ZOOLOGICAL RESEARCH

Reconstructing gene regulatory networks in single-cell transcriptomic data analysis

Hao Dai[1,2], Qi-Qi Jin[1,3,4], Lin Li[1,3], Luo-Nan Chen[1,4,5,6,*]

[1] *Key Laboratory of Systems Biology*, *Shanghai Institute of Biochemistry and Cell Biology*, *Center for Excellence in Molecular Cell Science*, *Chinese Academy of Sciences*, *Shanghai* 200031, *China*

[2] *Institute of Brain-Intelligence Technology*, *Zhangjiang Laboratory*, *Shanghai* 201210, *China*

[3] *University of Chinese Academy of Sciences*, *Beijing* 100049, *China*

[4] *School of Life Science and Technology*, *ShanghaiTech University*, *Shanghai* 201210, *China*

[5] *Key Laboratory of Systems Biology*, *Hangzhou Institute for Advanced Study*, *University of Chinese Academy of Sciences*, *Chinese Academy of Sciences*, *Hangzhou*, *Zhejiang* 310024, *China*

[6] *Center for Excellence in Animal Evolution and Genetics*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650223, *China*

ABSTRACT

Gene regulatory networks play pivotal roles in our understanding of biological processes/mechanisms at the molecular level. Many studies have developed sample-specific or cell-type-specific gene regulatory networks from single-cell transcriptomic data based on a large amount of cell samples. Here, we review the state-of-the-art computational algorithms and describe various applications of gene regulatory networks in biological studies.

Keywords: Gene regulatory network; Single-cell RNA sequencing; Computational algorithm; Sample-specific network; Cell-type-specific network; Cell-specific network

INTRODUCTION

Single-cell RNA sequencing (scRNA-seq) technology has made it possible to measure and compare gene transcriptomic profiles at single-cell resolution (Eberwine et al., 2014; Stegle et al., 2015). Based on scRNA-seq data, new cell types with distinct functions can be identified and cellular lineages during differentiation can be traced (Rozenblatt-Rosen et al., 2017; Villani et al., 2017). Many studies have focused on developing accurate and robust computational methods for scRNA-seq data analysis (Zeng & Dai, 2019), where a key problem is how to construct gene regulatory networks (GRNs) to pinpoint crucial factors, e.g., those that control cellular differentiation and determine phenotypes in disease progression (Iacono et al., 2019). In fact, scRNA-seq technology provides a large number of cell samples, making it possible to study gene-gene associations and transcriptional networks accurately (Dai et al., 2019). At present, many studies and algorithms have been developed to construct GRNs from scRNA-seq data based on various principles and perspectives.

INFERENCE METHODS FOR GRNs

Correlation networks

Correlation network analysis is one of the most widely used methods for scRNA-seq data. These networks measure gene-gene associations based on correlation coefficients and are

Received: 03 August 2020; Accepted: 20 October 2020; Online: 21 October 2020

Foundation items: This study was supported by the National Key Research and Development Program of China (2017YFA0505500), Strategic Priority Research Program of the Chinese Academy of Sciences (XDB38040400), National Science Foundation of China (31771476 and 31930022), and Shanghai Municipal Science and Technology Major Project (2017SHZDZX01)

*Corresponding author, E-mail: lnchen@sibs.ac.cn

DOI: 10.24272/j.issn.2095-8137.2020.215

suitable for large and high-dimensional datasets. In 2008, Langfelder & Horvath (2008) presented the popular weighted gene co-expression network (WGCNA) for weighted correlation network analysis. This method detects modules of highly correlated genes, identifies hub genes within these modules, and measures the relationships among the modules. Since then, many improved algorithms have been presented. For example, the partial and semi-partial correlation (PPCOR) method (Kim, 2015) substitutes the correlation coefficient in WGCNA by semi-partial correlation, which can measure the association between two variables after eliminating the effects of all other variables, i.e., the relationship between two genes is direct rather than influenced by other genes. The part mutual information (PMI) method (Zhang et al., 2015; Zhao et al., 2016) identifies direct associations based on the partial independence concept, and the partial information decomposition and context (PIDC) approach (Chan et al., 2017) uses partial information decomposition to determines the relationships between genes. These methods apply the concept of information theory and can measure nonlinear relationships. Usually, correlation networks are undirected, which means that the regulatory direction between two genes is unknown. However, the lag-based expression association for pseudotime-series (LEAP) approach (Specht & Li, 2017) can construct directed gene co-expression networks from pseudotime-ordered scRNA-seq data. This method computes the Pearson correlation coefficient over all possible time lags along the estimated pseudotime (no branch), and then uses the maximum correlation to construct the network. LEAP can capture the associations hidden by time lags and provides more accurate GRNs. Single-cell regularized inference using time-stamped expression profiles (SINCERITIES) (Gao et al., 2018) is another approach applied to construct directed GRNs from time-stamped single-cell transcriptional expression profiles. This method divides the single-cell data into several time points, uses Granger causality to infer regulatory networks centered by transcription factors (TFs), and uses ridge regression and partial correlation analyses to recover the directed regulatory relationships among genes. More accurate GRNs can be provided by this method. Several methods use statistical likelihood or Bayesian networks to infer GRNs, e.g., context likelihood of relatedness (CLR) (Faith et al., 2007) and first-order autoregressive moving-average and variational Bayesian expectation-maximization (AR1MA1-VBEM) (Sanchez-Castillo et al., 2018). These methods are similar to correlation networks but are considered more accurate for measuring gene-gene relationships based on their nonlinear principles.

Dynamic networks

In comparison to static (correlation) networks, dynamic networks are more suitable for describing changes in network dynamics, such as cellular lineages during differentiation. The Boolean model is one of the simplest methods, which takes the value of 0 or 1 to represent the absence or presence of gene expression and uses the Boolean operators AND, OR, and NOT to describe the interaction between two genes. The Boolean model is more robust to the effects of dropout, which makes it quite useful for scRNA-seq data. Many methods have reconstructed GRNs based on synchronous or asynchronous Boolean models, such as reduced ordered binary decision diagrams (ROBDD) (Garg et al., 2008), cellular network optimizer (CellNOptR) (Terfve et al., 2012), bool trainer (BTR) (Lim et al., 2016), single-cell network synthesis toolkit (SCNS) (Woodhouse et al., 2018), and gene modular network (GMN) (Zhang et al., 2020), which have been applied to find key regulators of cell fate and reveal network rewiring during cell differentiation (Moignard et al., 2015; Xu et al., 2014). However, the Boolean model must convert expression data into binary data, which may obscure gene-gene interactions. In contrast, differential equation-based models are more complex and offer high-precision predictions (Chen et al., 2009, 2010; Wang et al., 2006), but these methods must balance time complexity and prediction accuracy. Matsumoto et al. (2017) presented a highly efficient optimization algorithm (single-cell ordinary differentiation equations, SCODE) to reconstruct expression dynamics and infer GRNs from differentiating cells. This method integrates the transformation of linear ordinary differential equations (ODEs) and linear regression and can reconstruct the observed expression dynamics and GRNs accurately with remarkable efficiency.

Tree-based networks

Huynh-Thu et al. (2010) developed a tree-based algorithm (gene network inference with ensemble of trees, GENIE3), which adopted a distinctive way to infer regulatory networks. This method decomposes the prediction of GRNs into p regression models constructed by tree-based ensemble methods, e.g., Random Forests or Extra-Trees, where p is the number of genes. In each regression model, the expression pattern of gene x (target gene) is predicted from all other genes (input genes), and the weight of interaction between the target and input genes is determined by the importance of each input gene in the regression model. Several improvements to GENIE3 have been developed. For example, the GRN inference based on gradient boosting machine (GRNBoost2) method (Moerman et al., 2019) uses gradient boosting with GENIE3 architecture to improve algorithm efficiency. Jump3 (Huynh-Thu & Sanguinetti, 2015) combines the tree-based algorithm and dynamic systems to infer GRNs by exploiting the time series of expression data. The single-cell regulatory network inference and clustering (SCENIC) method (Aibar et al., 2017) removes the indirect targets from the GENIE3 modules based on TF motif enrichment analysis, and only retains those modules with enriched TF-binding motifs, called regulons. Generally, these methods are competitive with correlation models, and are able to construct directed networks (Chen & Mar, 2018; Pratapa et al., 2020).

Deep-learning-based networks

Deep-learning frameworks have also been used to infer gene relationships. The convolutional neural network for coexpression (CNNC) approach (Yuan & Bar-Joseph, 2019) is

a supervised and task-specific method, in which the network is trained by positive and negative samples, e.g. known targets of TFs, known pathways for specific biological processes, and known disease genes. Based on the data types used for training, CNNC can predict TF targets and identify disease-related genes.

Cell-specific networks

Recently, Dai et al. (2019) presented a new cell-specific network (CSN) method that can construct a network for each single cell from scRNA-seq data by considering statistical independence. Different from all other approaches, this method can identify gene-gene interactions and describe network heterogeneity at the single-cell level. CSN may help to find new cell types from a network perspective and reveal "dark" genes that play important roles in the network but are generally ignored by traditional differential analyses. Moreover, by considering partial independence, the conditional cell-specific network (CCSN) approach (Li et al., 2020) was developed to further reduce false positives in CSN method.

APPLICATION OF GENE REGULATORY NETWORKS

Method selection

The methods listed in this paper have their own advantages and disadvantages. How to choose the best method primarily depends on the scientific problem to be addressed. If the study focuses on time-series-related problems, such as development, cell differentiation, or disease procedures, the first choice would be an algorithm that constructs GRNs based on time-ordered data. If we only compare differences between two samples, e.g. the difference between a disease and normal state, and the difference before and after medication, the algorithm based on static data should be selected.

Directed networks provide information on the direction of a regulatory relationship, whereas undirected networks only measure the existence and strength of a regulatory relationship. Nonlinear algorithms can predict the strength of a regulatory relationship more accurately, but computational time will be longer; linear algorithms reduce the computational time, but accuracy also declines; binary algorithms can only show whether or not a regulatory relationship exists, but they are the fastest. Thus, if the purpose of a study is to explore the key regulatory factors controlling a biological process, research should focus on the changes or differences in network structure, instead of the strength of the regulatory relationship, and thus binary algorithms may be preferred. If we know certain regulatory factors are important and hope to identify their upstream and downstream genes, we can choose a directed network with linear or nonlinear algorithm. If we hope to simulate a biological process through a network, for example, by deleting network nodes to simulate gene knockout processes, then nonlinear algorithms are necessary. In addition, although the latest algorithms are often better than earlier ones, it is still important to build and compare networks constructed by different principles. Table 1 lists the type and principle of each method, and Table 2 lists the code and source of each method for reference.

In this paper, we selected several widely used algorithms to test whether they can identify proven gene regulatory relationships (Liu et al., 2020; Van Dijk et al., 2018). As shown in Table 3, most methods identified all six regulatory relationships, although two linear methods WGCNA and SINCERITIES did not perform well. This result is not

Table 1 Summary of inference methods for gene regulatory networks

Method	Type of edge		Input data	Principle	References
WGCNA	Linear	Undirected	Static	Pearson correlation	Langfelder & Horvath, 2008
PPCOR	Linear	Undirected	Static	Semi-partial correlation	Kim, 2015
PMI	Nonlinear	Undirected	Static	Part mutual information	Zhang et al., 2015; Zhao et al., 2016
PIDC	Nonlinear	Undirected	Static	Partial information decomposition	Chan et al., 2017
LEAP	Linear	Directed	Time-ordered	Pearson correlation	Specht & Li, 2017
SINCERITIES	Linear	Directed	Time-ordered	Ridge regression and partial correlation	Gao et al., 2018
AR1MA1 -VBEM	Nonlinear	Directed	Time-ordered	Bayesian framework	Sanchez-Castillo et al., 2018
ROBDD	Binary	Directed	Time-ordered	Boolean model	Garg et al., 2008
CellNOptR	Binary	Directed	Time-ordered	Boolean model	Terfve et al., 2012
BTR	Binary	Directed	Time-ordered	Boolean model	Lim et al., 2016
SCNS	Binary	Directed	Time-ordered	Boolean model	Woodhouse et al., 2018
SCODE	Nonlinear	Directed	Time-ordered	Ordinary differentiation equations	Matsumoto et al., 2017
GENIE3	Nonlinear	Directed	Static	Random Forests or Extra-Trees	Huynh-Thu et al., 2010
Jump3	Nonlinear	Directed	Time-ordered	Decision trees	Huynh-Thu & Sanguinetti, 2015
SCENIC	Nonlinear	Directed	Static	GENIE3 and TF motif enrichment analysis	Aibar et al., 2017
GRNBoost2	Nonlinear	Directed	Static	GENIE3 and gradient boosting	Moerman et al., 2019
CNNC	Nonlinear	Undirected	Static	Deep learning	Yuan & Bar-Joseph, 2019
CSN	Nonlinear	Undirected	Static / Time-ordered	Statistic independency	Dai et al., 2019
CCSN	Nonlinear	Undirected	Static / Time-ordered	Statistically partial independency	Li et al., 2020

unexpected as nonlinear algorithms usually predict gene regulatory relationships more accurately.

GRN analysis in biological studies

All methods listed in this paper use scRNA-seq data as input. Most GRN analyses need some prior information; algorithms based on time-ordered data need time-series information and algorithms based on static data need cell-type information. Both CSN and CCSN construct one network for each single cell, so they are suitable for either time-ordered data or static data. Some widely used data analysis software, e.g., Seurat (Butler et al., 2018) and Monocle (Qiu et al., 2017; Trapnell et al., 2014), can help obtain cell-type or time-series information based on clustering or pseudo-time analyses.

No matter which algorithm is used, network analysis is generally similar. For each network, the first step is to identify the modules it contains. Generally, a module represents a functional unit, as the genes performing the same function are often closely related to each other. In each module, the number of edges connected to a node, i.e., network degree, is an important indicator. If the network degree of a certain gene shows considerable differences between a disease and normal state, or shows significant changes during cell differentiation, this gene may be an important regulatory factor. If regulatory factors are known, the genes related to these factors should be considered. In addition, it should be noted that the genes linking different modules are often very important.

Gene regulatory networks have been widely used in biological studies. For example, based on correlation network analysis, Pina et al. (2015) identified a key regulatory gene (*Ddit3*) in erythroid lineage programming and found the *Ddit3*-*Gata2* regulatory axis could antagonize myeloid programs and enabled erythroid programs, which was validated experimentally. Xu et al. (2014) constructed Boolean networks composed of 30 genes related to the self-renewal and

Table 2 Sources of GRN inference methods

Method	Code	Source
WGCNA	R	R package: WGCNA
PPCOR	R	R package: ppcor
PMI	MATLAB	http://www.sysbio.ac.cn/cb/chenlab/software/PCA-PMI
PIDC	Julia	https://github.com/Tchanders
LEAP	R	R package: LEAP
SINCERITIES	R / MATLAB	http://www.cabsel.ethz.ch/tools/sincerities.html, https://github.com/CABSEL/SINCERITIES
AR1MA1- VBEM	MATLAB	https://github.com/mscastillo/GRNVBEM
ROBDD	Java	http://si2.epfl.ch/~garg/genysis.html
CellNOptR	R	http://www.bioconductor.org/packages/release/bioc/html/CellNOptR.html
BTR	R	R package: BTR
SCNS	R	https://github.com/swoodhouse/SCNS-GUI
SCODE	R	https://github.com/hmatsu1226/SCODE
GENIE3	R	http://www.montefiore.ulg.ac.be/~huynh-thu/software.html
Jump3	MATLAB	http://homepages.inf.ed.ac.uk/vhuynht/software.html
SCENIC	R	http://scenic.aertslab.org
GRNBoost2	Python	http://arboreto.readthedocs.io
CNNC	Python	https://github.com/xiaoyeye/CNNC
CSN	MATLAB	https://github.com/wys8c764/CSN
CCSN	MATLAB	http://sysbio.sibcb.ac.cn/cb/chenlab/soft/CCSN.zip

Table 3 Comparison of GRN inference methods

Method	Proven gene regulatory relationship					
	GSE114397			GSE139343		
	VIM- ZEB1	VIM- SNAI2	VIM- MYC	ARID1A-ZIC1	ARID1A-SOX1	ARID1A-MAP2
WGCNA	×	×	×	×	√	×
PPCOR	√	√	√	√	√	√
PMI	√	√	√	√	√	√
LEAP	√	√	√	√	√	√
SINCERITIES	√	×	×	×	√	√
SCODE	√	√	√	√	√	√
SCENIC	√	√	√	√	×	√
CSN	√	√	√	√	√	√

pluripotency of mouse embryonic stem cells (mESCs) cultured in serum/LIF or serum-free 2i/LIF conditions. They removed nodes from the Boolean network to simulate single and combinatorial RNA interference (RNAi) knockdown, with the predicted post-RNAi expression levels based on network analysis showing good agreement with experimental testing. In addition, Moignard et al. (2015) used diffusion maps to identify the developmental trajectory of the mesoderm toward blood in mouse based on scRNA-seq data, and then constructed Boolean networks to recapitulate blood development. The model predicted that the *Erg* gene is activated by *Sox17* or *Hoxb4*, which were validated by the observations that *Sox* and *Hox* factors control early expression of *Erg*. Harly et al. (2019) used LEAP to identify target genes of TCF-1 during innate lymphoid cell (ILC) development, and identified the role of TCF-1 in developmental progress of ILC precursors. Sagar et al. (2020) established a γδ T-cell differentiation map based on fetal and adult thymus scRNA-seq data using GENIE3 to construct GRNs and illustrate fetal and adult differences. Differentially expressed gene networks have also been successfully applied to recover and characterized distinct stages of γδ T-cell differentiation. Elyanow et al. (2020) presented a new computational method (netNMF-sc) using gene-gene co-expression networks as prior knowledge to perform dimensionality reduction and imputation of scRNA-seq data with high dropout rates, which was competitive with many other methods for dimension reduction and imputation.

FUTURE PERSPECTIVES

Although many GRN methods have been developed, GRN inference remains a challenging problem in bioinformatics and computational biology. A critical issue is the low quality of single-cell sequencing data. As RNA is obtained from only one cell, noise from amplification and dropout events in sequencing is a common problem. Recently, the integration of various single-cell-omics data, such as ATAC-seq and ChIP-seq, has attracted increasing attention (Li et al., 2017; Mimitou et al., 2019; Stuart et al., 2019), which may help in the development of next-generation GRN inference algorithms for various fields, including developmental and evolutionary biology.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

H.D. and L.N.C. conceived the review. H.D. prepared the draft. H.D., L.L., and Q.Q.J. collected the materials. All authors contributed to the discussions. All authors read and approved the final version of the manuscript.

REFERENCES

Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. 2017. SCENIC: single-cell regulatory network inference and clustering. *Nature Methods*, **14**(11): 1083−1086.

Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nature Biotechnology*, **36**(5): 411−420.

Chan TE, Stumpf MPH, Babtie AC. 2017. Gene regulatory network inference from single-cell data using multivariate information measures. *Cell Systems*, **5**(3): 251−267.e3.

Chen LN, Wang RQ, Li CG, Aihara K. 2010. Modeling Biomolecular Networks in Cells: Structures and Dynamics. London: Springer-Verlag.

Chen LN, Wang RS, Zhang XS. 2009. Biomolecular Networks: Methods and Applications in Systems Biology. New York: John Wiley & Sons.

Chen SN, Mar JC. 2018. Evaluating methods of inferring gene regulatory networks highlights their lack of performance for single cell gene expression data. *BMC Bioinformatics*, **19**(1): 232.

Dai H, Li L, Zeng T, Chen LN. 2019. Cell-specific network constructed by single-cell RNA sequencing data. *Nucleic Acids Research*, **47**(11): e62.

Eberwine J, Sul JY, Bartfai T, Kim J. 2014. The promise of single-cell sequencing. *Nature Methods*, **11**(1): 25−27.

Elyanow R, Dumitrascu B, Engelhardt BE, Raphael BJ. 2020. netNMF-sc: leveraging gene-gene interactions for imputation and dimensionality reduction in single-cell expression analysis. *Genome Research*, **30**(2): 195−204.

Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, et al. 2007. Large-scale mapping and validation of escherichia coli transcriptional regulation from a compendium of expression profiles. *PLoS Biology*, **5**(1): e8.

Gao NP, Ud-Dean SMM, Gandrillon O, Gunawan R. 2018. SINCERITIES: inferring gene regulatory networks from time-stamped single cell transcriptional expression profiles. *Bioinformatics*, **34**(2): 258−266.

Garg A, Di Cara A, Xenarios I, Mendoza L, De Micheli G. 2008. Synchronous versus asynchronous modeling of gene regulatory networks. *Bioinformatics*, **24**(17): 1917−1925.

Harly C, Kenney D, Ren G, Lai BB, Raabe T, Yang Q, et al. 2019. The transcription factor TCF-1 enforces commitment to the innate lymphoid cell lineage. *Nature Immunology*, **20**(9): 1150−1160.

Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. 2010. Inferring regulatory networks from expression data using tree-based methods. *PLoS One*, **5**(9): e12776.

Huynh-Thu VA, Sanguinetti G. 2015. Combining tree-based and dynamical systems for the inference of gene regulatory networks. *Bioinformatics*, **31**(10): 1614−1622.

Iacono G, Massoni-Badosa R, Heyn H. 2019. Single-cell transcriptomics unveils gene regulatory network plasticity. *Genome Biology*, **20**(1): 110.

Kim S. 2015. ppcor: an R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods*, **22**(6): 665−674.

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**(1): 559.

Li L, Dai H, Fang ZY, Chen LN. 2020. CCSN: single Cell RNA sequencing data analysis by conditional cell-specific network. *bioRxiv*, doi: 10.1101/2020.01.25.919829.

Li XY, Chen WZ, Chen Y, Zhang XG, Gu J, Zhang MQ. 2017. Network embedding-based representation learning for single cell RNA-seq data.

Nucleic Acids Research, **45**(19): e166.
Lim CY, Wang HG, Woodhouse S, Piterman N, Wernisch L, Fisher J, et al. 2016. BTR: training asynchronous Boolean models using single-cell expression data. *BMC Bioinformatics*, **17**(1): 355.
Liu JL, Liu S, Gao HY, Han L, Chu XN, Sheng Y, et al. 2020. Genome-wide studies reveal the essential and opposite roles of ARID1A in controlling human cardiogenesis and neurogenesis from pluripotent stem cells. *Genome Biology*, **21**(1): 169.
Matsumoto H, Kiryu H, Furusawa C, Ko MSH, Ko SBH, Gouda N, et al. 2017. SCODE: an efficient regulatory network inference algorithm from single-cell RNA-Seq during differentiation. *Bioinformatics*, **33**(15): 2314−2321.
Mimitou EP, Cheng A, Montalbano A, Hao S, Stoeckius M, Legut M, et al. 2019. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nature Methods*, **16**(5): 409−412.
Moerman T, Santos SA, González-Blas CB, Simm J, Moreau Y, Aerts J, et al. 2019. GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics*, **35**(12): 2159−2161.
Moignard V, Woodhouse S, Haghverdi L, Lilly AJ, Tanaka Y, Wilkinson AC, et al. 2015. Decoding the regulatory network of early blood development from single-cell gene expression measurements. *Nature Biotechnology*, **33**(3): 269−276.
Pina C, Teles J, Fugazza C, May G, Wang DP, Guo YP, et al. 2015. Single-cell network analysis identifies DDIT3 as a nodal lineage regulator in hematopoiesis. *Cell Reports*, **11**(10): 1503−1510.
Pratapa A, Jalihal AP, Law JN, Bharadwaj A, Murali TM. 2020. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. *Nature Methods*, **17**(2): 147−154.
Qiu XJ, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. 2017. Reversed graph embedding resolves complex single-cell trajectories. *Nature Methods*, **14**(10): 979−982.
Rozenblatt-Rosen O, Stubbington MJT, Regev A, Teichmann SA. 2017. The human cell atlas: from vision to reality. *Nature*, **550**(7677): 451−453.
Sagar, Pokrovskii M, Herman JS, Naik S, Sock E, Zeis P, et al. 2020. Deciphering the regulatory landscape of fetal and adult γδ T-cell development at single-cell resolution. *The EMBO Journal*, **39**(13): e104159.
Sanchez-Castillo M, Blanco D, Tienda-Luna IM, Carrion MC, Huang YF. 2018. A bayesian framework for the inference of gene regulatory networks from time and pseudo-time series data. *Bioinformatics*, **34**(6): 964−970.
Specht AT, Li J. 2017. LEAP: constructing gene co-expression networks for single-cell RNA-sequencing data using pseudotime ordering. *Bioinformatics*, **33**(5): 764−766.
Stegle O, Teichmann SA, Marioni JC. 2015. Computational and analytical challenges in single-cell transcriptomics. *Nature Reviews Genetics*, **16**(3): 133−145.
Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck III WM, et al. 2019. Comprehensive integration of single-cell data. *Cell*, **177**(7): 1888−1902.e21.
Terfve C, Cokelaer T, Henriques D, MacNamara A, Goncalves E, Morris MK, et al. 2012. CellNOptR: a flexible toolkit to train protein signaling networks to data using multiple logic formalisms. *BMC Systems Biology*, **6**(1): 133.
Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li SQ, Morse M, et al. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology*, **32**(4): 381−386.
Van Dijk D, Sharma R, Nainys J, Yim K, Kathail P, Carr AJ, et al. 2018. Recovering gene interactions from single-cell data using data diffusion. *Cell*, **174**(3): 716−729.e27.
Villani AC, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, et al. 2017. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science*, **356**(6335): eaah4573.
Wang Y, Joshi T, Zhang XS, Xu D, Chen L. 2006. Inferring gene regulatory networks from multiple microarray datasets. *Bioinformatics*, **22**(19): 2413−2420.
Woodhouse S, Piterman N, Wintersteiger CM, Göttgens B, Fisher J. 2018. SCNS: a graphical tool for reconstructing executable regulatory networks from single-cell genomic data. *BMC Systems Biology*, **12**(1): 59.
Xu HL, Ang YS, Sevilla A, Lemischka IR, Ma'ayan A. 2014. Construction and validation of a regulatory network for pluripotency and self-renewal of mouse embryonic stem cells. *PLoS Computational Biology*, **10**(8): e1003777.
Yuan Y, Bar-Joseph Z. 2019. Deep learning for inferring gene relationships from single-cell expression data. *Proceedings of the National Academy of Sciences of the United States of America*, **116**(52): 27151−27158.
Zeng T, Dai H. 2019. Single-cell RNA sequencing-based computational analysis to describe disease heterogeneity. *Frontiers in Genetics*, **10**: 629.
Zhang S, Zhao J, Lv XD, Fan JL, Lu Y, Zeng T, et al. 2020. Analysis on gene modular network reveals morphogen-directed development robustness in Drosophila. *Cell Discovery*, **6**(1): 43.
Zhang XJ, Zhao J, Hao JK, Zhao XM, Chen LN. 2015. Conditional mutual inclusive information enables accurate quantification of associations in gene regulatory networks. *Nucleic Acids Research*, **43**(5): e31.
Zhao J, Zhou YW, Zhang XJ, Chen LN. 2016. Part mutual information for quantifying direct associations in networks. *Proceedings of the National Academy of Sciences of the United States of America*, **113**(18): 5130−5135.