for diseases or developmental states [19]. Although expression specificity is a general distinct characteristic between lncRNAs and mRNAs, lncRNAs share common transcriptional and biochemical features with mRNAs [20]. For example, like mRNAs, lncRNAs are transcribed by RNA polymerase II (Pol II) [21] and often 5′-capped, spliced and polyadenylated. Moreover, lncRNAs often showed correlated expression patterns with mRNAs [20,21], suggesting that both lncRNAs and mRNAs may be co-regulated in same gene networks or biological pathways. Therefore, gene regulation network analysis that has been used to decipher the gene function in complex biological process can be also expanded to lncRNAs.

High-throughput RNA sequencing (RNA-seq) technology rapidly generates enormous amount of data in different tissues or organs. RNA-seq offers the opportunity to capture even low-copy transcripts in samples and it thus allow biologists to probe lncRNAs in a transcriptome-wide manner [18]. Numerous lncRNAs have been discovered thanks to this high-throughput sequencing techniques. However, for the lncRNAs that only express in a specific cell type or a certain cellular state, information can easily be override in bulk data due to averaging transcripts from many types of cells. Recently, the emerging single cell omics technologies open another window for lncRNA study. Single cell RNA sequencing

(scRNA-seq) is one of the widely used techniques, which can reveal the transcriptome (including mRNAs and lncRNAs) heterogeneity of different cells (Fig. 2). By focusing on investigation of cell-type specific expression and gene regulatory networks (GRNs) for lncRNAs and mRNAs, transient state of cells and the heterogeneous in multicellular organisms can be resolved in single cell studies. To our best knowledge, there is still lack of attention for lncRNAs in single cell related studies. Therefore, in this review we discussed the state-of-the-art lncRNA studies which benefit from single cell sequencing techniques, with a special focus on studies related to lncRNA expression dynamics and lncRNA-associated GRNs at the single cell level. We also summarized the bioinformatics tools utilized for single cell GRN (scGRN) constructions, and proposed a computational framework for single cell multi-omics data integration on purpose of lncRNA scGRN analysis. In sum, scGRN analysis based on single cell multi-omics data will be an essential approach to understand lncRNA functions with an unprecedented resolution and accuracy.

2. Single-cell sequencing for lncRNA analysis: From gene to networks

Single cell transcriptomics (scRNA-seq) was first introduced in 2009 [22]. Afterwards, numerous of researches benefit from a series of single cell-based techniques. The assistive role of scRNA-seq for lncRNA studies is emerging. As one exciting example, scRNA-seq has been used to identify that the limb-expressed lncRNA *Maenli* co-expressed with the *En1* gene exclusively in the ectoderm cell cluster. Transcriptional activity of *Maenli* is required for limb-specific *En1* activation in *cis* for fine-tuning the GRNs controlling limb malformation in mouse and human [23]. In another study, myeloid signature lncRNA PIRAT was confirmed to be a CD14 + -
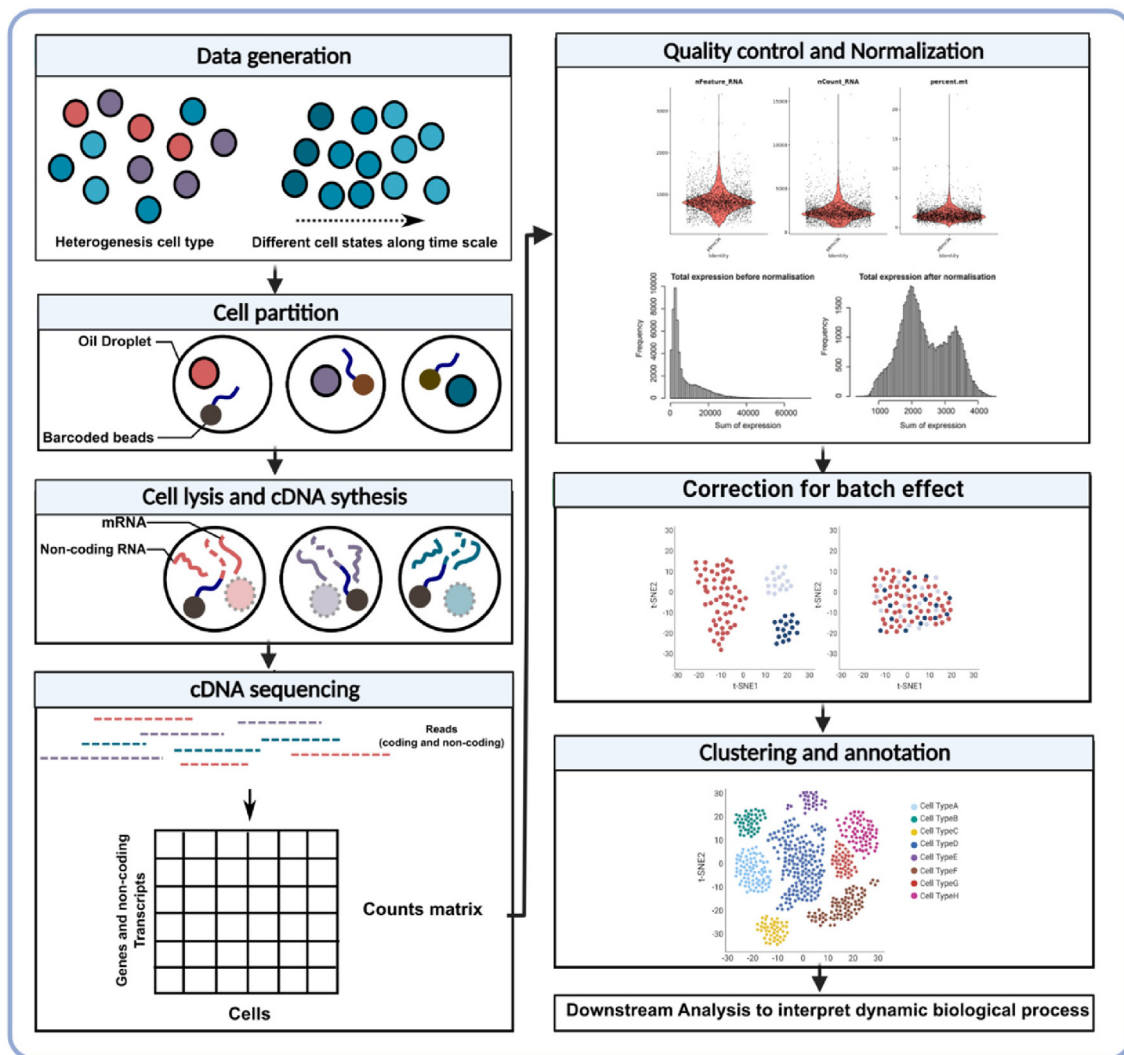
Fig. 2. Workflow of sing-cell RNA sequencing (scRNA-seq) data analysis. Both mRNAs and lncRNAs can be quantified at a single-cell resolution.

monocyte-specific myeloid lincRNA by customized scRNA-seq analysis in sever COVID-19 patients [24]. Here we focused on studies which use single cell techniques as a major strength to illustrate the contribution to lncRNA biology (Fig. 3).

2.1. An expanding lncRNA database predicted by single cell RNA sequencing

LncRNAs are newly defined transcripts that make up the majority of transcriptome. Until recently 270,044 lncRNAs have been reported in human [25], yet the database is still expanding rapidly. scRNA-seq is ideal technology for detection of short-lived, cell type specific lncRNAs that are previously neglected in bulk RNA-seq data. Pioneering studies of lncRNA prediction at the single level discovered 2,733 novel lncRNAs in 124 individual cells during human embryonic development, many of which were only expressed in specific developmental stages [26]. Later, See et al. found that 30.4 % of lincRNAs (intergenic lncRNAs) identified in 359 nuclei of cardiomyocytes are novel and only detectable by single nuclei transcriptome (snRNA-seq) but not by bulk transcriptom [27]. This indicates the advantage of scRNA-seq/snRNA-seq towards discovering nuclear retained lncRNAs that may be diluted by large pool of cytoplasmic mRNA generated from bulk RNA-seq. Nevertheless, attention needs to be paid in choosing snRNA-seq

since a large fraction of cytoplasmic lncRNAs that should not be overlooked. In this regard, snRNA-seq is obviously more suitable for exploration of nuclear retained lncRNAs. Recently, Luo et al. updated 9433 novel lncRNA transcripts in human T cells across three types of cancer with a full-length scRNA-seq technology SMART-seq2 and developed a computational pipeline for *de novo* transcriptome assembly [28]. Indeed, single-cell techniques have been questioned for the insufficient coverage of transcriptome in each cell, most of the annotated lncRNAs so far were identified by traditional bulk RNA-seq. However, there is still improving space for single cell sequencing techniques to uncover new lncRNAs in certain cell types or transient cell states. In this regard, combination of bulk RNA-seq and scRNA-seq could be necessary to catalogue thoroughly the whole lncRNA database in a given species.

2.2. lncRNA expression heterogeneity reveal by single-cell RNA sequencing

LncRNAs generally have lower expression than protein coding genes (PCGs) in bulk transcriptome analysis [17]. However, such observation could due to the dilution effect of cell mixtures with low average abundance of lncRNAs [29]. Firstly, Cabili et al. denied this hypothesis using single-molecule RNA-FISH to investigate the
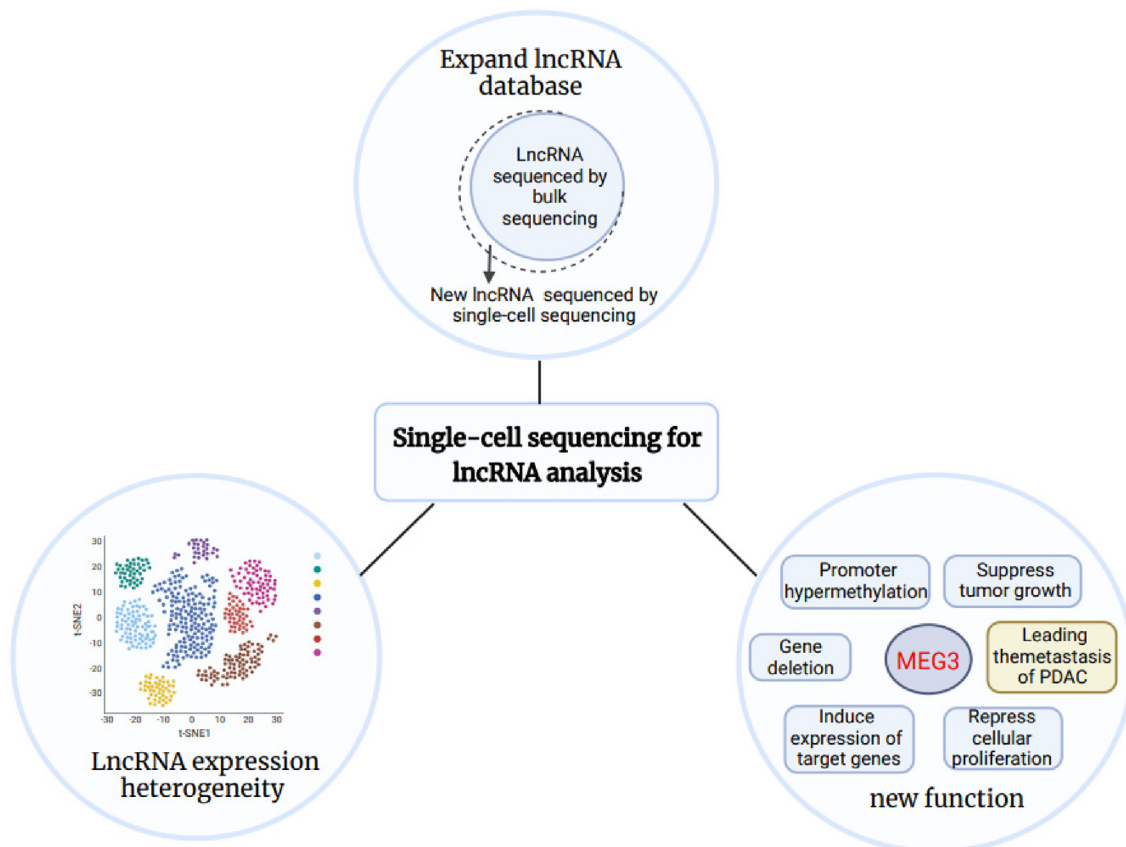
Fig. 3. Advancement of lncRNAs studies benefit from single-cell omics studies.

subcellular localization patterns of 61 lncRNAs in three different cell lines and reported the similar pattern of cell-to-cell variation between lncRNAs and PCGs [30]. However, several recent studies support the heterogenic expression of lncRNAs with more comprehensive single-cell transcriptome data analysis. For instance, Xue et al. reported that lncRNAs in kidney cancer cells are cancer type specific, and they tend to have lower expression level but higher differences compared to PCGs [31]. Similar observation was reported in human neocortex [32]. The analysis by Wang et al. showed the transcriptional or translational specific of lncRNAs and their cell-type specificity by function analysis [33]. Moreover, recently allele-sensitive scRNA-seq demonstrated that expression of lncRNAs has higher cell-to-cell variability compared with mRNA [34]. A fraction of cerebral cortex organoids lncRNAs are transiently expressed (TrEx). scRNA-seq detected those TrEx lncRNAs only exists in specific cell types [35]. The above studies together point out cell type specificity of lncRNAs, high cell-to-cell variation of lncRNA expression, indicating that lncRNAs may function in a cell-type specific way.

With enormous amount of data generated by scRNA-seq in different tissues in a species, it is appealing to integrate single-cell transcriptomic data from different tissues and conditions to draw a reference lncRNA expression landscape at the cell level. Similar attempt has been made by Wang et al. by collecting lncRNA-associated competing endogenous RNA (ceRNA) regulation network from 94 000 cells across 25 types of cancers and documented 9000 experimentally supported lncRNA biomarkers [36]. Indeed, one could expect a particular (or a set of) lncRNAs as new potential cell markers, each cell type or cells at certain stage could have a barcode-like lncRNAs signature which does not necessarily have clear biological function but consistently expressed in certain patterns. In fact, the ColorCells database contains greater than 200

scRNA-seq datasets from six species. Initial analyses revealed that lncRNAs could also be used to classify cell type comparable with mRNAs and there are even certain cell groups can only be identified by lncRNAs [37].

2.3. Versatile roles of lncRNAs revealed by single-cell gene regulatory networks

In addition to lncRNA expression, investigation of lncRNA regulation would allow to identify transcription factors (TFs) that may regulate the expression of a target lncRNA and the *cis*-regulatory elements relevant for the regulation. GRN analysis based on single-cell sequencing data is an emerging and powerful approach to dissect novel biological functions of lncRNAs in certain cell types. Here we endeavour to discuss recent advances of lncRNA functional exploration in cancer research using scGRN analyses.

MEG3 is a star oncogenic lncRNA that suppresses tumour cell proliferation in multiple types of cancer by regulating the major tumour suppressor genes p53 and Rb [38]. However, Pan et al. illustrated that MEG3 leads to metastasis of pancreatic ductal adenocarcinoma (PDAC) tumour cells in their scRNA-seq data analysis [39], suggested an unexpected role of MEG3 in cancer cell biology. Another example is lincRNAs-p21, which is a p53-activated lncRNA that triggers apoptosis. Yang et al. found that differentially expression of lincRNA-p21 in human non-small cell lung cancer cells can cause distinct affects, their scRNA-seq data showed that upregulation of lincRNA-p21 considerably inhibited cell apoptosis while the downregulation of lincRNA-p21 showed the opposite effect [40]. MALAT1 is a highly conserved lncRNA that is known for multiple roles in a context-depend way in different human cancer cell lines [41], LINC-PINT is a lncRNA interact with PRC2 and suppresses gene expression of cancer cell invasion [42]. Both of them were

found to have a new function in mediating neoadjuvant chemotherapy resistance in triple-negative breast cancer (TNBC) by scanning the lncRNA profiles of persistence sub-cluster at single cell resolution [43]. lncRNA GAS5, another multi-function suppressor, has been found in several cardiomyocytes nuclei lincRNAs (LINCMs) as hub regulators in heart disease related module in GRNs constructed by scRNA-seq. Further experiments confirmed the key regulatory role in coordinating the stress-response in cardiomyocyte [27]. Above all, scGRN has its own advantage in pinpointing the particular role a lncRNA played in specific modality and deciphering the biology function of cell-type specific lncRNAs in different circumstance.

3. Computational frameworks for lncRNA scGRN inference

3.1. Single-cell omics resources for lncRNA-contained scGRN construction

Despite the fact that a handful types of single-cell omics data (Table 1) are available for scGRN construction in theory, scRNA-seq data is still the most widely used resources currently. Knowing that different layers of single-cell omics data could be used in a combination manner to infer GRNs, we listed several single-cell omics data without exhaustively describe all the state-of-art techniques or go deep into technique details, since they have been reviewed nicely elsewhere [44]. Instead, we simply discussed the

ones which have potential to construct meritorious lncRNA-based GRNs.

The emerged techniques were sorted into two categories: mono- and multi-omics single-cell techniques. Mono-omics techniques are measurement of single layer information such as scRNA-seq that provides transcriptome information, scChIP-seq/scChIC-seq that provide specific TF binding information, scATAC-seq and scDNase-seq that provide chromatin accessibility information. Regarding to lncRNA-included GRN construction, mono-omics such as scRNA-seq can be used to build gene co-expression networks which may include both protein-coding genes and lncRNAs, while scChIP-seq can be used to predict potential TF binding sites (TFBSs) for lncRNAs in specific cell types. Multi-omics techniques on the other hand measure two or more types of information from one cell simultaneously, and therefore provide multi-layer information in a unified way. DR-seq, G&T-seq and SCTG are techniques that sequence genome and transcriptome from a single cell in parallel. DR-seq pre-amplify DNA and mRNA with in the same tube but G&T-seq capture mRNA using streptavidin magnetic beads before amplification. Genomic variations of lncRNAs may affect the expression of downstream genes through multiple ways [45]. Therefore, genomic features could be taken into account for GRN construction in lncRNA studies. SNARE-seq, sciCAR-seq and scCAT-seq provides combined information of chromatin accessibility and transcriptome. These data coupled with TF motif analysis, provide indirect evidence of chromatin binding for hundreds of

Table 1
Single cell omics used for lncRNA-based gene regulatory network analysis.

Mono single cell omics	
Genome	Single cell genome sequencing [46]
Transcriptome	scRNA-seq [22]
TFs binding	scChIP-seq [47], scChIC-seq [48]
Chromatin accessibility	scATAC-seq [49], scDNase-seq [50]
Multi single cell omics	
Genome and transcriptome	G&T-seq [51], DR-seq [52], SCTG [53], TARGET-Seq [54], SIDR [55]
Epigenome and transcriptome	SNARE-seq [56], scMT-seq [57], SCM&T-seq [58], scNMT-seq [59], sci-CAR [60], scCAT-seq [61]

Table 2
Recent computational methods for single-cell omics network analysis.

Bioinformatics tools	Features	Methods/Models	Reference
SCNS toolkit	limited in small-size GRNs.	Boolean network	[62]
GENIE3	Follow the expression of all genes can be summarized as a simple weighted linear equation	Random forest regression	[67]
SCODE	Focuses on a series of discrete states to capture the dynamics of the network	Ordinary differential equation (ODE)	[86]
SCENIC	Combination of co-expression network and *cis*-regulatory analysis	Random forest regression (GENIE3) & motif enrichment	[72]
LinkedSOMs	Coupled scATAC-seq and scRNA-seq, generates chromatin and gene expression maps separately and combines them using a linking function	Self-organizing map (SOM)	[79]
Coupled NMF	A systematic mapping of *cis*-regulatory elements (CREs)to genes and TFs to CREs, informative for downstream analyses	Nonnegative matrix factorizations (NMF)	[80]
SINGE	Ordered scRNA-seq data pseudotemorally before construction of GRNs	Kernel-based Granger causality regression	[73]
CNNC	Supervised by small set of labeled positive gene pairs for gene-gene relationship predictions	Convolutional neural network	[75]
scTenifoldNet	Compare constructed scGRNs from two samples to detect changes in gene expression	Principal component regression & tensor decomposition & manifold alignment	[69]
DGRNS	Hybrid two deep learning method for GRN inference, exploring both time-dependent and spatially related information	recurrent neural network & convolutional neural network	[76]
SCGRNs	Integrated three machine learning approaches to infer the regulation network of various diseases	Tree boosting & support vector machine (SVM) & deepboost	[87]
scGNN	Modeling cell–cell relationships and their underlying complex gene expression pattern	Graph neural network & left-truncated mixture Gaussian (LTMG)	[88]
KPNN	Modified generic neural network to enhance the interpretability of the network	Knowledge-primed neural networks	[89]

TFs at a genome-wide manner and are thus useful to link TFs and their target lncRNAs.

3.2. Current computational tools for scGRN construction and their potential applications to lncRNAs analysis

The aim of scGRN construction is to investigate the regulatory mechanism of gene expression dynamics either in specific cell types or during the transition of different cell states. To achieve this goal, various algorithms have been developed to infer the relationship of different transcripts based on their expression patterns from scRNA-seq data. Mathematical models underling these tools include Boolean networks [62], Bayesian networks [63], ordinary differential equations (ODEs) [64], the information theory [65], regression-based [66,67] or correlation-based models [65,68], and more recently machine learning methods [69] (reviewed in [70,71]). To avoid repetitive documentation, we focused on the recent published frameworks of scGRN construction (Table 2). SCE-NIC is one of the widely used methods which combines results from co-expression analysis and predition of TF binding sites, which enables to identify key regulators of cell identities [72]. SINGE is a novel method developed to order the cells over irregular pseudotime using modified Granger causality and to ensemble them for GRN inference, which are reported to be more resilient to dropout events -- the major challenge in scRNA-seq data analysis [73]. Similarly, GRISLI, an RNA velocity estimator infers GRNs
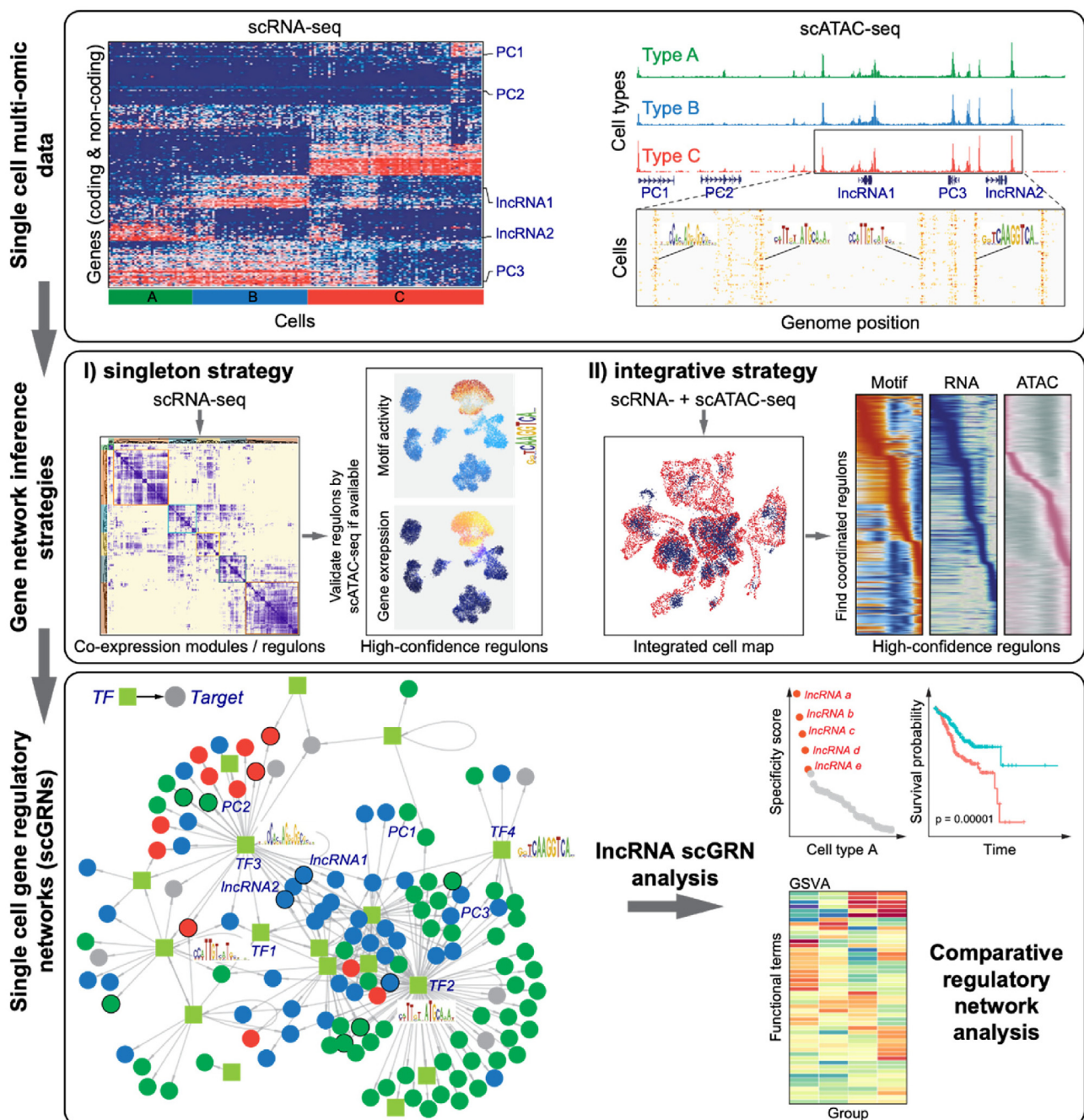

Fig. 4. Concept of integrative strategies for single-cell omics constructed gene regulation network for lncRNA analysis, I) singleton strategy, reproduced with permission [84], copyright 2022, Springer Nature. II) integrative strategy, reproduced with permission [85], copyright 2021 Elsevier Inc.

through dynamics of cell trajectories [74]. To be noted, an increasing number of current scGRN reconstruction tools were developed based on deep learning (DL) models (Table 2), Convolutional neural network (CNN) is one of the frequently used models. CNNC is a typical CNN-based approach for reconstructing GRN with training datasets [75], while DGRNS is a hybrid DL framework with the combination of CNN and recurrent neural network (RNN) [76]. Interestingly, scTenifoldNet was built to compare scGRNs between two different samples and to detect regulatory rewriting events. scGNN is a powerful multi-model-based tool that are hypothesis-free, which means that dropout events and statistical distribution of data won't be presumed to formulate cell–cell relationships and gene expression patterns.

Yet, in order to obtain high-confidential cell-type specific gene regulations, the inferred regulatory relationship need to be validated using epigenomic data such as chromatin accessibility in matched cell types (Fig. 4; singleton strategy). Whenever possible, scGRN interpretation accuracy can and should be improved by pairing single cell transcriptomic and epigenomic data. In order to mine gene regulation at the cell-type level in a multi-omics manner, one of the widely used approach is based on paired single-cell expression and accessibility analysis, where scRNA-seq and scATAC-seq experiments are performed on two different samples of cells from the same cell populations [77,78]. Otherwise, multi-omics techniques listed in Table 1 have advantage towards obtaining matched modalities of each cell since the DNA/RNA sequencing are extracted for the same cell. In terms of analysis, joint clustering and embedding is firstly performed to identify cell populations with distinct expression and accessibility profiles. In this way, the integrated cell map provides both expression and regulation information of investigated genes for a given cell population. Then cell population-specific regulatory relations can be inferred using gene expression profile, chromatin accessibility of regulatory elements, as well as TF-motif enrichment scores in accessible regulatory elements (Fig. 4; integrative strategy). Based on the above principle, several computational tools such as Linked-SOMs [79] and Coupled NMF [80] are recently developed to integrate and analyse scRNA-seq and scATAC-seq data for scGRN inference (Table 2). Construction of scGRNs based on multi-omics data is more challenging than singleton strategy, since robust integration of multimodal data is an essential prerequisite for GRN inference. Several reviews well summarized algorithm and bioinformatics tools for single cell omics data integration [81,82]. Besides that, although a range of computational approaches aimed at jointly analysing different layers of omics data were developed as mentioned above, there is no all-in-one solution of analytical tools for scGRN inference based on multi-omics data so far.

Given the emerging patterns of lncRNAs in cell-type specific expression and regulation [24,33,35,43,83], it would be meaningful to build lncRNA-associated GRN in a systematic manner for lncRNA studies. The strategies for scGRN inference discussed above can directly be applied for lncRNA-associated scGRN inference by including both PCGs and lncRNAs into the analysis. The resulted GRNs could be a collection of gene pairs in a cell-type specific manner where TFs are upstream regulators that regulate target genes of either PCs or lncRNAs (Fig. 4). Based on the inferred networks, the functions of lncRNAs can be predicted through network analysis and/or derived from co-regulated PCGs (Fig. 4). For example, we can prioritize the functional importance of lncRNAs by node degree analysis. Biological pathways that lncRNAs may involve in can be estimated based on functional enrichment analysis using co-regulated PCGs. Moreover, lncRNA-based GRNs can be compared under different conditions to reveal functional specificity of lncRNAs using comparative regulatory analysis tools such as Sc-compReg and scTenifoldNet [78]. scGNN could be an applicable tool as well because of its hypothesis-free strategy.

4. Challenges and future perspectives

Single-cell multi-omics data are powerful recourses to study the expression and regulation of genes as well as the highly cell-type specific lncRNAs. However, current development of single-cell technologies may not yet be optimized for lncRNA detection and quantification. For example, only a minority of expressed genes can be detected in a cell by scRNA-seq due to dropout issue, which remains a big obstacle for single cell techniques-based studies. Since lncRNAs generally have lower expression level than protein-coding genes, the dropout events would affect more the analysis. Full-length techniques like Smart-seq2 outperformed drop-based techniques like 10X Genomics in term of sensitivity [90], and they should be better options for lncRNA studies. Recently, Bocchi et al. utilized bulk RNA-seq to discover lncRNAs in neocortex then combined scRNA-seq analysis of cell-type specific GRNs to decrypt the co-regulation network with novel identified lncRNAs and protein-coding genes [32]. This provides an alternative idea of impute low coverage of scRNA-seq data.

GRN analysis at the single-cell level provides an important strategy to understand the molecular principle that controls cellular differentiation and cell type or state transitions. However, inferring GRNs from single-cell omics data is of great computational challenges owing to several aspects of difficulties, including substantial cellular heterogeneity [91], cell-to-cell variation in sequencing depth, the high sparsity caused by dropouts [63] and cell-cycle-related effects from the data [92]. There are even more challenges for lncRNA-associated scGRN inference since the estimated lncRNA expression information is sparse and noise. Although a dozen of algorithms (Table 2) have been developed or used to infer GRNs from single-cell data, none of these methods has been tested or evaluated for lncRNA-based GRN mapping. Single-cell expression data are especially useful for lncRNA expression analysis and thus for constructing scGRNs. However, there are inherent shortcomings of using expression correlations to infer regulatory interactions. In this regard, integration analysis of multi-omics data could reduce the impacts of noise and enhance the performance of GRN inference by cross-validating the regulatory interactions in multiple datasets [71], and thus would be promising for the reliable inference of lncRNA-based GRNs. Based on the fact that lncRNAs are an emerging type of players in GRNs [93] and largely less investigated compared to protein-coding genes, lncRNA-based scGRN analysis would provide a frontier area for lncRNA biology in the near future.

CRediT authorship contribution statement

Xue Zhao: Writing – original draft, Writing – review & editing. **Yangming Lan:** . **Dijun Chen:** Conceptualization, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Ransohoff JD, Wei Y, Khavari PA. The functions and unique features of long intergenic non-coding RNA. Nat Rev Mol Cell Biol 2018;19:143–57. https://doi.org/10.1038/nrm.2017.104.

[2] Tuck AC, Natarajan KN, Rice GM, Borawski J, Mohn F, Rankova A, et al. Distinctive features of lincRNA gene expression suggest widespread RNA-independent functions. Life Science Alliance 2018;1:e201800124. 10.26508/lsa.201800124.

[3] Fedak H, Palusinska M, Krzyczmonik K, Brzezniak L, Yatusevich R, Pietras Z, et al. Control of seed dormancy in Arabidopsis by a cis -acting noncoding antisense transcript. Proceedings of the National Academy of Sciences 2016;113. 10.1073/pnas.1608827113.

[4] Heo JB, Sung S. Vernalization-Mediated Epigenetic Silencing by a Long Intronic Noncoding RNA. Science 1979;2011(331):76–9. https://doi.org/10.1126/science.1197349.

[5] Kim E-D, Sung S. Long noncoding RNA: unveiling hidden layer of gene regulatory networks. Trends Plant Sci 2012;17:16–21. https://doi.org/10.1016/j.tplants.2011.10.008.

[6] Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, et al. Target mimicry provides a new mechanism for regulation of microRNA activity. Nat Genet 2007;39:1033–7. https://doi.org/10.1038/ng2079.

[7] Zou Y, Chen B. Long non-coding RNA HCP5 in cancer. Clin Chim Acta 2021;512:33–9. https://doi.org/10.1016/j.cca.2020.11.015.

[8] Lin R, Maeda S, Liu C, Karin M, Edgington TS. A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. Oncogene 2007;26:851–8. https://doi.org/10.1038/sj.onc.1209846.

[9] Lazorthes S, Vallot C, Briois S, Aguirrebengoa M, Thuret J-Y, Laurent G, et al. A vlincRNA participates in senescence maintenance by relieving H2AZ-mediated repression at the INK4 locus. Nature. Communications 2015;6:5971. https://doi.org/10.1038/ncomms6971.

[10] Koerner MV, Pauler FM, Huang R, Barlow DP. The function of non-coding RNAs in genomic imprinting. Development 2009;136:1771–83. https://doi.org/10.1242/dev.030403.

[11] Postepska-Igielska A, Giwojna A, Gasri-Plotnitsky L, Schmitt N, Dold A, Ginsberg D, et al. LncRNA Khps1 Regulates Expression of the Proto-oncogene SPHK1 via Triplex-Mediated Changes in Chromatin Structure. Mol Cell 2015;60:626–36. https://doi.org/10.1016/j.molcel.2015.10.001.

[12] Atianand MK, Hu W, Satpathy AT, Shen Y, Ricci EP, Alvarez-Dominguez JR, et al. A Long Noncoding RNA lincRNA-EPS Acts as a Transcriptional Brake to Restrain Inflammation. Cell 2016;165:1672–85. https://doi.org/10.1016/j.cell.2016.05.075.

[13] Zhang Y, He Q, Hu Z, Feng Y, Fan L, Tang Z, et al. Long noncoding RNA LINP1 regulates repair of DNA double-strand breaks in triple-negative breast cancer. Nat Struct Mol Biol 2016;23:522–30. https://doi.org/10.1038/nsmb.3211.

[14] An Y, Furber KL, Ji S. Pseudogenes regulate parental gene expression via ceRNA network. J Cell Mol Med 2017;21:185–92. https://doi.org/10.1111/jcmm.12952.

[15] Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, et al. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. Science 2005;309:1570–3. https://doi.org/10.1126/science.1115901.

[16] Ninomiya K, Iwakiri J, Aly MK, Sakaguchi Y, Adachi S, Natsume T, et al. m [6] A modification of HSATIII lncRNAs regulates temperature-dependent splicing. The EMBO Journal 2021;40. 10.15252/embj.2021107976.

[17] Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 2011;25:1915–27. https://doi.org/10.1101/gad.17446611.

[18] Rappoport N, Shamir R. Multi-omic and multi-view clustering algorithms: review and cancer benchmark. Nucleic Acids Res 2018;46:10546–62. https://doi.org/10.1093/nar/gky889.

[19] Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 2012;22:1760–74. https://doi.org/10.1101/gr.135350.111.

[20] Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. Nat Rev Genet 2016;17:47–62. https://doi.org/10.1038/nrg.2015.10.

[21] Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009;458:223–7. https://doi.org/10.1038/nature07672.

[22] Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods 2009;6:377–82. https://doi.org/10.1038/nmeth.1315.

[23] Allou L, Balzano S, Magg A, Quinodoz M, Royer-Bertrand B, Schöpflin R, et al. Non-coding deletions identify Maenli lncRNA as a limb-specific En1 regulator. Nature 2021. https://doi.org/10.1038/s41586-021-03208-9.

[24] Aznaourova M, Schmerer N, Janga H, Zhang Z, Pauck K, Hoppe J, et al. Single cell RNA-seq uncovers the nuclear decoy lincRNA PIRAT as a regulator of systemic monocyte immunity during COVID-19 n.d. 10.1101/2021.11.05.467458.

[25] Ma L, Cao J, Liu L, Du Q, Li Z, Zou D, et al. LncBook: a curated knowledgebase of human long non-coding RNAs. Nucleic Acids Res 2019;47:D128–34. https://doi.org/10.1093/nar/gky960.

[26] Yan L, Yang M, Guo H, Yang L, Wu J, Li R, et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol 2013;20:1131–9. https://doi.org/10.1038/nsmb.2660.

[27] See K, Tan WLW, Lim EH, Tiang Z, Lee LT, Li PYQ, et al. Single cardiomyocyte nuclear transcriptomes reveal a lincRNA-regulated de-differentiation and cell cycle stress-response in vivo. Nature. Communications 2017;8. https://doi.org/10.1038/s41467-017-00319-8.

[28] Luo H, Bu D, Shao L, Li Y, Sun L, Wang C, et al. Single-cell Long Non-coding RNA Landscape of T Cells in Human Cancer Immunity. Genomics, Proteomics and Bioinformatics 2021;19:377–93. https://doi.org/10.1016/j.gpb.2021.02.006.

[29] Dinger ME, Amaral PP, Mercer TR, Mattick JS. Pervasive transcription of the eukaryotic genome: functional indices and conceptual implications. Brief Funct Genomics Proteomics 2009;8:407–23. https://doi.org/10.1093/bfgp/elp038.

[30] Cabili MN, Dunagin MC, McClanahan PD, Biaesch A, Padovan-Merhar O, Regev A, et al. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. Genome Biol 2015;16:20. https://doi.org/10.1186/s13059-015-0586-4.

[31] Li X, Meng X, Wei C, Zhou Y, Chen H, Huang H, et al. Dissecting LncRNA Roles in Renal Cell Carcinoma Metastasis and Characterizing Genomic Heterogeneity by Single-Cell RNA-seq. Mol Cancer Res 2018;16:1879–88. https://doi.org/10.1158/1541-7786.MCR-17-0776.

[32] Bocchi VD, Conforti P, Vezzoli E, Besusso D, Cappadona C, Lischetti T, et al. The coding and long noncoding single-cell atlas of the developing human fetal striatum. Science 1979;2021:372. https://doi.org/10.1126/science.abf5759.

[33] Wang H, Wang Y, Xie S, Liu Y, Xie Z. Global and cell-type specific properties of lincRNAs with ribosome occupancy. Nucleic Acids Res 2016. https://doi.org/10.1093/nar/gkw909.

[34] Johnsson P, Ziegenhain C, Hartmanis L, Hendriks G-J, Hagemann-Jensen M, Reinius B, et al. Transcriptional kinetics and molecular functions of long noncoding RNAs. Nat Genet 2022;54:306–17. https://doi.org/10.1038/s41588-022-01014-1.

[35] Field AR, Jacobs FMJ, Fiddes IT, Phillips APR, Reyes-Ortiz AM, LaMontagne E, et al. Structurally Conserved Primate LncRNAs Are Transiently Expressed during Human Cortical Differentiation and Influence Cell-Type-Specific Genes. Stem Cell Rep 2019;12:245–57. https://doi.org/10.1016/j.stemcr.2018.12.006.

[36] Wang P, Guo Q, Hao Y, Liu Q, Gao Y, Zhi H, et al. LnCeCell: a comprehensive database of predicted lncRNA-associated ceRNA networks at single-cell resolution. Nucleic Acids Res 2021;49:D125–33. https://doi.org/10.1093/nar/gkaa1017.

[37] Zheng L-L, Xiong J-H, Zheng W-J, Wang J-H, Huang Z-L, Chen Z-R, et al. ColorCells: a database of expression, classification and functions of lncRNAs in single cells. Briefings Bioinf 2021;22. https://doi.org/10.1093/bib/bbaa325.

[38] Al-Rugeebah A, Alanazi M, Parine NR. MEG3: an Oncogenic Long Non-coding RNA in Different Cancers. Pathol Oncol Res 2019;25:859–74. https://doi.org/10.1007/s12253-019-00614-3.

[39] Pan H, Diao H, Zhong W, Wang T, Wen P, Wu C. A Cancer Cell Cluster Marked by LincRNA MEG3 Leads Pancreatic Ductal Adenocarcinoma Metastasis. Frontiers in Oncology 2021;11. 10.3389/fonc.2021.656564.

[40] Yang T, Zhang W, Wang L, Xiao C, Guo B, Gong Y, et al. Long intergenic noncoding RNA-p21 inhibits apoptosis by decreasing PUMA expression in non-small cell lung cancer. J Int Med Res 2019;47:481–93. https://doi.org/10.1177/0300060518816592.

[41] Yoshimoto R, Mayeda A, Yoshida M, Nakagawa S. MALAT1 long non-coding RNA in cancer. Biochimica et Biophysica Acta (BBA). - Gene Regulatory Mechanisms 2016;1859:192–9. https://doi.org/10.1016/j.bbagrm.2015.09.012.

[42] Marín-Béjar O, Mas AM, González J, Martinez D, Athie A, Morales X, et al. The human lncRNA LINC-PINT inhibits tumor cell invasion through a highly conserved sequence element. Genome Biol 2017;18:202. https://doi.org/10.1186/s13059-017-1331-y.

[43] Shaath H, Vishnubalaji R, Elango R, Khattak S, Alajez NM. Single-cell long noncoding RNA (lncRNA) transcriptome implicates MALAT1 in triple-negative breast cancer (TNBC) resistance to neoadjuvant chemotherapy. Cell Death Discovery 2021;7:23. https://doi.org/10.1038/s41420-020-00383-y.

[44] Lee J, Hyeon DY, Hwang D. Single-cell multiomics: technologies and data analysis methods. Exp Mol Med 2020;52:1428–42. https://doi.org/10.1038/s12276-020-0420-2.

[45] Wang P, Li X, Gao Y, Guo Q, Ning S, Zhang Y, et al. LnCeVar: a comprehensive database of genomic variations that disturb ceRNA network regulation. Nucleic Acids Res 2019. https://doi.org/10.1093/nar/gkz887.

[46] Gawad C, Koh W, Quake SR. Single-cell genome sequencing: current state of the science. Nat Rev Genet 2016;17:175–88. https://doi.org/10.1038/nrg.2015.16.

[47] Rotem A, Ram O, Shoresh N, Sperling RA, Goren A, Weitz DA, et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat Biotechnol 2015;33:1165–72. https://doi.org/10.1038/nbt.3383.

[48] Ku WL, Nakamura K, Gao W, Cui K, Hu G, Tang Q, et al. Single-cell chromatin immunocleavage sequencing (scChIC-seq) to profile histone modification. Nat Methods 2019;16:323–5. https://doi.org/10.1038/s41592-019-0361-7.

[49] Chen H, Lareau C, Andreani T, Vinyard ME, Garcia SP, Clement K, et al. Assessment of computational methods for the analysis of single-cell ATAC-seq data. Genome Biol 2019;20. https://doi.org/10.1186/s13059-019-1854-5.

[50] Jin W, Tang Q, Wan M, Cui K, Zhang Y, Ren G, et al. Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. Nature 2015;528:142–6. https://doi.org/10.1038/nature15740.

[51] Macaulay IC, Teng MJ, Haerty W, Kumar P, Ponting CP, Voet T. Separation and parallel sequencing of the genomes and transcriptomes of single cells using

G&T-seq. Nat Protoc 2016;11:2081–103. https://doi.org/10.1038/nprot.2016.138.

[52] Dey SS, Kester L, Spanjaard B, Bienko M, van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. Nat Biotechnol 2015;33:285–9. https://doi.org/10.1038/nbt.3129.

[53] Li W, Calder RB, Mar JC, Vijg J. Single-cell transcriptogenomics reveals transcriptional exclusion of ENU-mutated alleles. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 2015;772:55–62. https://doi.org/10.1016/j.mrfmmm.2015.01.002.

[54] Rodriguez-Meira A, Buck G, Clark S-A, Povinelli BJ, Alcolea V, Louka E, et al. Unravelling Intratumoral Heterogeneity through High-Sensitivity Single-Cell Mutational Analysis and Parallel RNA Sequencing. Mol Cell 2019;73:1292–1305.e8. https://doi.org/10.1016/j.molcel.2019.01.009.

[55] Han KY, Kim K-T, Joung J-G, Son D-S, Kim YJ, Jo A, et al. SIDR: simultaneous isolation and parallel sequencing of genomic DNA and total RNA from single cells. Genome Res 2018;28:75–87. https://doi.org/10.1101/gr.223263.117.

[56] Chen S, Lake BB, Zhang K. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. Nat Biotechnol 2019;37:1452–7. https://doi.org/10.1038/s41587-019-0290-0.

[57] Hu Y, Huang K, An Q, Du G, Hu G, Xue J, et al. Simultaneous profiling of transcriptome and DNA methylome from a single cell. Genome Biol 2016;17:88. https://doi.org/10.1186/s13059-016-0950-z.

[58] Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. Nat Methods 2016;13:229–32. https://doi.org/10.1038/nmeth.3728.

[59] Clark SJ, Argelaguet R, Kapourani C-A, Stubbs TM, Lee HJ, Alda-Catalinas C, et al. scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. Nat Commun 2018;9:781. https://doi.org/10.1038/s41467-018-03149-4.

[60] Cao J, Cusanovich DA, Ramani V, Aghamirzaie D, Pliner HA, Hill AJ, et al. Joint profiling of chromatin accessibility and gene expression in thousands of single cells. Science 1979;2018(361):1380–5. https://doi.org/10.1126/science.aau0730.

[61] Liu L, Liu C, Quintero A, Wu L, Yuan Y, Wang M, et al. Deconvolution of single-cell multi-omics layers reveals regulatory heterogeneity. Nat Commun 2019;10:470. https://doi.org/10.1038/s41467-018-08205-7.

[62] Moignard V, Woodhouse S, Haghverdi L, Lilly AJ, Tanaka Y, Wilkinson AC, et al. Decoding the regulatory network of early blood development from single-cell gene expression measurements. Nat Biotechnol 2015;33:269–76. https://doi.org/10.1038/nbt.3154.

[63] Kharchenko PV, Silberstein L, Scadden DT. Bayesian approach to single-cell differential expression analysis. Nat Methods 2014;11:740–2. https://doi.org/10.1038/nmeth.2967.

[64] Kana O, Bhattacharya S. Single Cell Universal Differential Equations: A Machine Learning Framework for Extracting Ordinary Differential Equations for Gene Regulatory Network Inference. The FASEB Journal 2021;35:fasebj.2021.35.S1.03839. 10.1096/fasebj.2021.35.S1.03839.

[65] Chan TE, Stumpf MPH, Babtie AC. Gene regulatory network inference from single-cell data using multivariate information measures. Cell Systems 2017;5:251–267.e3. https://doi.org/10.1016/j.cels.2017.08.014.

[66] Papili Gao N, Ud-Dean SMM, Gandrillon O, Gunawan R. SINCERITIES: inferring gene regulatory networks from time-stamped single cell transcriptional expression profiles. Bioinformatics 2018;34:258–66. https://doi.org/10.1093/bioinformatics/btx575.

[67] Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. Inferring regulatory networks from expression data using tree-based methods. PLoS ONE 2010;5:e12776.

[68] Specht AT, Li JLEAP. constructing gene co-expression networks for single-cell RNA-sequencing data using pseudotime ordering. Bioinformatics 2016. https://doi.org/10.1093/bioinformatics/btw729.

[69] Osorio D, Zhong Y, Li G, Huang JZ, Cai JJ. scTenifoldNet: A Machine Learning Workflow for Constructing and Comparing Transcriptome-wide Gene Regulatory Networks from Single-Cell Data. Patterns 2020;1:. https://doi.org/10.1016/j.patter.2020.100139100139.

[70] Fiers MWEJ, Minnoye L, Aibar S, Bravo González-Blas C, Kalender Atak Z, Aerts S. Mapping gene regulatory networks from single-cell omics data. Briefings in Functional Genomics 2018;17:246–54. https://doi.org/10.1093/bfgp/elx046.

[71] Hu X, Hu Y, Wu F, Leung RWT, Qin J. Integration of single-cell multi-omics for gene regulatory network inference. Comput Struct Biotechnol J 2020;18:1925–38. https://doi.org/10.1016/j.csbj.2020.06.033.

[72] Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods 2017;14:1083–6. https://doi.org/10.1038/nmeth.4463.

[73] Deshpande A, Chu L-F, Stewart R, Gitter A. Network inference with Granger causality ensembles on single-cell transcriptomics. Cell Reports 2022;38:. https://doi.org/10.1016/j.celrep.2022.110333110333.

[74] Aubin-Frankowski P-C, Vert J-P. Gene regulation inference from single-cell RNA-seq data with linear differential equations and velocity inference. Bioinformatics 2020;36:4774–80. https://doi.org/10.1093/bioinformatics/btaa576.

[75] Yuan Y, Bar-Joseph Z. Deep learning for inferring gene relationships from single-cell expression data. Proc Natl Acad Sci 2019;116:27151–8. https://doi.org/10.1073/pnas.1911536116.

[76] Zhao M, He W, Tang J, Zou Q, Guo F. A hybrid deep learning framework for gene regulatory network inference from single-cell transcriptomic data. Briefings Bioinf 2022;23. https://doi.org/10.1093/bib/bbab568.

[77] Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive Integration of Single-Cell Data. Cell 2019;177:1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

[78] Duren Z, Lu WS, Arthur JG, Shah P, Xin J, Meschi F, et al. Sc-compReg enables the comparison of gene regulatory networks between conditions using single-cell data. Nat Commun 2021;12:4763. https://doi.org/10.1038/s41467-021-25089-2.

[79] Jansen C, Ramirez RN, El-Ali NC, Gomez-Cabrero D, Tegner J, Merkenschlager M, et al. Building gene regulatory networks from scATAC-seq and scRNA-seq using Linked Self Organizing Maps. PLoS Comput Biol 2019;15:e1006555.

[80] Duren Z, Chen X, Zamanighomi M, Zeng W, Satpathy AT, Chang HY, et al. Integrative analysis of single-cell genomics data by coupled nonnegative matrix factorizations. Proc Natl Acad Sci 2018;115:7723–8. https://doi.org/10.1073/pnas.1805681115.

[81] Ma A, McDermaid A, Xu J, Chang Y, Ma Q. Integrative Methods and Practical Challenges for Single-Cell Multi-omics. Trends Biotechnol 2020;38:1007–22. https://doi.org/10.1016/j.tibtech.2020.02.013.

[82] Adossa N, Khan S, Rytkönen KT, Elo LL. Computational strategies for single-cell multi-omics integration. Comput Struct Biotechnol J 2021;19:2588–96. https://doi.org/10.1016/j.csbj.2021.04.060.

[83] Liu SJ, Horlbeck MA, Cho SW, Birk HS, Malatesta M, He D, et al. CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. Science 1979;2017:355. https://doi.org/10.1126/science.aah7111.

[84] Qu J, Yang F, Zhu T, Wang Y, Fang W, Ding Y, et al. A reference single-cell regulomic and transcriptomic map of cynomolgus monkeys. Nat Commun 2022;13:4069. https://doi.org/10.1038/s41467-022-31770-x.

[85] Marand AP, Chen Z, Gallavotti A, Schmitz RJ. A cis-regulatory atlas in maize at single-cell resolution. Cell 2021;184:3041–3055.e21. https://doi.org/10.1016/j.cell.2021.04.014.

[86] Matsumoto H, Kiryu H, Furusawa C, Ko MSH, Ko SBH, Gouda N, et al. SCODE: an efficient regulatory network inference algorithm from single-cell RNA-Seq during differentiation. Bioinformatics 2017;33:2314–21. https://doi.org/10.1093/bioinformatics/btx194.

[87] Turki T, Taguchi Y. SCGRNs: Novel supervised inference of single-cell gene regulatory networks of complex diseases. Comput Biol Med 2020;118:. https://doi.org/10.1016/j.compbiomed.2020.103656103656.

[88] Wang J, Ma A, Chang Y, Gong J, Jiang Y, Qi R, et al. scGNN is a novel graph neural network framework for single-cell RNA-Seq analyses. Nat Commun 2021;12:1882. https://doi.org/10.1038/s41467-021-22197-x.

[89] Fortelny N, Bock C. Knowledge-primed neural networks enable biologically interpretable deep learning on single-cell sequencing data. Genome Biol 2020;21:190. https://doi.org/10.1186/s13059-020-02100-5.

[90] Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, et al. Comparative Analysis of Single-Cell RNA Sequencing Methods. Mol Cell 2017;65:631–643.e4. https://doi.org/10.1016/j.molcel.2017.01.023.

[91] Wagner A, Regev A, Yosef N. Revealing the vectors of cellular identity with single-cell genomics. Nat Biotechnol 2016;34:1145–60. https://doi.org/10.1038/nbt.3711.

[92] Buettner F, Natarajan KN, Casale FP, Proserpio V, Scialdone A, Theis FJ, et al. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. Nat Biotechnol 2015;33:155–60. https://doi.org/10.1038/nbt.3102.

[93] Anastasiadou E, Jacob LS, Slack FJ. Non-coding RNA networks in cancer. Nat Rev Cancer 2018;18:5–18. https://doi.org/10.1038/nrc.2017.99.