From time-series transcriptomics to gene regulatory networks: A review on inference methods

Malvina Marku1, Vera Pancaldi1,2*

1 CRCT, Université de Toulouse, Inserm, CNRS, Université Toulouse III-Paul Sabatier, Centre de Recherches en Cancérologie de Toulouse, Toulouse, France, 2 Barcelona Supercomputing Center, Barcelona, Spain

* malvina.marku@inserm.fr (MM); vera.pancaldi@inserm.fr (VP)

Abstract

Inference of gene regulatory networks has been an active area of research for around 20 years, leading to the development of sophisticated inference algorithms based on a variety of assumptions and approaches. With the ever increasing demand for more accurate and powerful models, the inference problem remains of broad scientific interest. The abstract representation of biological systems through gene regulatory networks represents a powerful method to study such systems, encoding different amounts and types of information. In this review, we summarize the different types of inference algorithms specifically based on time-series transcriptomics, giving an overview of the main applications of gene regulatory networks in computational biology. This review is intended to give an updated reference of regulatory networks inference tools to biologists and researchers new to the topic and guide them in selecting the appropriate inference method that best fits their questions, aims, and experimental data.

1. Introduction

In complex system theory, a system is defined as complex if certain properties, such as nonlinearity, feedback loops, adaptation, and nontrivial behavior, emerge from the collective interactions between the system components and the surrounding environment [1]. As such, all biological systems, especially molecular systems, are inherently complex, and the global structure and behavior of the system cannot be straightforwardly inferred from the (local) properties of its components. For example, characterizing the connection between genotype and phenotype and, furthermore, pathology, requires not only the identification of the molecules involved in the process and their specific characteristics but also a description of the ways in which these molecules interact with each other across spatiotemporal scales. To facilitate the representation and study of such complex systems, their interacting components can be represented as a network, commonly visualized as a graph of nodes (vertices) connected by edges (links) (Fig 1). In a molecular network, the nodes represent molecular objects of interest (e.g. genes, mRNAs, transcription factors (TFs), whereas the edges represent the interactions between them (e.g. protein binding, gene co-expression, TF–target regulation, etc.). Therefore,
depending on the types of nodes and edges considered, different molecular networks exist, like protein–protein networks (PPI) [2–5], gene regulatory networks (hereafter, GRNs) [6,7], signal transduction networks [8,9], etc. Strictly speaking, we define GRNs as networks including any type of regulatory interactions between regulatory and target molecular entities (miRNAs-targets, RBP-targets, kinases/phosphatases-substrates). Here we will focus mostly on GRNs that describe interactions between molecular components at the transcriptomics level, thus including TFs and their targets.

In studying GRNs, 2 main approaches for extracting information exist: (i) static network analysis, and (ii) dynamical modeling, each of which offers different amounts and types of information regarding the network organization, topology, and behavior [10,11]. Inferring and modeling these regulatory networks, however, is a challenging reverse engineering process and requires the combination of both a thorough biological understanding of the system, as well as accurate and advanced computational inference methods [12]. Moreover, the advanced technological improvements in measuring gene expression, in cellular populations or even single cells, and the increasing interest in clinical applications of genomics, confer importance and relevance to data-driven GRN inference methods. These approaches can ultimately provide considerably more insights on gene regulation mechanisms, drugs’ mode of action, pathway perturbation, etc. than the original data alone.

With the continuous advances in biotechnology, inference methods have also evolved from using bulk gene expression data [13–16] to single-cell transcriptomics [17–22]. Additionally, some methods were adapted to infer the regulatory networks from time-series and/or pseudo-temporal single-cell transcriptomics, in which case more accurate knowledge on gene–gene interaction can be inferred [20,23,24].

Consequently, highly performing algorithms have been developed over the years, with increasing accuracy, robustness, and applicability, and incorporating additional computational
analyses, such as filtering for the presence of TF motifs in promoters of their inferred targets [25,26]. It is important to note, however, that GRN inference is not the final aim. Our goal in performing such calculations is to produce insight on the biological processes considered, aiming to uncover new functionally important molecular regulatory interactions or propose new drug targets [6]. Additionally, network inference is a prerequisite to applying dynamical models to GRNs, in order to increase the models’ power to understand and predict the system’s temporal behavior [12].

Inference of data-driven accurate and powerful GRNs remains an open and evolving computational challenge, and new inference methods are being published continuously. In this review, we give a description of GRNs and describe the algorithms behind a selection of the available methods on network inference from experimental data. Several reviews on GRN inference based on different data sources have been published [27–30], many of them reviewing and performing benchmarking analysis on their performance based on simulated or real case datasets [31–34]. While almost all of them consider methods based on steady-state bulk or single-cell datasets, here, we will focus particularly on pseudo/time-series transcriptomics-based inference algorithms, including some methods exploiting the pseudotemporal formalism applied in single-cell datasets, but mostly focusing on inference based on bulk RNAseq time courses. We aim to give researchers a roadmap from the simplification of biological processes in the form of regulatory networks, to an algorithmic review of GRN inference methods based on time-series transcriptomics. Our goal is to guide researchers, but the choice of a method will ultimately depend on the scope and purpose of their research. Furthermore, we list some of the recent applications of GRN inference methods in various fields, particularly in cancer research, and give an introduction to dynamical modeling of GRNs, as a promising tool to exploit GRNs in the generation of new biological hypotheses. We conclude by discussing some pitfalls and existing challenges in data-driven GRN inference methods and speculate on possible directions for future developments in this field.

2. Gene maps: Network representation of gene regulation

Although a single definition does not exist, we define GRNs as topological maps representing relationships and interactions between biological entities. Interactions between proteins, TFs, and genes can be represented as directed graphs of nodes and edges. Notably, the directionality of the edges (defining the source and the target in the interaction) is not present in all biological networks, but it is an important feature in the case of the regulatory networks or metabolic networks, defining the direction of information flow. In these networks, additional information is added by indicating specific types of interactions, represented by signed edges (Fig 1B). In the simplest case, interactions are categorized as activations and inhibitions, represented as positive and negative edges accordingly. GRNs are composed of regulatory nodes (source/cause nodes) and regulated nodes (target/effect nodes), generally mapped as TF–target gene networks, with the incoming connectivity (in-degree) estimated using gene-centered approaches, and the outgoing connectivity (out-degree) using TF-centered approaches.

The structure of the network enables the calculation of various quantities that capture different features of the network topology that can reveal important information on the underlying biology of the system. Centrality measures, such as eigenvector centrality, Page/Chebi Rank, Burt’s constraint, or alpha centrality, have been shown to highlight key nodes in a network based on connectivity and influence [35–38]. Particularly of interest, in combination with mathematical methods such as linear approximation, Signal Flow Analysis, and Feedback Vertex Set Control [39], the network topology can give valuable information on the determinant nodes that define the system’s temporal behavior [40]. It is important to note that interpreting
the topological results in GRNs requires careful consideration of other implicit factors linked to the biological context of the system under study. For example, the role of a gene in a GRN and its regulatory interactions can depend on the cellular or environmental conditions in which it is expressed. These context-dependent factors are not captured by the centrality measures, resulting in incomplete descriptions of gene regulation. More on these open challenges will be discussed in Section 8.

3. Inference methods

By definition, the process of inferring the network structure of the system based on experimental data is a reverse engineering process, usually referred to as GRN inference. In general, depending on the input data used, GRN inference methods can be categorized into 2 types: (i) steady state gene expression, and (ii) time-series gene expression inference methods. In the first category, in principle, the GRN inference is obtained by considering perturbations or instances of the system and estimating the gene expression after it reaches an equilibrium. In the second category, the input data consist of gene expression measured at several time points after a perturbation, thus providing a temporal evolution of the transcriptome. Consequently, time series can be more informative than static data in a wide range of situations to infer gene functionalities, interactions, and causal relationships, as well as to establish potential clinical implications of gene expression dynamics determined by these relationships [41]. Both methods, however, have various limitations, mostly stemming from technical issues that arise from the experimental protocols: sampling time points, cost, cell synchronization, sparsity of gene expression data, etc. Therefore, several computational methods that combine both steady-state and time-series approaches have been developed, usually based on advanced machine learning algorithms. Additionally, new technologies providing expression at the single-cell level have led to the development of inference methods that are specifically adapted to single-cell transcriptomics [42].

Inferring the functional relationships between genes requires the estimation of gene functional dependencies, which can be broadly achieved by 2 different reverse engineering approaches for GRN inference:

1. **Model-free methods**: In this approach, gene dependencies are inferred using several statistical and machine learning methods, such as mutual information [43–45], random forest [46–49], or network deconvolution [50,51].

2. **Model-based methods**: In this approach, a quantitative dynamical model (for example, ODE [24,52,53]), regression method [54], or Bayesian reasoning [55] is defined to model the dynamical properties of the system, while the regulatory network is inferred by optimization of the model parameters based on the time-series data. In this way, model-based GRN methods can reveal some dynamical features of the system, increasing the model interpretability, especially from a biological point of view.

Depending on the scope and biological question for which the GRN is required, model-free or model-based inference methods can be applied. It is important to note that model-based methods are considerably more computationally intensive and might be limited by the underlying nonrealistic linear models of gene expression dynamics. On the other hand, the Dialogue on Reverse Engineering Assessment and Methods (DREAM) project, an initiative to benchmark multiple inference methods, indicated that no single method can perform the best in every possible dataset and across different settings. Instead, a high-confidence consensus network, inferred from different methods in a process referred to as Wisdom of Crowds [14], is the most accurate, while also providing an estimation of robustness and performance. For this
reason, instead of using a single inference method, other researchers have been developing computational tools combining several GRN inference methods [56], providing a ranking of the methods according to their performance.

The increasing demand for improving the inference accuracy in real datasets has warranted the adaptation of some steady-state inference methods to consider time-series data. This led to the development of new promising approaches, whose applications can generate novel results in various molecular systems. In the next sections, we will elaborate on this category of inference tools specifically and give an overview of several different algorithms and their performance.

3.1 Inference from time-series

Inference of gene networks from time-series data provides a more complete picture of the system than using just steady-state data, as the biological system is intrinsically complex and biological function relies on coordinated sets of genes whose expression evolves over time. In general, inference methods based on time-series transcriptomics use a dataset containing a list of genes with their expression measured at several time points. Let us define the dataset $D_{TS}$ as a matrix with dimensions $N \times P$, where $N$ is the number of genes and $P$ is the number of time points at which their expression is measured:

$$D_{TS} = X(t_1), X(t_2), \ldots, X(t_p)$$

where $X(t_p), p = 1, 2, \ldots, P$ is a vector of $N$ genes with their expression at time $t_p$. The main goal of inference methods is to assign a weight $w_{ij}, i, j = 1, 2, \ldots, N$ to any putative interaction between gene $i$ (target) and $j$ (source), representing a regulatory interaction in the biological system. To this end, different inference methods use various regression tools to model the expression of a gene as a function of its regulators. It is important to note that, usually, the time-series inference methods combine time-series with steady-state datasets, in order to increase the accuracy and predictive power of the inferred network. Independent of the method chosen, the goal is to reconstruct the GRN that would produce the observed profile of expression across time, in the form of a directed graph, in which each edge is associated with its characteristic weight (Fig 2).

3.2 Inference algorithms

Based on the inference model they use, inference methods can be grouped into 7 categories, namely, (i) mutual information (MI); (ii) dynamical Bayesian; (iii) Granger causality; (iv) Boolean; (v) ordinary differential equation (ODE); (vi) graphical Gaussian; and (vii) regression. It is important to note, however, that many methods apply a combination of different models and novel approaches. In the following sections, we will introduce the basic concept of each inference algorithm, while some of the available tools for using each of the methods are summarized in Table 1.

Mutual information–based models

In these methods, the putative interaction between 2 genes is inferred by estimating the mutual information (MI) between the expression $X_i$ of gene $i$ at time point $t_p$ and the the expressions $X_j$ of gene $j$ at the $m^{th}$ previous time points:

$$I_m(g_i; g_j) = \sum_{p-m \leq t} P(X_i(t_p), X_j(t_p - m)) \log \frac{P(X_i(t_p)X_j(t_p - m))}{P(X_i(t_p))P(X_j(t_p - m))}$$
Fig 2. From time-series transcriptomics to GRN inference and model validation. The process usually starts with processing the time-series datasets and identifying the time-points that are relevant to describe the biological process (1). As a second step, data formatting, normalization and gene filtering and/or binarization might be necessary (2), depending on the inference method to be used (3). The inferred network consists of a weighted directed GRN (4). Model validation (5) is then necessary to quantify the method performance in inferring the GRN when compared to gold-standard datasets, evidence from literature or prior biological knowledge. The inferred network can further be used for the application of dynamical models in cellular or multilevel modeling (6).

Table 1. GRN inference tools from time-series transcriptomics categorized by their inferring algorithm. The characteristics of the inferred network are indicated as follows: undirected; ▶ directed and unsigned; ▼ directed and signed.

| Tool          | Method                          | Language | Reference |
|---------------|---------------------------------|----------|-----------|
|▶ Time-delayed ARACNE | Mutual Information               | R        | [57]      |
|▸ MRNET        | Mutual Information              | minet R package | [58]      |
|▸ CLR          | Mutual Information              | minet R package | [43]      |
|▸ Jump3        | Regression                      | MATLAB   | [79]      |
|▸ dynGENIE3    | Regression (+ ODE)              | R        | [24]      |
|▸ SWING-RF     | Granger causality + Regression  | Python   | [65]      |
|▸ BETS         | Granger causality               | Python   | [64]      |
|▸ CGC-2SPR     | Granger causality               | R/MATLAB | [66]      |
|▸ scanBMA      | Dynamical Bayesian Model         | R        | [55]      |
|▸ Package GeneNet | Dynamical Bayesian Model        | R        | [81]      |
|▸ Package G1DBN | Dynamical Bayesian Model        | R        | [82]      |
|▸ PROB         | Dynamical Bayesian Model        | MATLAB   | [83]      |
|▸ MERLIN       | Dynamical Bayesian Model        | C        | [84]      |
|▸ REVEAL       | Boolean Network + MI            | C        | [70]      |
|▸ BoolNet      | Boolean Network                 | R        | [73]      |
|▸ ATEN         | Boolean Network + Regression    | R        | [74]      |
|▸ TSNI         | ODE                             | MATLAB   | [77]      |
|▸ Inferelator  | ODE                             | R        | [76]      |
|▸ DryNetMC     | ODE                             | R        | [85]      |
|▸ BINGO        | Gaussian process                | MATLAB   | [53]      |

https://doi.org/10.1371/journal.pcbi.1011254.t001
where $P(X_i(t_p), X_j(t_{p-m}))$ is the joint probability of observing gene $i$ at time point $t_p$ with expression $X_i(t_p)$ and the expression $X_j(t_{p-m})$ of gene $j$ at time point $t_{p-m}$, whereas $P(X_i(t_p))$ and $P(X_i(t_{p-m}))$ are the marginal probabilities of observing gene $i$ at time point $t_p$ with expression $X_i(t_p)$ and the expression $X_i(t_{p-m})$ of gene $j$ at time point $t_{p-m}$ independently. Estimating $I_m(g_i, g_j)$ for each pair of genes and each time point, an interaction $g_i \rightarrow g_j$ is being considered if $I_m(g_i, g_j)$ exceeds a defined threshold. Among the methods using MI algorithm for GRN inference, we can highlight TimeDelay-ARACNE [57], MRNET [45,58], and CLR [43]. One of the main advantages of these methods lies in their simplicity and speed of computation. On the other hand, their main limitation is that they do not give information on the nature of inferred interactions (activation or inhibition).

**Dynamical Bayesian models**

Dynamical Bayesian network (DBN) model GRN inference methods from time-series datasets \([59,60]\) come as an extension of simple Bayesian network (BN) models applied to steady-state datasets, in order to overcome the limitation of prohibiting the existence of feedback loops in the network. Given a GRN defined as a set of vertices and edges $G = (V,E)$, the method searches for any relationship between the expression of a target gene (defined as a random variable) and its regulators (defined as parent genes) by calculating a joint probability distribution. In a time-series experiment, consisting of $N$ genes and $P$ time points, the joint probability distribution of each gene is defined as:

$$P(X_1(t_1), X_1(t_2), \ldots, X_1(t_p), \ldots, X_N(t_1), X_N(t_2), \ldots, X_N(t_p)) = P(X(t_1))P(X(t_2)|X(t_1)) \cdots P(X(t_P)|P(X(t_{P-1}))) \quad (3)$$

where $X(t_p), p = 1, 2, \ldots, P$ is a vector representing the gene expression values at time point $t_p$. The network $G$ is then inferred by identifying the network structure with the highest posterior probability of each edge $E$ from the data, assuming a linear dependency between previous and current gene expression. The network $G$ is Markovian, meaning that each gene is dependent only on the regulation by its regulators (parents). An edge is then included in the network if the marginal posterior probability of an observation (left hand side of Eq 3 when written for each gene) exceeds a given threshold. Despite having a generally good performance when benchmarked against other time-series GRN inference methods (DREAM5 challenge \[14\]), DBNs are generally computationally expensive and usually limited to small size networks. To overcome this limit, several methods like CAS \[61\] or scanBMA \[55\], implemented as a function in networkBMA R package \[62\], have been developed, combining DBN with machine learning methods or prior knowledge to guide the search for the regulators of each gene.

**Granger causality models**

In these methods, 2 assumptions in defining the expression $X_i$ of a gene $i$ are made: (1) the expression $X_i$ of gene $i$ at time point $t_p$ is a function of its expressions at the $m$ previous time points; (2) the expression $X_i$ of gene $i$ at time point $t_p$ is a function of its expressions at the $m$ previous time points, combined with the expressions $X_j$ of gene $j$ at the $m$ previous time points, thus indicating the effect of gene $j$ on gene $i$. If the second assumption is significantly more successful than the first assumption, it is said that gene $j$ Granger causes gene $i$.

$$X_i(t_p) = \alpha_0 + \alpha_1 X_i(t_p - 1) + \ldots + \alpha_m X_i(t_p - m) + \eta(t_p) \quad (4A)$$

$$X_j(t_p) = \beta_0 + \beta_1 X_j(t_p - 1) + \ldots + \beta_m X_j(t_p - m) + \eta(t_p) \quad (4B)$$
where $\alpha, \beta$ are coefficients and $\eta(t_p)$ is the residual noise at time $t_p$. Generalizing for $N$ genes, gene $j$ ($j = 1, 2, \ldots, N, j \neq i$) is said to be causal for gene $i$ if considering the expression of gene $j$ in the previous time points significantly improves the prediction of the expression of gene $i$ at the current time point. Notably, when referring to the “previous time points,” several previous time points ($t_p-1, t_p-2$, etc.) can be considered. Considering $N$ genes in the dataset, the vector autoregressive (VAR) model is used to estimate the Granger causality over the entire list of genes, where a linear dependency between genes is being assumed [63]. Several methods using Granger causality for GRN network inference have been developed, such as BETS [64], SWING [65], CGC-2SPR [66], among others. One of the notable advantages of this approach is the computational efficiency and speed, compared to other inference methods, with a similar performance. However, due to the definition and underlying algorithm, Granger causality inference methods can be used only on time-series datasets, which often suffer from sparsity and nonuniformity of time-points. For this reason, a flexible time lag between consecutive time-points can be applied as suggested in [65].

Boolean models
In these methods, a GRN is represented as a graph $G = (V,E)$ and $<i,j,s>$ indexes, in which $i$ represents the target and $j$ represents the source gene. Each edge is characterized by a sign $s \in \{+, -\}$ indicating the type of interaction between the regulator and the source node. Generally, a positive sign indicates a positive (activating, promoting) relationship between a regulator and its target gene (i.e., the regulator gene $j$ contributes in increasing the expression of target gene $i$ through specific biological processes), whereas a negative sign indicates a negative (inhibiting, repression) relationship (i.e., the more expressed the regulator gene $j$ is, the less expressed the target gene $i$ will be). Notably, the expression of each gene is given in Boolean (binary) values $\{0,1\}$, thus providing only a qualitative description of gene regulation and expression. Given a time-series dataset, the expression $X_i$ of gene $i$ at time $t_p$ is given by a Boolean function of its regulators:

$$X_i(t_p) = F_B^i(X_{i1}(t_p - 1), X_{i2}(t_p - 1), \ldots, X_{ik}(t_p - 1))$$

Ordinary differential equation (ODE) models
In this formalism, the variation of expression $X_i$ of gene $i$ at time $t_p$ is given by a nonlinear function of its regulators (including self-regulation) at the same time steps:

$$\frac{dX_i}{dt} = \sum_{k=1,k \neq i}^{N-1} f(X_{ik}) + \eta_i$$
where the functions \( f(X_{i,k}) \) of \( k \) regulators of gene \( i \) with expression \( X_{i,k} \) are usually assumed as polynomial functions \([75]\), and \( \eta \) represents the nondeterministic term (noise) in the regulatory function of \( X_i \). In this way, the regulatory network is inferred by solving a system of ODEs (one equation for each gene). From its definition, solving the system of ODEs Eq 6 is computationally and conceptually challenging, even in a linear space, as these systems are defined by a large number of interacting coefficients, for which the prior information is limited or missing. Consequently, ODE methods are limited to network inference on a small number of genes and are usually combined with other inference methods in order to reduce the computational complexity \([44,45]\). In principle, a causal relationship between 2 genes is considered if the interaction is described by a relatively high interacting coefficient. Despite the difficulty of applying them to real datasets, several methods using ODE models have been developed, including Inferelator \([76]\) and TSNI \([77]\), but they often display a relatively lower performance than alternative approaches, as reported in \([64]\).

### Regression models

Generally, in these methods, a nonlinear regression or ODE formulation is used to model the expression of a gene at a certain time point/condition as a nonparametric function of the expression of other genes at the same time point/condition:

\[
X_i(t_p) = f(X_{j_6=0}(t_p)) + \eta(t_p)
\]

where \( \eta(t_p) \) is a random noise for the \( t_p \) time point. The inference of \( f_i \) functions is then performed by implementing a feature selection algorithm, like LASSO, least angle regression (LARS), or random forest algorithm, to “learn these functions from an ensemble of regression trees” \([78]\). Identifying the candidate genes on the regulation of \( N \) genes starts by splitting the problem into \( N \) distinct subproblems, thus assuming that each gene can be a target gene and a potential regulator. Causal interactions are quantified by estimating the confidence levels of each \( j \rightarrow i \) interaction, representing the weights \( w_{j,i} \). Importantly, in order to increase the accuracy of the inferred network, some tree-based methods have been adapted to use both time-series and steady-state datasets. Some inference methods based on decision-tree algorithms can be highlighted for having particularly good performance: dynGENIE3 \([24]\) and Jump3 \([79]\), both being extensions of GENIE3 \([46]\) on time-series datasets.

### Gaussian process (GP) models

In these methods, a GP regression is used to model the relationship between the current expression of the genes and their previous expression levels, with the functional relationships between targets and their regulators modeled as Gaussian processes:

\[
X_i(t_p) = f(X_i(t_p)) + \eta(t).
\]

A causal relationship (edge) between 2 genes is considered based on the sum of the posterior probabilities of the existence of this relationship (edge) over the time points. In this way, the Gaussian process model is generally considered as nonlinear DBN. Consequently, the developed inference methods, such as the algorithm presented by Aijo and Lähdesmäki \([80]\) and BINGO \([53]\), combine both DBN and GP.

### 4. Inference from time-series of single-cell transcriptomics

For some biological processes, like cell differentiation or phenotypic reprogramming, a higher resolution of the temporal dynamics of gene expression is necessary to identify major
phenotypic transitions in complex tissues while characterizing the phenotypic spectrum of individual cells. In this regard, single-cell RNA-sequencing technology enables deeper investigation of the molecular interactions and identification of novel molecular mechanisms that orchestrate biological processes at the single-cell level. Computationally, this technological revolution has led to the development of several algorithms to analyze single-cell RNA-seq and—as a part of it—inference of GRNs (Fig 3). Intuitively, observing genes’ dynamics at single-cell resolution would lead to an increased accuracy in inferring the functional interactions between genes that define the biological process. However, due to the limitations in some of the most widely available single-cell technologies, the heterogeneity and sparsity of single-cell data lead to limitations and challenges for GRN inference methods and put reliability in question.

An important point in using single-cell RNA-seq data for GRN inference is often the lack of time-resolved expression measurements. Instead, many inference methods exploit the multiplicity of RNAseq profiles at one single time point across cells as a proxy for temporal evolution of the phenotype, as is the case for trajectory inference based on pseudo-time ordering of the cells [25], assuming ergodicity of the phenotypes. In this process, the pseudotemporal trajectory is generated by linearly ordering the single-cell profiles from a specific time point based on their transcriptional similarity, thus enabling the identification of gene patterns along the developmental trajectory of continuously ordered cells [86]. Subsequently, to extract discrete time points from pseudotemporal ordering, different techniques can be employed: cell cluster time-point assignment (Slingshot [87], TSCAN [88], or Palantir [89]), partition of pseudotemporal trajectories into discrete time points, or differential gene expression time-point assignment (Monocle [90]). It is important to note, however, that the temporal representation obtained from pseudotemporal analysis in single-cell lacks the equivalence between pseudo-time and real time. In addition, this pseudotemporal representation is better suited in developmental systems, in which cells undergo differentiation processes recognized by the presence of bifurcation points in the pseudotemporal trajectory. This is not always the case, in which occasion trajectory inference may lead to inaccurate results. Nonetheless, this representation helps implement the GRN inference algorithms in modeling the expression levels at a given (pseudo)time-point as a function of gene expression at the previous (pseudo)time-point(s). Accordingly, a subset of the inference methods require specific information about the pseudotemporal ordering of the cells (CellOracle [91], SCODE [20], SINCERITIES [92], SINGE [93], LEAP [94], Scribe [95], etc.), having a significant difference in performance.

Fig 3. GRN inference from single-cell (pseudo)time-series. Starting from single-cell (pseudo) time-series (1), the inference...
when such information is not available. Other methods, like GENIE3 [46], GRNBoost2 [96], or PPCOR [97], do not require a temporal ordering of the cells as input and have relatively good performance when tested on some published curated models [98]. However, the incomplete equivalence between bulk gene expression time courses and pseudo-time time-series implies that these 2 types of inference cannot always be performed by the same tools.

Several benchmarking papers on the performance of single-cell RNA-seq inference methods have been published, facilitating their comparison. We refer the reader to [17,18,21] for an extensive review and comparison of single-cell RNA-seq inference methods, [19,98,99] for some benchmarking libraries, and to [29] for an algorithmic review. In principle, these methods follow a similar classification as in bulk transcriptomics, described in Section 3.2, as many of them were adapted for usage in single-cell from previous existing methods for bulk data. For this reason, here, we simply give a list of some available inference methods (Table 2). We note that there are many more published inference methods, and the list is constantly increasing.

For a user, the choice between all the different inference methods will depend on both their overall performance, the type and amount of information they require, and the type of reconstructed network they provide. For example, SCODE, PPCOR, and SINCERITIES infer a directed and signed GRN, which can further be used easily to make dynamical models, if necessary.

### 5. GRN validation

A crucial and essential step in performing GRN inference is validation of the inferred network. Having a validation protocol that allows evaluation of the score of each proposed model is very important in order to choose the optimal inference procedure among the variety of existing

---

**Table 2. GRN inference tools from single-cell transcriptomics categorized by their inferring algorithm.** The characteristics of the inferred network are indicated as follows: ▶ directed; ◆ directed and unsigned; △ directed and signed.

| Tool          | Method                                      | Language          | Reference |
|---------------|---------------------------------------------|-------------------|-----------|
| ◆◆ SINCERITIES| Correlation ensemble                         | R/MATLAB          | [92]      |
| ◆◆ SCIMITAR   | Correlation ensemble                         | Python            | [100]     |
| ◆◆ SCENIC     | ODE + Correlatorit                          | R                 | [25]      |
| ◆◆ GRNBoost2  | Regression                                  | Python            | [96]      |
| ◆◆ CellOracle | Regression                                  | Python            | [91]      |
| ◆◆ Pando      | Regression                                  | R                 | [101]     |
| ◆◆ WASABI     | Regression + wave propagation               | Python            | [102]     |
| CARDAMOM, HARISSA | Regression + wave propagation              | Python            | [103–105]|
|◆◆ PPCOR      | Semi-partial correlation                    | R                 | [97]      |
|◆◆ SINCERA    | Correlation                                 | R                 | [106]     |
| LEAP          | Correlation                                 | R                 | [94]      |
|◆◆ SINGE      | Granger causality                           | MATLAB            | [93]      |
|◆◆ ARIMA1—VBEM| Bayesian Dynamics                           | MATLAB            | [107]     |
|◆◆ SCODE      | ODE                                         | R/Julia/Ruby      | [20]      |
|◆◆ GRISLI     | ODE                                         | MATLAB            | [23]      |
|◆◆ SCOUP      | ODE                                         | C++               | [107]     |
|◆◆ SCRIBE     | Mutual Information                         | Python            | [95]      |
| PIDC          | Mutual Information                         | Julia             | [109]     |
|◆◆ Boolean Pseudo-time | Boolean Model                           | Python            | [110]     |
|◆◆ InferenceSnapshot | Boolean Model                      | C++/MATLAB        | [111]     |

https://doi.org/10.1371/journal.pcbi.1011254.t002
algorithms. Despite advances in this direction, evaluating the effectiveness of the inference method remains an open challenge, mostly due to limitations in ground truth/gold standard datasets.

A common approach to evaluate the accuracy of an inference method is to compare the inferred edges to TF–target interactions annotated in public databases, which can be of different kinds, depending on the type of information they contain [30,112]. Some databases will provide information on functional interactions, i.e., evidence that 2 genes/proteins, in this case a TF–target pair, can be related based on any type of link between them, evidenced by gene expression correlation, coevolution of the genes, comention in scientific abstracts, etc. These functional interactions can be found in databases such as String [113], Reactome (the functional interaction network [114], TRRUST [115], or RegNetwork [116], the latest selecting information from around 25 databases. Moreover, there are efforts to map more specific TF–target interactions based on experimental evidence, either collecting results of ChIP-seq experiments that show in which target gene promoters the binding peak of the TF can be found or also combining these experimental results with a crosscheck of the presence of the TF binding motif in the peak regions. Among the databases including these interactions, we mention ChEA [117] or ReMap [118]. Additionally, other databases, generally used when GRN inference is performed on single-cell RNAseq datasets, the validation of inferred interactions is performed on data on the same or similar cell types. Databases providing such information include ENCODE [119], ESCAPE [120], or ChIP-Atlas [121].

However, it is important to note that repositories that report experimental data contain reference data for a reduced number of interactions, limiting the validation to a subset of the inferred network. This implies that the interactions for which the reference data are absent are considered as not existing, raising important questions on the implication of literature biases or the potential to infer novel interactions and regulators. Focusing on prior knowledge for validation might reduce our chances of identifying new regulatory pathways involved in the system under consideration, for which gold standard references with high scores are absent or sparse.

Another possibility for network validation comes from using simulated data, which can be engineered to include several conditions and measurements [14]—yet the extent of coverage of the inferred interactions remains limited to small network size.

Other than validating the inferred network with a gold-standard one, another important point in GNR inference is network comparison between different inference methods, on the same dataset, thus providing an estimate of the “robustness of the inferred network.” Additionally, one may perform this benchmark analysis in order to choose the algorithm/method that best suits a given dataset.

Quantitatively, the algorithm’s performance can be evaluated as for any prediction by 2 metrics typically used in prediction tests: (i) area under the precision-recall curve for estimating the performance of a certain algorithm and (ii) area under the receiver operating characteristics curve for comparing the GRN inferred by the algorithm against a gold-standard network. Considering every inferred interaction in the GRN as a positive (true—TP, or false—FP) and any missing one as negative (true—TN, or false—FN), the 2 measures are defined as:

1. **area under precision-recall curve (AUPRC):** starting from the inferred weighted GRN, one might start by filtering by the highest values of the weights \( w_{ij} \in [0, 1] \), for which the precision (the fraction of TP over total predictions) will be high and recall (or sensitivity, the fraction of TP over total true elements) will be low. Adding interactions with lower weights will lead to a decrease of precision, until at recall = 1 (no filter, fully connected network), the precision will represent the fraction of true positives over the total number of interactions.
edges in the fully connected network. Graphically, this procedure is represented by a precision-recall curve: A good algorithm will have a high precision even when increasing the recall, and an AUPRC close to 1. Generally, this measure is used as a global performance estimator for most of the inference algorithms.

2. **area under the receiver operating characteristic curve (AUROC)**, defined as the ratio between the true positives and false positives, graphically represented with the TP on the y axis and the FP on the x axis. A good algorithm will have high TP while maintaining low FP values and an AUROC close to 1.

There are numerous other measures of performance that take into account class imbalance more explicitly (F1 value, Matthew’s correlation coefficient, etc.) [122]. In addition, the structural features of inferred networks, such as PageRank, heat diffusion, the shortest path, etc., can be used to compare GRNs inferred from different methods [123], whereas other metrics can be used to define the method performance, such as stability (across simulations, artificially removing measurements), identification of network motifs, and computational time and memory usage [98].

An important question arises when comparing the structure of the same cellular network of 2 different cellular states, e.g., different phenotypes of a certain cell or healthy versus disease conditions. This comparative analysis, referred to as differential network analysis (DiNA), remains a challenge, in parallel with the development of GRN inference methods, especially since different data types other than RNAseq can be included in GRN inference in order to increase the accuracy of the inferred network. For example, using single-cell ATAC-seq, patterns of gene expression can be combined with chromatin accessibility profiles, thus identifying cell subpopulations and groups of cells at different developmental stages [124]. Consequently, several differential network analysis algorithms have been developed, such as DiffRank [125], dcanr R package [126], DiffK [127], DINA [128], to name a few. We refer the reader to [129] for a statistical perspective of DiNA and [130] for a DiNA algorithm comparison. In principle, these algorithms consist of combining the information from 2 main computational tasks: (i) **differential expression analysis**, which estimates the gene expression abundance for each gene in the GRN between the 2 conditions, and (ii) **network expression analysis**, which estimates the importance of each gene in the GRN, based on the topological properties of the network. In this way, DiNA algorithms aim at identifying genes or subnetworks of genes whose expression changes the most across conditions, and they are especially suitable in cases when the changes in the network structure lead to phenotypic changes in the system. Numerous publications have demonstrated the power of this analysis, such as for identifying key TFs involved in cancer, when compared to healthy controls [124].

### 6. Applications

Although inferring regulatory networks is a trending topic and despite the multitude of different algorithms available, GRN inference methods struggle to reach high performance in real-world studies, on both bulk and single-cell RNA-seq data, as reported in [42,131]. Therefore, their application to biologically relevant datasets remains limited. Recent effort has been focused on combining information from multiple data sources to improve predictions [132,133], where gene expression data are combined with DNA methylation, copy number variation, genome-wide binding data, or CHIP-seq data. As a consequence, many inference methods introduced above combine multiple data sources to complement GRN inference and provide more insightful results on the regulatory processes. For example, in SCENIC [25], results of a GRN analysis are combined with TF binding motif information from RcisTarget to
identify a subset of high-confidence interactions. In CellOracle [91] and Pando [101], single-cell RNAseq from a time-series is combined with ATACseq data. These papers illustrate diverse strategies for integrating multiple data sources to enhance predictions and gain a deeper understanding of complex biological systems.

Different applications of inferring GRNs for novel discoveries in biology have been presented in most of the cited works on GRN inference methods, including studies on cancer, cell development, and cell fate decision. In cancer research, special attention has been given to identifying driver genes in cancer progression, and drug resistance. For example, in [85], the authors perform GRN inference from time-series RNAseq in gliomas to build sensitive and resistant networks, found to exhibit significant differences with respect to network topology, local entropy, and gene expression dynamics. Based on the node importance score, they developed a differential regulatory network–based biomarker model to identify the most influential genes for predicting and controlling drug resistance. Going a step further, [134] use time-series of single-cell RNAseq to study epithelial–mesenchymal transition, following a multilayer network approach and linking the intracellular gene regulation to cell–cell communications in ovarian cell lines. Other applications of GRN inference in cancer research focused on identifying putative key regulators and gene modules in PDAC disease progression [135], building cancer cell expression networks in liver hepatocellular carcinoma (LIHC) and bladder urothelial carcinoma (BLCA) [136], analyzing the functional components by extracting subnetworks and investigating the local landscape of prostate cancer genes [137], etc.

Other applications have been focusing on cell differentiation and development [25,103], by integrating different layers of information in GRN inference. For example, in [138], time-series of RNAseq data is combined with ATAC-seq data to derive dynamic gene regulatory networks for human myeloid differentiation, specifically promyelocytes differentiating into macrophages, neutrophils, monocytes, and monocyte-derived macrophages. In Bocci and colleagues [139], single-cell transcriptomics is reinforced with mRNA splicing analysis to identify the key molecular drivers leading to different final states when starting from a common initial state during pancreas endocrinogenesis and epithelial/mesenchymal state transition. In another application, Thorne [140] used public time-series datasets from recount2 database [141] to infer the GRN of neural progenitor cell differentiation. Performing structural analysis on the inferred GRN, they identified key genes like CDON and MCUR1, which were experimentally observed to influence neuronal differentiation. In another approach, Kamimoto and colleagues [91] used single-cell RNAseq and ATACseq datasets to infer cell state–specific GRNs that emerge during the differentiation process of fibroblasts. In this way, analyzing the changes in the GRNs during cell reprogramming or development can help understanding how the TF interactions regulate and define cell identity. In cell fate decision studies, Fleck and colleagues [101] used multiomics datasets, including RNAseq and ATACseq combined with transcription factor binding sites analysis, to infer a GRN describing brain organoid development, leading to the identification of the GLI3 transcription factor as a key regulator of cell fate. Similar studies are presented in [142–144].

In other efforts, data-driven GRNs have been used to build molecular disease maps [145,146], phenotypic characterization of a cell in a given microenvironment [147,148], identifying predictive or prognostic biomarkers [64], performing extensive studies on performance of the available methods on different datasets/conditions [22,124,149], and many more.

### 7. From static to dynamic networks

The inferred GRNs can serve as a backbone for constructing dynamical models: One can start from a temporal analysis of the system behavior under multiple conditions, describing how
the expression of genes in the network changes due to their regulatory interactions. Dynamical models of GRNs cover a spectrum as wide as the GRNs inference methods one. Many dynamical models exist, ranging from continuous quantitative ODE models to discrete logic quantitative models (Fig 4A). In continuous models, the temporal evolution of the state $X_i$ of gene $i$ ($i = 1, 2, \ldots, N$) in the network is given by a continuous ODE function of its regulators. In this way, the regulatory interactions between genes are modeled as chemical reactions, or species interactions in ecological models [150,151]. Beneficially, these models can capture the temporal evolution of the system at the level of individual reactions. However, their usage in dynamical modeling of GRN remains limited because of the detailed and complete mathematical and parametric description they require. On the other side of the spectrum, in discrete models, such as multivariate logic models [152], Petri nets [153], or Boolean models [67,154], a discrete logic of interactions is applied and the temporal evolution of expression of each gene is given by a discrete (and often logic) function of their regulators. Contrary to continuous models, discrete models can be applied with no or considerably fewer parameters and fragmented mechanistic description, making them suitable for large networks. However, the system behavior is only described (semi)quantitatively. Other methods use the network structure to infer the asymptotic behavior of the system, i.e., identifying only the steady states of the system, considering only the static GRN and a set of initial conditions, without requiring a dynamical description [39,40].

Fig 4. (A) Dynamical models of gene regulatory networks, represented as a spectrum of models ranging from continuous to discrete. (B) Steady states of a system composed of 3 genes: (left) ODE model, as a system of 3 interacting species; (right) Boolean model, as logic-based interacting entities. The attractors are identified as steady states in the long-term behavior of the system.

https://doi.org/10.1371/journal.pcbi.1011254.g004

For a modeler, the choice of the modeling method is going to depend on the type of questions being asked on the system, the type of available description of the system dynamics (quantitative or qualitative), the type and amount of available data, prior knowledge, etc. Whichever the dynamical model used, the main goal is to identify the phenotypic changes of a cell in response to certain extracellular environmental conditions, to drugs, to cellular interactions in the microenvironment, or in experimental knockout/overexpression experiments.
Graphically, these phenotypic changes can be identified as the system’s steady states or attractors (fixed points or limit cycles) (Fig 4B) [155], defined as a state (a vector of expression values of each gene), which remains unvaried even in the presence of perturbations. Usually, a GRN can have multiple attractors, representing the possible phenotypes or cell states that can be reached when starting from the given initial conditions. In this case, further analysis must be performed on the attractors, to study their stability, their biological relevance, or—additionally—their categorization into known phenotypes.

Particularly in Boolean models, similar to using time-series of expression data for GRN inference, the temporal information in the changes of the genes’ expression is also used to infer the Boolean rules that govern these changes [115,156–161], which can be then studied using several available tools [162]. More recently, patient-specific Boolean models have been developed to suggest targeted therapy to patients based on their omics profile [163]. Importantly, these intracellular models can be further integrated with other cell population models, like agent-based models or metabolic models, thus providing a multilevel description of the system dynamics, including processes like cell motility, cytokine diffusion, tissue expansion, spatial organization, etc. [164–167]. This combined approach enables addressing more complex tasks, such as drug design, or prediction of therapy response, from the cell to the tissue scale.

8. Discussion, open challenges

In this review, we give a basic introduction to GRNs, their topological characteristics and, most importantly, we describe the main GRN inference methods. Our aim is to give life scientists an overview of how to use the abstract concept of GRNs to investigate complex cellular molecular interactions more thoroughly and identify how specific interactions determine cells’ behavior and response to the environment. With the increasing abundance of transcriptomics data, data-derived GRNs have the potential to capture novel gene interactions and help to expand our knowledge of important molecular pathways. Despite the increasing interest in developing highly performing GRN inference methods, and this field of research having been active for more than 20 years, often the application of even the best performing methods in real-world studies raises questions about their reliability [131]. Purely data-driven GRN inference remains an open challenge, especially in single-cell RNAseq. Recent promising research addresses the task of network inference in the presence of incomplete multiomics datasets, resulting in the development of advanced computational methods [158]. We speculate that perhaps the information provided by transcriptomics data, even in single cell, is not sufficient to cover the complex cellular processes giving rise to phenotypes and that the inherited concept that coexpression patterns might reveal putative gene interactions might not be universally applicable. For example, due to posttranscriptional and posttranslation processes, a highly expressed mRNA may not lead to a highly abundant functional protein, and vice versa—a synthesized protein does not necessarily possess the necessary activation state or conformation or localization to affect its downstream targets. Therefore, there is a necessity to produce multiple types of omics data, consider different biological entities and levels of regulation, and the challenge in performing GNR inference will consist of integrating the data effectively.

Additional limitations in the performance of GRN inference methods can come from other—often ignored—sources. For example, gathering experimental samples from various patients can suppress the variability, since patients have individual histories, immune system, or genetic profiles. Experimental protocols can present restrictions as well, such as limited
measurements of expression (transcription, degradation, sequencing capture, etc.), heterogeneity in bulk datasets, or incompleteness in single-cell RNAseq transcriptomics.

Another challenge in data-driven GRN inference is result interpretability and the difficulty in dealing with the high complexity of the resulting networks increasing with their size. When considering cell state transitions (like differentiation or polarization) in a multicellular system, we can imagine that the interactions between TFs and their target genes can be cell type specific, potentially requiring different GRNs. However, all of the inference methods produce a single GRN, thus failing to provide the dynamics of the underlying mechanisms during cell state transition captured by the time-series. One possibility in dealing with this challenge is to infer stage-specific GRNs, thus having a time-evolving GRN that can help understanding how the involvement and interactions of specific genes/TFs lead to cell state transition. However, studying and understanding time-evolving GRNs remains an unexplored field of high complexity. Additionally, tracking specific regulatory pathways or interactions rapidly becomes a real challenge when dealing with large networks consisting of thousands of genes. We assume that topological network analysis might help in reducing the network size by selecting the most important nodes (by their centrality measures), although there is still a gap in our understanding between the structural and dynamical properties of the network [13,168–171]. Centrality measures can be sensitive to noisy or incomplete data that affect the structure of the inferred network. While such centrality measures are widely used to study the structural properties of many types of networks (social networks, communication networks, etc.), they might fail to capture the complex nonlinear interactions of gene regulation that are key in biological regulatory networks. We speculate that the existing challenges of data-driven inference methods discussed above are tightly intertwined with the limitations displayed by centrality measures derived from structural analysis of GRNs. An alternative could be the combination of multiple genes into functional units (gene modules, pathways) that could be nodes of a smaller GRN, thus reducing the dimensionality of the problem. This would still require quite arbitrary choices about how to group genes, but identifying functional modules could be achieved [172].

Using advanced computational methods, recent efforts have helped tackle the challenge of filling the gap between static and dynamic properties of GRNs. For example, in their work, Marazzi and colleagues [40] developed a method to combine several layers of information: gene expression data, the static representation of GRNs, the concept of system’s asymptotic behavior, and machine learning methods, to identify cell fate reprogramming targets. In another approach, Meena and colleagues [173] use the information on network topology to apply the linear approximation of dynamical Jacobian ensemble, thus qualitatively investigating and classifying the system’s asymptotic behavior around the fixed points.

One of the most difficult challenges in performing data-driven GRN inference is validating the results and estimating the method performance in real case studies. Benchmarking the inference methods with simulated datasets from a prior knowledge network (PKN) and a small list of genes can be an efficient way to estimate the methods’ performance. However, this is far from real biological systems, in which the actual regulatory mechanisms are mostly unknown and—in some cases—experimentally unexplored. Unfortunately, using public databases (usually built on bulk-RNAseq) for validating the inferred GRNs can be limited to small-size GRNs and narrowed around the highly studied regulatory pathways. Consequently, many approaches can lead to biased conclusions and a real difficulty in identifying novel regulatory pathways that might play an important role in the system under study.

We believe that many of these challenges will be addressed in integrative inference methods to be developed in the future. Despite the considerable improvements and the rapid growth in numbers of GRN inference methods, this remains a relatively new and highly complex field. Feeding the methods with different types of ‘omics data and prior knowledge, GRN inference
can help discover the importance of—yet uncharacterized—pathways, biological components, and interactions, thus further increasing our knowledge that will serve as a starting point for even more complete models, in a positive feedback loop.

Acknowledgments
The authors thank Hugo Chenel for exploring the computational tools for structural analysis on GRN, and Benoît Aliaga and Hugo Chenel for critical reading of the manuscript.

Author Contributions
Conceptualization: Malvina Marku.
Writing – original draft: Malvina Marku.
Writing – review & editing: Malvina Marku, Vera Pancaldi.

References
1. Artimo O, De Domenico M. From the origin of life to pandemics: emergent phenomena in complex systems. Philos Trans R Soc A Math Phys Eng Sci. 2022; 380(2227):20200410. https://doi.org/10.1098/rsta.2020.0410 PMID: 35999559
2. Koh GCKW, Porras P, Arandia B, Hermjakob H, Orchard SE. Analyzing Protein-Protein Interaction Networks. J Proteome Res. 2012; 11(4):2014–2031. https://doi.org/10.1021/pr201211w PMID: 22365417
3. Pellegrini M, Haynor D, Johnson JM. Protein Interaction Networks. Expert Rev Proteomics. 2004; 1(2):239–249. https://doi.org/10.1586/14789430.1.2.239 PMID: 1596618
4. Schwikowski B, Uetz P, Fields S. A network of protein-protein interactions in yeast. Nat Biotechnol. 2000; 18(12):1257–1261. https://doi.org/10.1038/82360 PMID: 11101803
5. Vazquez A, Flammini A, Maritan A, Vespignani A. Global protein function prediction from protein-protein interaction networks. Nat Biotechnol. 2003; 21(6):697–700. https://doi.org/10.1038/nbt825 PMID: 12740586
6. Emmert-Streib F, Dehmer M, Haibe-Kains B. Gene regulatory networks and their applications: understanding biological and medical problems in terms of networks. Front Cell Dev Biol. 2014; 2. https://doi.org/10.3389/fcell.2014.00038 PMID: 25364745
7. Vijesh N, Chakrabarti SK, Sreekumar J. Modeling of gene regulatory networks: A review. J Biomed Sci Eng. 2013; 06(02):223. https://doi.org/10.4236/jbise.2013.62A027
8. Hu Y, Peng T, Gao L, Tan K. CytoTalk: De novo construction of signal transduction networks using single-cell transcriptomic data. Sci Adv. 2021; 7(16):eabb1356. https://doi.org/10.1126/sciadv.abb1356 PMID: 33853780
9. Kolch W, Halas M, Granovskaya M, Kolodendro BN. The dynamic control of signal transduction networks in cancer cells. Nat Rev Cancer. 2015; 15(9):515–527. https://doi.org/10.1038/nrc3983 PMID: 26289315
10. He Q, Macauley M, Davies R. Chapter 5—Dynamics of Complex Boolean Networks: Canalization, Stability, and Criticality. In: Robeva RS, editor. Algebraic and Discrete Mathematical Methods for Modern Biology. Boston: Academic Press; 2015. p. 93–119. Available from: https://www.sciencedirect.com/science/article/pii/B9780128012130000058.
11. Jungck JR, Viswanathan R. Chapter 1—Graph Theory for Systems Biology: Interval Graphs, Motifs, and Pattern Recognition. In: Robeva RS, editor. Algebraic and Discrete Mathematical Methods for Modern Biology. Boston: Academic Press; 2015. p. 1–27. Available from: https://www.sciencedirect.com/science/article/pii/B9780128012130000010.
12. Angeline-Bonnet O, Biggs PJ, Vignes M. Gene Regulatory Networks: A Primer in Biological Processes and Statistical Modelling. Methods Mol Biol. 2019; 1883:347–383. https://doi.org/10.1007/978-1-4939-8882-2_15 PMID: 30547408
13. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet. 2011; 12(1):56–68. https://doi.org/10.1038/nrg2918 PMID: 21164525
14. Marbach D, Costello JC, Kuffner R, Vega NM, Prill RJ, Camacho DM, et al. Wisdom of crowds for robust gene network inference. Nat Methods. 2012; 9(8):796–804. https://doi.org/10.1038/nmeth.22796662
15. Silverman EK, Schmidt HHHW, Anastasiadou E, Altucci L, Angelini M, Badimon L, et al. Molecular networks in Network Medicine: Development and applications. Wiley Interdiscip Rev Syst Biol Med. 2020; 12(6):e1489. https://doi.org/10.1002/wsbm.1489 PMID: 32307915

16. Sonawane AR, Weiss ST, Glass K, Sharma A. Network Medicine in the Age of Biomedical Big Data. Front Genet. 2019; 10. https://doi.org/10.3389/fgene.2019.00294 PMID: 31031797

17. Blencowe M, Arneson D, Ding J, Chen YW, Saleem Z, Yang X. Network modeling of single-cell omics data: challenges, opportunities, and progresses. Emerg Top Life Sci. 2019; 3(4):379–398. https://doi.org/10.1042/ets20180176 PMID: 32270049

18. Fiers MWEJ, Minnoye L, Aibar S, Bravo Gonzalez-Blas C, Kalender Atak Z, Aerts S. Mapping gene regulatory networks from single-cell omics data. Brief Funct Genomics. 2018; 17(4):246–254. https://doi.org/10.1093/bfgp/elx046 PMID: 29342231

19. Kang Y, Thieffry D, Cantini L. Evaluating the Reproducibility of Single-Cell Gene Regulatory Network Inference Algorithms. Front Genet. 2021; 12. https://doi.org/10.3389/fgene.2021.617282 PMID: 33828580

20. 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 (15):2314–2321. https://doi.org/10.1093/bioinformatics/btx194 PMID: 28379368

21. Nguyen H, Tran D, Tran B, Pehlivan B, Nguyen T. A comprehensive survey of regulatory network inference methods using single cell RNA sequencing data. Brief Bioinform. 2021; 22(3):bbaa190. https://doi.org/10.1093/bib/bbaa190 PMID: 34020546

22. Raharirinina NA, Peppert F, von Kleist M, Schütte C, Sunkara V. Inferring gene regulatory networks from single-cell RNA-seq temporal snapshot data requires higher-order moments. Patterns. 2021; 2 (9):100332. https://doi.org/10.1016/j.patter.2021.100332 PMID: 34553172

23. Aubin-Frankowski PC, Vert JP. Gene regulation inference from single-cell RNA-seq data with linear differential equations and velocity inference. Bioinformatics. 2020; 36(18):4774–4780. https://doi.org/10.1093/bioinformatics/btaa576 PMID: 33026066

24. Huynh-Thu VA, Geurts P. dynGENIE3: dynamical GENIE3 for the inference of gene networks from time series expression data. Sci Rep. 2018; 8(1):3384. https://doi.org/10.1038/s41598-018-21715-0 PMID: 29467401

25. Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: Single-cell regulatory network inference and clustering. Nat Methods. 2017; 14(11):1083–1086. https://doi.org/10.1038/nmeth.4463 PMID: 28991892

26. Alvarez MJ, Shen Y, Giorgi FM, Lachmann A, Ding BB, Ye BH, et al. Network-based inference of protein activity helps functionalize the genetic landscape of cancer. Nat Genet. 2016; 48(8):838–847. https://doi.org/10.1038/ng.3593 PMID: 27322546

27. Androulakis IP, Yang E, Almon RR. Analysis of Time-Series Gene Expression Data: Methods, Challenges, and Opportunities. Annu Rev Biomed Eng. 2007; 9:205–228. https://doi.org/10.1146/annurev.bioeng.9.060906.151904 PMID: 17341157

28. Delgado FM, Gómez-Vela F. Computational methods for Gene Regulatory Networks reconstruction and analysis: A review. Artif Intell Med. 2019; 95:133–145. https://doi.org/10.1016/j.artmed.2018.10.006 PMID: 30420244

29. Huynh-Thu VA, Sanguinetti G. Gene Regulatory Network Inference: An Introductory Survey. In: Sanguinetti G, Huynh-Thu VA, editors. Gene Regulatory Networks. vol. 1883. New York, NY: Springer New York; 2019. p. 1–23. Available from: http://link.springer.com/10.1007/978-1-4939-8882-2_1.

30. Mercatelli D, Scalambra L, Triboli L, Ray F, Giorgi FM. Gene regulatory network inference resources: A practical overview. Biochim Biophys Acta Gene Regul Mech. 2020; 1863(6):194430. https://doi.org/10.1016/j.bbagrm.2019.194430 PMID: 31678629

31. Cassan O, Lèbre S, Martin A. Inferring and analyzing gene regulatory networks from multi-factorial expression data: a complete and interactive suite. BMC Genomics. 2021; 22(1):387. https://doi.org/10.1186/s12864-021-07659-2 PMID: 34039282

32. Maetschke SR, Madhamshettiwar PB, Davis MJ, Ragan MA. Supervised, semi-supervised and unsupervised inference of gene regulatory networks. Brief Bioinform. 2014; 15(2):195–211. https://doi.org/10.1093/bib/bbt034 PMID: 23698722

33. Saint-Antoine MM, Singh A. Network inference in systems biology: recent developments, challenges, and applications. Curr Opin Biotechnol. 2020; 63:89–98. https://doi.org/10.1016/j.copbio.2019.12.002 PMID: 31927423

34. Zhao M, He W, Tang J, Zou Q, Guo F. A comprehensive overview and critical evaluation of gene regulatory network inference technologies. Brief Bioinform. 2021; 22(5):bbab009. https://doi.org/10.1093/bib/bbab009 PMID: 33539514
35. Newman M, Barabási AL, Watts DJ. The Structure and Dynamics of Networks. Princeton University Press; 2006.
36. Cohen R, Havlin S. Complex Networks: Structure, Robustness and Function. Cambridge: Cambridge University Press; 2010. Available from: https://www.cambridge.org/core/books/complex-networks/97986C242A6F1D8668A5E0E9194B53EC.
37. Newman M. Networks: An Introduction. Oxford: Oxford University Press; 2010. Available from: https://oxford.universitypressscholarship.com/10.1093/acprof:oso/9780199206650.001.0001/acprof-9780199206650.
38. Ashiani M, Mirzaie M, Jafari M. CINNA: an R/CRAN package to decipher Central Informative Nodes in Network Analysis. Bioinformatics. 2019; 35(8):1436–1437. https://doi.org/10.1093/bioinformatics/bty819 PMID: 30239607
39. Lee D, Cho KH. Topological estimation of signal flow in complex signaling networks. Sci Rep. 2018; 8(1):5262. https://doi.org/10.1038/s41598-018-23643-5 PMID: 29588498
40. Marazzi L, Shah M, Balakrishnan S, Patil A, Vera-Licona P. NETISC: a network-based tool for cell fate reprogramming. NPJ Syst Biol Appl. 2022; 8(1):1–15. https://doi.org/10.1038/s41540-022-00231-y PMID: 35725577
41. Bar-Joseph Z, Gitter A, Simon I. Studying and modelling dynamic biological processes using time-series gene expression data. Nat Rev Genet. 2012; 13(8):552–564. https://doi.org/10.1038/nrg3244 PMID: 22805708
42. Chen S, Mar JC. Evaluating methods of inferring gene regulatory networks highlights their lack of performance for single cell gene expression data. BMC Bioinformatics. 2018; 19(1):232. https://doi.org/10.1186/s12859-018-2217-z PMID: 29914350
43. Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, et al. Large-Scale Mapping and Validation of Escherichia coli Transcriptional Regulation from a Compendium of Expression Profiles. PLoS Biol. 2007; 5(1):e8. https://doi.org/10.1371/journal.pbio.0050008 PMID: 17214507
44. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Favera RD, et al. ARACNE: An Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context. BMC Bioinformatics. 2006; 7(1):S7. https://doi.org/10.1186/1471-2105-7-S1-S7 PMID: 16723010
45. Meyer PE, Kontos K, Lafitte F, Bontempi G. Information-Theoretic Inference of Large Transcriptional Regulatory Networks. EURASIP J Bioinform Syst Biol. 2007; 2007:1–9. https://doi.org/10.1155/2007/18354736
46. Huynh-Thu VA, Irie HM, Wehenkel L, Geurts P. Inferring Regulatory Networks from Expression Data Using Tree-Based Methods. PLoS ONE. 2010; 5(9):e12776. https://doi.org/10.1371/journal.pone.0012776 PMID: 20927193
47. Kimura S, Fukutomi R, Tokuhisa M, Okada M. Inference of Genetic Networks From Time-Series and Static Gene Expression Data: Combining a Random-Forest-Based Inference Method With Feature Selection Methods. Front Genet. 2020; 11. https://doi.org/10.3389/fgene.2020.595912 PMID: 33384716
48. Park S, Kim JM, Shin W, Han SW, Jeon M, Jang HJ, et al. BTNET: boosted tree based gene regulatory network inference algorithm using time-course measurement data. BMC Syst Biol. 2018; 12(2):20. https://doi.org/10.1186/s12918-018-0547-0 PMID: 29560827
49. Petralia F, Wang P, Yang J, Tu Z. Integrative random forest for gene regulatory network inference. Bioinformatics. 2015; 31(12):i197–i205. https://doi.org/10.1093/bioinformatics/btv268 PMID: 26072483
50. Chen H, Munda PA, Zhao LN, Lin F, Zheng J. Highly sensitive inference of time-delayed gene regulation by network deconvolution. BMC Syst Biol. 2014; 8(4):S6. https://doi.org/10.1186/1752-0509-8-S4-S6 PMID: 25521243
51. Feizi S, Marbach D, MÅ©dard M, Kellis M. Network deconvolution as a general method to distinguish direct dependencies in networks. Nat Biotechnol. 2013; 31(8):726–733. https://doi.org/10.1038/nbt.2635 PMID: 23851448
52. Iglesias-Martinez LF, Kolch W, Santra T. BGRMI: A method for inferring gene regulatory networks from time-course gene expression data and its application in breast cancer research. Sci Rep. 2016; 6(1):37140. https://doi.org/10.1038/srep37140 PMID: 27876826
53. Aalto A, Vitasaaari L, Ilmonen P, Mombaerts L, Goncalves J. Gene regulatory network inference from sparsely sampled noisy data. Nat Commun. 2020; 11(1):3493. https://doi.org/10.1038/s41467-020-17217-1 PMID: 32661225
54. Michailidis G, d’Alché-Buc F. Autoregressive models for gene regulatory network inference: Sparsity, stability and causality issues. Math Biosci. 2013; 246(2):326–334. https://doi.org/10.1016/j.mbs.2013.10.003 PMID: 24176667
55. Young WC, Raftery AE, Yeung KY. Fast Bayesian inference for gene regulatory networks using ScanBMA. BMC Syst Biol. 2014; 8(1):47. https://doi.org/10.1186/1752-0509-8-47 PMID: 24742092
56. Manica M, Bunne C, Mathis R, Cadow J, Ahsen ME, Stolovitzky GA, et al. COSIFER: a Python package for the consensus inference of molecular interaction networks. Bioinformatics. 2021; 37 (14):2070–2072. https://doi.org/10.1093/bioinformatics/btaa942 PMID: 33241320
57. Zoppoli P, Morganella S, Ceccarelli M. TimeDelay-ARACNE: Reverse engineering of gene networks from time-course data by an information theoretic approach. BMC Bioinformatics. 2010; 11:154. https://doi.org/10.1186/1752-0509-11-154 PMID: 20338053
58. Liu W, Zhu W, Liao B, Chen H, Ren S, Cai L. Improving gene regulatory network structure using redundancy reduction in the MRNET algorithm. RSC Adv. 2017; 7(37):23222–23233. https://doi.org/10.1039/C7RA01557G
59. Friedman N, Koller D. Being Bayesian About Network Structure. A Bayesian Approach to Structure Discovery in Bayesian Networks. Machine Learning. 2003; 50(1):95–125. https://doi.org/10.1023/A:1020249912095
60. Murphy K, Mian S. Modelling gene expression data using dynamic Bayesian networks; 1999.
61. Xing L, Guo M, Liu X, Wang C, Wang L, Zhang Y. An improved Bayesian network method for reconstructing gene regulatory network based on candidate auto selection. BMC Genomics. 2017; 20(9):844. https://doi.org/10.1186/s12864-017-4228-y PMID: 29219084
62. Yeung KY, Fraley C, Raftery AE. Uncovering gene regulatory relationships using networkBMA. 2014.
63. Tam GHF, Chang C, Hung YS. Gene regulatory network discovery using pairwise Granger causality. IET Syst Biol. 2013; 7(5):195–204. https://doi.org/10.1049/iet-syb.2012.0063 PMID: 24067420
64. Lu J, Dumitrascu B, McDowell IC, Jo B, Barrera A, Hong LK, et al. Causal network inference from gene transcriptional time-series response to glucocorticoids. PLoS Comput Biol. 2021; 17(1): e1008223. https://doi.org/10.1371/journal.pcbi.1008223 PMID: 39513136
65. Finkle JD, Wu JJ, Bagheri N. Windowed Granger causal inference strategy improves discovery of gene regulatory networks. Proc Natl Acad Sci. 2018; 115(9):2252–2257. https://doi.org/10.1073/pnas.1710936115 PMID: 29440433
66. Yao S, Yoo S, Yu D. Prior knowledge driven Granger causality analysis on gene regulatory network discovery. BMC Bioinformatics. 2015; 16(1):273. https://doi.org/10.1186/s12859-015-0710-1 PMID: 26316173
67. Kauffman SA. The Origins of Order: Self-organization and Selection in Evolution. Oxford University Press; 1993.
68. Shmulevich I, Dougherty ER, Kim S, Zhang W. Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. Bioinformatics. 2002; 18(2):261–274. https://doi.org/10.1093/bioinformatics/18.2.261 PMID: 11847074
69. Pušnik Ž, Mraz M, Zimic N, Moškon M. Review and assessment of Boolean approaches for inference of gene regulatory networks. Helikon. 2022; 8(8):10222. https://doi.org/10.1016/j.helikon.2022.10222 PMID: 36033302
70. Liang S, Fuhrman S, Somogyi R. Reveal, a general reverse engineering algorithm for inference of genetic network architectures. Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing. 1998; p. 18–29. PMID: 9897168
77. Bansal M, Gatta GD, di Bernardo D. Inference of gene regulatory networks and compound mode of action from time course gene expression profiles. Bioinformatics. 2006; 22(7):815–822. https://doi.org/10.1093/bioinformatics/btl003 PMID: 16418235

78. Breiman L. Random Forests. Machine Learning. 2001; 45(1):5–32. https://doi.org/10.1023/A:1010933404324

79. Huynh-Thu VA, Sanguinetti G. Combining tree-based and dynamical systems for the inference of gene regulatory networks. Bioinformatics (Oxford, England). 2015; 31(10):1614–1622. https://doi.org/10.1093/bioinformatics/btu863 PMID: 25573916

80. Aijó T, Lähdesmäki H. Learning gene regulatory networks from gene expression measurements using non-parametric molecular kinetics. Bioinformatics (Oxford, England). 2009; 25(22):2937–2944. https://doi.org/10.1093/bioinformatics/btp511 PMID: 19706742

81. Schäfer J, Opgen-Rhein R, Strimmer K. Reverse engineering genetic networks using the GeneNet package. The Newsletter of the R Project Volume 6/5, December 2006. 2006; 6(9):50.

82. Lebre S, Lebre MS. Package 'G1DBN'. 2012.

83. Sun X, Zhang J, Nie Q. Inferring latent temporal progression and regulatory networks from cross-sectional transcriptomic data of cancer samples. PLoS Comput Biol. 2021; 17(3):e1008379. https://doi.org/10.1371/journal.pcbi.1008379 PMID: 33667222

84. Siahpirani AF, Roy S. A prior-based integrative framework for functional transcriptional regulatory network inference. Nucleic Acids Res. 2016;gkw963. https://doi.org/10.1093/nar/gkw963 PMID: 27794550

85. Zhang J, Zhu W, Wang Q, Gu J, Huang LF, Sun X. Differential regulatory network-based quantification and prioritization of key genes underlying cancer drug resistance based on time-course RNA-seq data. PLoS Comput Biol. 2019; 15(11):e1007435. https://doi.org/10.1371/journal.pcbi.1007435 PMID: 31682596

86. Zeng C, Mulas F, Sui Y, Guan T, Miller N, Tan Y, et al. Pseudo temporal Ordering of Single Cells Reveals Metabolic Control of Postnatal Cell Proliferation. Cell Metab. 2017; 25(5):1160–1175.e11. https://doi.org/10.1016/j.cmet.2017.04.014 PMID: 28467932

87. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudo-time inference for single-cell transcriptomics. BMC Genomics. 2018; 19(1):477. https://doi.org/10.1186/s12864-018-4772-0 PMID: 29914354

88. Ji Z, Ji H. TSCAN: Pseudo-time reconstruction and evaluation in single cell RNA-seq analysis. Nucleic Acids Res. 2016; 44(13):e117. https://doi.org/10.1093/nar/gkw963 PMID: 27179027

89. Setty M, Kiseliovas V, Levine J, Gayoso A, Mazutis L, Pe’er D. Characterization of cell fate probabilities in single-cell data with Palanit. Nat Biotechnol. 2019; 37(4):451–460. https://doi.org/10.1038/s41587-019-0068-4 PMID: 30899105

90. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudo-temporal ordering of single cells. Nat Biotechnol. 2014; 32(4):381–386. https://doi.org/10.1038/nbt.2859 PMID: 24658644

91. Kamimoto K, Stringa B, Hoffmann CM, Jindal K, Solnica-Krezel L, Morris SA. Dissecting cell identity via network inference and in silico gene perturbation. Nature. 2023; 614(7949):742–751. https://doi.org/10.1038/s41586-022-05688-9 PMID: 36755098

92. 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(2):258–266. https://doi.org/10.1093/bioinformatics/btx575 PMID: 28968704

93. Deshpande A, Chu LF, Stewart R, Gitter A. Network inference with Granger causality ensembles on single-cell transcriptomics. Cell Rep. 2022; 38(6):110333. https://doi.org/10.1016/j.celrep.2022.110333 PMID: 35139376

94. Specht AT, Li J. LEAP: constructing gene co-expression networks for single-cell RNA-seq data using pseudotime ordering. Bioinformatics. 2016;btw729. https://doi.org/10.1093/bioinformatics/btw729 PMID: 27993778

95. Qiu X, Rahimzamani A, Wang L, Ren B, Mao Q, Durham T, et al. Inferring Causal Gene Regulatory Networks from Coupled Single-Cell Expression Dynamics Using Scribe. Cell Systems. 2020; 10 (3):265–274.e11. https://doi.org/10.1016/j.cels.2020.02.003 PMID: 32135093

96. Moerman T, Aibar Santos S, Bravo González-Blas C, Simm J, Moreau Y, Aerts J, et al. GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. Bioinformatics. 2019; 35 (12):2159–2161. https://doi.org/10.1093/bioinformatics/bty916 PMID: 30445495

97. Kim S, ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. Commun Stat Appl Methods. 2015; 22(6):665–674. https://doi.org/10.5351/CSAM.2015.22.6.665 PMID: 26688802
98. Pratapa A, Jalihal AP, Law JN, Bharadwaj A, Murali TM. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. Nat Methods. 2020; 17(2):147–154. https://doi.org/10.1038/s41592-019-0690-6 PMID: 31907445

99. Bellot P, Olsen C, Salembier P, Oliveras-Vergés A, Meyer PE. NetBenchmark: a bioconductor package for reproducible benchmarks of gene regulatory network inference. BMC Bioinformatics. 2015; 16(1):312. https://doi.org/10.1186/s12859-015-0728-4 PMID: 26415849

100. Cordero P, Stuart JM. Tracing co-regulatory network dynamics in noisy, single-cell transcriptome trajectories. In: Pacific Symposium on Biocomputing 2017. World Scientific. 2017:576–587. https://doi.org/10.1142/9789813207813-0053 PMID: 27897008

101. Fleck JS, Jansen SMJ, Wollny D, Zenk F, Seimiya M, Jain A, et al. Inferring and perturbing cell fate regulomes in human brain organoids. Nature. 2022;1–8. https://doi.org/10.1038/s41586-022-05279-8 PMID: 36198796

102. Ventre E, Herbach U, Espinasse T, Benoit G, Gandrillon O. One model fits all: combining inference and simulation of gene regulatory networks. PLoS Comput Biol. 2023; 19(3):e1010962. https://doi.org/10.1371/journal.pcbi.1010962 PMID: 36972296

103. Bonnaffoux A, Herbach U, Richard A, Guillemín A, Gonin-Giraud S, Gros PA, et al. WASABI: a dynamic iterative framework for gene regulatory network inference. BMC Bioinformatics. 2019; 20(1):220. https://doi.org/10.1186/s12859-019-2798-1 PMID: 31046682

104. Herbach U, Bonnaffoux A, Espinasse T, Gandrillon O. Inferring gene regulatory networks from single-cell data: a mechanistic approach. BMC Syst Biol. 2017; 11(1):105. https://doi.org/10.1186/s12918-017-0487-0 PMID: 29157246

105. Ventre E. Reverse engineering of a mechanistic model of gene expression using metastability and temporal dynamics. In Silico Biology. 2021; 14(3-4):89–113. https://doi.org/10.3233/ISB-210226

106. Guo M, Wang H, Potter SS, Whitsett JA, Xu Y. SINCERA: A Pipeline for Single-Cell RNA-Seq Profiling Analysis. PLoS Comput Biol. 2015; 11(11):e1004575. https://doi.org/10.1371/journal.pcbi.1004575 PMID: 26600239

107. Sanchez-Castillo M, Blanco D, Tienda-Luna IM, Carrion MC, Huang Y. A Bayesian framework for the inference of gene regulatory networks from time and pseudo-time series data. Bioinformatics. 2018; 34(6):964–970. https://doi.org/10.1093/bioinformatics/btx605 PMID: 29028984

108. Matsumoto H, Kiyur H. SCOPU: a probabilistic model based on the Ornstein-Uhlenbeck process to analyze single-cell expression data during differentiation. BMC Bioinformatics. 2016; 17(1):232. https://doi.org/10.1186/s12859-016-1109-3 PMID: 27277014

109. Chan TE, Stumpf MPH, Babtie AC. Gene Regulatory Network Inference from Single-Cell Data Using Multivariate Information Measures. Cell Systems. 2017; 5(3):251–267.e3. https://doi.org/10.1016/j.cels.2017.08.014 PMID: 28957658

110. Harney FK, Nestorowa S, Kinston SJ, Kent DG, Wilson NK, Göttgens B. Reconstructing blood stem cell regulatory network models from single-cell molecular profiles. Proc Natl Acad Sci. 2017; 114(23):5822–5829. https://doi.org/10.1073/pnas.1610609114 PMID: 28584094

111. Ocone A, Haghverdi L, Mueller NS, Theis FJ. Reconstructing gene regulatory dynamics from high-dimensional single-cell snapshot data. Bioinformatics. 2015; 31(12):i89–i96. https://doi.org/10.1093/bioinformatics/btv257 PMID: 26072513

112. Garcia-Alonso L, Holland CH, Ibrahim MM, Turei D, Saez-Rodriguez J. Benchmark and integration of resources for the estimation of human transcription factor activities. Genome Res. 2019; 29(8):1363–1375. https://doi.org/10.1101/gr.240663.118 PMID: 31340985

113. van Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: a database of predicted functional associations between proteins. Nucleic Acids Res. 2003; 31(1):258–262. https://doi.org/10.1093/nar/gkg034 PMID: 12519996

114. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The Reactome Pathway Knowledgebase. Nucleic Acids Res. 2018; 46:D649–D655. https://doi.org/10.1093/nar/gkx1132 PMID: 29145629

115. Hall N. Data integration in logic-based models of biological mechanisms | Elsevier Enhanced Reader; 2021. Available from: https://www.sciencedirect.com/science/article/pii/S2452310021000809.

116. Liu ZP, Wu C, Miao H, Wu H. RegNetwork: an integrated database of transcriptional and post-transcriptional regulatory networks in human and mouse. Database. 2015; 2015:tvb095. https://doi.org/10.1093/database/tvb095 PMID: 26420482

117. Lachmann A, Xu H, Krishnan J, Berger SI, Mazloom AR, Ma’ayan A. ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. Bioinformatics. 2010; 26(19):2438–2444. https://doi.org/10.1093/bioinformatics/btq466 PMID: 20709693
118. Chénéby J, Gheorghe M, Artufel M, Mathelier A, Ballester B. ReMap 2018: an updated atlas of regulatory regions from an integrative analysis of DNA-binding ChIP-seq experiments. Nucleic Acids Res. 2018; 46:D267–D275. https://doi.org/10.1093/nar/gkx1092 PMID: 29126285

119. Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, et al. The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res. 2018; 46:D794–D801. https://doi.org/10.1093/nar/gkx1092 PMID: 29126249

120. Xu H, Baroukh C, Dannenfelser R, Chen EY, Tan CM, Kou Y, et al. ESCAPE: database for integrating high-content published data collected from human and mouse embryonic stem cells. Database. 2013; 2013:bat045. https://doi.org/10.1093/database/bat045 PMID: 30413482

121. Oki S, Ohta T, Shioi G, Hatanaka H, Ogasawara O, Okuda Y, et al. ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. EMBO Rep. 2018; 19(12):e46255. https://doi.org/10.15252/embr.201846255 PMID: 30413482

122. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics. 2020; 21(1):6. https://doi.org/10.1186/s12864-019-6413-7 PMID: 31898477

123. Arici MK, Tuncbag N. Performance Assessment of the Network Reconstruction Approaches on Various Interactomes. Front Mol Biosci. 2021; 8:666705. https://doi.org/10.3389/fmolb.2021.666705 PMID: 34676243

124. 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(1):4763. https://doi.org/10.1038/s41467-021-25089-2 PMID: 34362918

125. Odbat O, Reddy CK. Ranking differential genes in co-expression networks. In: Proceedings of the 2nd ACM Conference on Bioinformatics, Computational Biology and Biomedicine. BCB ‘11. Association for Computing Machinery; 2011. p. 350–354. Available from: https://doi.org/10.1145/2147805.2147849

126. Bhuvu DD, Cursons J, Smyth GK, Davis MJ. Differential co-expression-based detection of conditional relationships in transcriptional data: comparative analysis and application to breast cancer. Genome Biol. 2019; 20(1):236. https://doi.org/10.1186/s13059-019-1851-8 PMID: 31727119

127. Fuller TF, Ghazalpour A, Atten JE, Drake TA, Lusis AJ, Horvath S. Weighted gene coexpression network analysis strategies applied to mouse weight. Mamm Genome. 2007; 18(6):463–472. https://doi.org/10.1007/s00335-007-9043-3 PMID: 17668265

128. Gambardella G, Moretti MN, de Cegli R, Cardone L, Peron A, di Bernardo D. Differential network analysis for the identification of condition-specific pathway activity and regulation. Bioinformatics. 2013; 29(14):1776–1785. https://doi.org/10.1093/bioinformatics/btt290 PMID: 23749957

129. Shojaie A. Differential network analysis: A statistical perspective. WIREs Computational Statistics. 2021; 13(2). https://doi.org/10.1002/wics.1508 PMID: 37050915

130. Lichtblau Y, Zimmermann K, Haldemann B, Lenze D, Hummel M, Leser U. Comparative assessment of differential network analysis methods. Brief Bioinform. 2017; 18(5):837–850. https://doi.org/10.1093/bib/bbw061 PMID: 27473063

131. Kernfeld E, Keener R, Cahan P, Battle A. Model-X knockoffs reveal data-dependent limits on regulatory network identification; 2023. Available from: https://www.biorxiv.org/content/10.1101/2023.05.23.541948v1.

132. Liang X, Young WC, Hung LH, Raftery AE, Yeung KY. Integration of Multiple Data Sources for Gene Network Inference using Genetic Perturbation Data: Extended Abstract. In: Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics. BCB ’18. New York, NY, USA: Association for Computing Machinery; 2018. p. 601–602. Available from: https://doi.org/10.1145/3233547.3233692.

133. Yuan L, Guo LH, Yuan CA, Zhang Y, Han K, Nandi AK, et al. Integration of Multi-Omics Data for Gene Regulatory Network Inference and Application to Breast Cancer. IEEE/ACM Trans Comput Biol Bioinform. 2019; 16(3):782–791. https://doi.org/10.1109/TCBB.2018.2866836 PMID: 30130712

134. Sha Y, Wang S, Bocci F, Zhou P, Nie Q. Inference of Intercellular Communications and Multilayer Gene-Regulations of Epithelial-Mesenchymal Transition From Single-Cell Transcriptomic Data. Front Genet. 2021; 11. https://doi.org/10.3389/fgene.2020.604586 PMID: 33486773

135. Hossain SMM, Halsana AA, Khatun L, Ray S, Mukhopadhyay A. Discovering key transcriptomic regulators in pancreatic ductal adenocarcinoma using Dirichlet process Gaussian mixture model. Sci Rep. 2021; 11(1):7853. https://doi.org/10.1038/s41598-021-87254-7 PMID: 33846515

136. Zhang Y, Chang X, Liu X. Inference of gene regulatory networks using pseudo-time series data. Bioinformatics. 2021; 37(16):2423–2431. https://doi.org/10.1093/bioinformatics/btab099 PMID: 33576767
137. Moore D, de Matos Simoes R, Dehmer M, Emmert-Streib F. Prostate Cancer Gene Regulatory Network Inferred from RNA-Seq Data. Curr Genomics. 2019; 20(1):38–48. https://doi.org/10.2174/13892029196618107122005 PMID: 31015790

138. Ramirez RN, El-Al NC, Mager MA, Wyman D, Conesa A, Mortazavi A. Dynamic Gene Regulatory Networks of Human Myeloid Differentiation. Cell Systems. 2017; 4(4):416–429.e3. https://doi.org/10.1016/j.cels.2017.03.005 PMID: 28365152

139. Bocci F, Zhou P, Nie Q. spliceJAC: transition genes and state-specific gene regulation from single-cell transcriptome data. Mol Syst Biol. 2022; 18(11):e11176. https://doi.org/10.15252/msb.202211176 PMID: 36321549

140. Thorne T. Approximate inference of gene regulatory network models from RNA-Seq time series data. BMC Bioinformatics. 2018; 19(1):127. https://doi.org/10.1186/s12859-018-2125-2 PMID: 29642837

141. Collado-Torres L, Nellore A, Kammers K, Ellis SE, Taub MA, Hansen KD, et al. Reproducible RNA-seq analysis using recount2. Nat Biotechnol. 2017; 35(4):319–321. https://doi.org/10.1038/nbt.3838 PMID: 28398307

142. Dai H, Li L, Zeng T, Chen L. Cell-specific network constructed by single-cell RNA sequencing data. Nucleic Acids Res. 2019-06-20; 47(11):e62. https://doi.org/10.1093/nar/gkz172 PMID: 30864667

143. Skok Gibbs C, Jackson CA, Saldi GA, Tjarnberg A, Shah A, Watters A, et al. High-performance single-cell gene regulatory network inference at scale: the Inferelator 3.0. Bioinformatics. 2022; 38(9):2519–2528. https://doi.org/10.1093/bioinformatics/btac117 PMID: 35188184

144. Watters E, Boeckx B, Albar S, Nittner D, Burton O, et al. Phenotypic molding of stromal cells in the lung tumor microenvironment. Nat Med. 2018; 24(8):1277–1289. https://doi.org/10.1038/s41591-018-0096-5 PMID: 29988129

145. Glass L, Kauffman SA. The logical analysis of continuous, non-linear biochemical control networks. J Theor Biol. 1973; 39(1):103–129. https://doi.org/10.1016/0022-5193(73)90208-7 PMID: 4741704

146. Shah OS, Chaudhary MFA, Awon HA, Fatima F, Arshad Z, Amina B, et al. ATLANTIS—Attractor Landscape Analysis Toolbox for Cell Fate Discovery and Reprogramming. Sci Rep. 2018; 8(1):3554. https://doi.org/10.1038/s41598-018-22031-3 PMID: 29476134

147. Gjerga E, Trairatphisan P, Gabor A, Koch H, Chevalier C, Ceccarelli F, et al. Converting networks to predictive logic models from perturbation signalling data with CellInOpt. Bioinformatics. 2020; 36 (16):4523–4524. https://doi.org/10.1093/bioinformatics/bty584 PMID: 30423074
158. Henao JD, Lauber M, Azevedo M, Grekova A, List M, Ogris C, et al. Multi-Omics Regulatory Network Inference in the Presence of Missing Data. bioRxiv. 2022. Available from: http://biorxiv.org/lookup/doi/10.1101/2022.04.14.488153.

159. Ostrowski M, Paulevé L, Schaub T, Siegel A, Guziolowski C. Boolean network identification from perturbation time series data combining dynamics abstraction and logic programming. Biosystems. 2016; 149:139–153. https://doi.org/10.1016/j.biosystems.2016.07.009 PMID: 27484338

160. Razzaq M, Paulevé L, Siegel A, Saez-Rodriguez J, Bourdon J, Guziolowski C. Computational discovery of dynamic cell line specific Boolean networks from multiplex time-course data. PLoS Comput Biol. 2018; 14(10):e1006538. https://doi.org/10.1371/journal.pcbi.1006538 PMID: 30372442

161. Paulevé L, Kolčák J, Chatain T, Haar S. Reconciling qualitative, abstract, and scalable modeling of biological networks. Nat Commun. 2020; 11(1):4256. https://doi.org/10.1038/s41467-020-18112-5 PMID: 32848126

162. Naldi A, Hernandez C, Levy N, Stoll G, Monteiro PT, Chaouiya C, et al. The CoLoMobo Interactive Notebook: Accessible and Reproducible Computational Analyses for Qualitative Biological Networks. Front Physiol. 2018; 9. https://doi.org/10.3389/fphys.2018.00680 PMID: 29971009

163. Montagud A, Béal J, Tobalina L, Traynard P, Subramanian V, Szalai B, et al. Patient-specific Boolean models of signalling networks guide personalised treatments. Elife. 2022; 11:e72626. https://doi.org/10.7554/eLife.72626 PMID: 35164900

164. Aguilar B, Gibbs DL, Reiss DJ, McConnell M, Danziger SA, Dervan A, et al. A generalizable data-driven multicellular model of pancreatic ductal adenocarcinoma. GigaScience. 2020; 9(7):giaa075. https://doi.org/10.1093/gigascience/giaa075 PMID: 32696951

165. Letort G, Montagud A, Stoll G, Heiland R, Barillot E, Macklin P, et al. PhysiBoSS: a multi-scale agent-based modelling framework integrating physical dimension and cell signalling. Bioinformatics. 2019; 35(7):1188–1196. https://doi.org/10.1093/bioinformatics/bty766 PMID: 30169736

166. Stoll G, Naldi A, Noel V, Viara E, Barillot E, Kroemer G, et al. UPMaBoSS: A Novel Framework for Dynamic Cell Population Modeling. Front Mol Biosci. 2022; 9. https://doi.org/10.3389/fmolb.2022.800152 PMID: 35309516

167. Wertheim KY, Punia BL, Fleur AL, Shah AR, Barberis M, Helikar T. A multi-approach and multi-scale platform to model CD4+ T cells responding to infections. PLoS Comput Biol. 2021; 17(8):e1009209. https://doi.org/10.1371/journal.pcbi.1009209 PMID: 34334169

168. Luscombe NM, Madan Babu M, Yu H, Snyder M, Teichmann SA, Gerstein M. Genomic analysis of regulatory network dynamics reveals large topological changes. Nature. 2004; 431(7006):308–312. https://doi.org/10.1038/nature02782 PMID: 15372033

169. Klamt S, Saez-Rodriguez J, Lindquist JA, Simeoni L, Gilles ED. A methodology for the structural and functional analysis of signaling and regulatory networks. BMC Bioinformatics. 2006; 7:56. https://doi.org/10.1186/1471-2105-7-56 PMID: 16464248

170. Yang G, Gómez Tejeda Zarriñu J, Albert R. Target Control in Logical Models Using the Domain of Influence of Nodes. Front Physiol. 2018; 9. https://doi.org/10.3389/fphys.2018.00454 PMID: 29867523

171. Radde N, Bar NS, Banaji M. Graphical methods for analysing feedback in biological networks—A survey. Int J Syst Sci. 2010; 41(1):35–46. https://doi.org/10.1080/00207720903151326

172. Singh AJ, Ramsey SA, Fitz TM, Kioussi C. Differential gene regulatory networks in development and disease. Cell Mol Life Sci. 2018; 75(6):1013–1025. https://doi.org/10.1007/s00018-017-2679-6 PMID: 29018888

173. Meena C, Hens C, Acharya S, Haber S, Boccaletti S, Barzel B. Emergent stability in complex network dynamics. Nat Phys. 2023;1–10. https://doi.org/10.1038/s41567-023-02020-8