

<https://doi.org/10.1038/s41540-024-00372-2>

# Decoding the principle of cell-fate determination for its reverse control

Check for updates

Jonghoon Lee<sup>1,3</sup>, Namhee Kim<sup>1,2,3</sup> & Kwang-Hyun Cho<sup>1</sup>

Understanding and manipulating cell fate determination is pivotal in biology. Cell fate is determined by intricate and nonlinear interactions among molecules, making mathematical model-based quantitative analysis indispensable for its elucidation. Nevertheless, obtaining the essential dynamic experimental data for model development has been a significant obstacle. However, recent advancements in large-scale omics data technology are providing the necessary foundation for developing such models. Based on accumulated experimental evidence, we can postulate that cell fate is governed by a limited number of core regulatory circuits. Following this concept, we present a conceptual control framework that leverages single-cell RNA-seq data for dynamic molecular regulatory network modeling, aiming to identify and manipulate core regulatory circuits and their master regulators to drive desired cellular state transitions. We illustrate the proposed framework by applying it to the reversion of lung cancer cell states, although it is more broadly applicable to understanding and controlling a wide range of cell-fate determination processes.

Cell-fate determination is an evolutionarily well-conserved process through which cells make critical decisions regarding their ultimate roles within a multicellular organism<sup>1</sup>. This foundational process underpins functions that are indispensable for all multicellular organisms, including the precise orchestration of normal developmental pathways, the maintenance of internal equilibrium (homeostasis), and the facilitation of adult tissue regeneration. Due to its vital importance, cell-fate determination has evolved to be highly resilient to various perturbations. Nevertheless, seminal experimental findings have revealed that the predetermined destiny of a cell can be dramatically reshaped by a few molecular modifications. For instance, the overexpression of Yamanaka factors OCT4, SOX2, KLF4, and c-MYC (OSKM)-can reprogram differentiated fibroblasts into induced pluripotent stem cells<sup>2</sup>. Moreover, ectopic activation of Yes-associated protein can transdifferentiate terminally differentiated hepatocytes into biliary epithelial-like cells<sup>3</sup>. In addition, adenomatous polyposis coli restoration is able to reverse colon carcinoma cells back to a functionally normal state despite the presence of potent oncogenic mutations<sup>4</sup>. These findings collectively highlight that predisposed cell fates, in principle, can be changed by manipulating a few specific molecules, called master regulators, while remaining highly robust against most other molecular perturbations. This perspective further raises the following challenges: how can we identify the master regulators, and through what molecular regulatory mechanisms do they induce cell-fate changes?

Cells are dynamic systems composed of intricate signaling pathways interconnected by various feedbacks and crosstalks, forming a complex network<sup>5,6</sup>. The regulatory relationships within this network are predominantly nonlinear, adding layers of complexity that defy simple, intuitive predictions about how altering a specific molecule might impact cellular functions<sup>7,8</sup>. The inherent complexity and nonlinearity present significant challenges in understanding the principles underlying cell-fate determination and its control. To comprehend the intricate networks of intracellular molecular regulations, a systems biology approach that integrates quantitative mathematical modeling with molecular experimentation is indispensable.

In this perspective, we argue that the integration of newly emerging dynamic information with mathematical models enables us to not only decode the fundamental principles embedded in the process of cell-fate determination, but also to exert control over this intricate process to the degree that was previously unachievable. In particular, we propose that although cell-fate determination accompanies genome-wide molecular state changes, it might be underpinned by only specific subnetworks within the network, which we refer to as ‘core regulatory circuits’. These circuits are instrumental in orchestrating the intricate sequence of events that determine cell fate. Based on this, we present a conceptual framework that employs single-cell RNA-sequencing (scRNA-seq) data to identify and manipulate core regulatory circuits determining cell fate. By applying this

<sup>1</sup>Laboratory for Systems Biology and Bio-inspired Engineering, Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea. <sup>2</sup>Present address: biorevert, Inc., Daejeon, Republic of Korea. <sup>3</sup>These authors contributed equally: Jonghoon Lee, Namhee Kim. e-mail: [ckh@kaist.ac.kr](mailto:ckh@kaist.ac.kr)

framework to scRNA-seq data of tumorigenesis, we propose a system-level approach to identify molecular candidates for cancer reversion and their molecular mechanisms through which deregulated gene regulatory dynamics are rewired, and cancer hallmarks can be compromised to reestablish normal phenotypes.

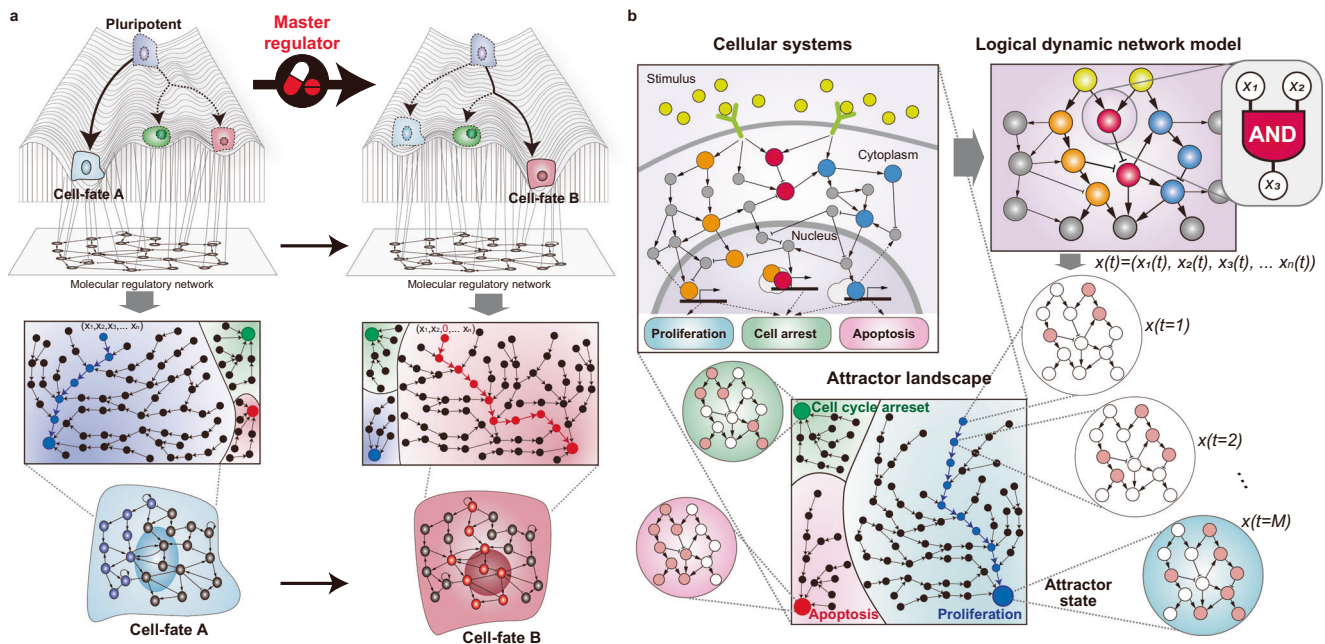
### Representing cell-fate changes by complex molecular network dynamics

Cellular functions are orchestrated by a gene regulatory network (GRN) composed of tens of thousands of genes, intertwined through intricate nonlinear interactions. Network dynamics can be conceptualized through Waddington’s landscape, an intuitive model that illustrates how cells navigate through various states within a high-dimensional state space<sup>9</sup> (Fig. 1a). In this landscape, valleys correspond to specific cell types, known as ‘attractors’, representing stable states that cells naturally settle into<sup>10</sup>. Building on this, the concept of ‘attractor landscape’ further illustrates the array of potentially stable states that cells can adopt. The basins surrounding these attractors indicate the probability of cells adopting each phenotype, offering insights into the dynamics and probabilities of cell-fate transitions. In this context, cell fates can be regarded as the most probable states a cell can occupy within the attractor landscape, reflecting their potential transition trajectories. In principle, the predetermined cell fates can be changed by altering the attractor landscapes of cells by rewiring the core regulatory circuits that underlie their complex molecular interactions<sup>11,12</sup> (Fig. 1a). Can Waddington’s landscape metaphor be used for a quantitative description of actual cell-fate determination? Moreover, can it be applied to cell-fate inference and control?

Such a metaphorical landscape can be quantitatively described through mathematical models (Fig. 1b). These models formalize the state of a cell at time  $t$  by the collective value of each molecular state in the network. For example, a network composed of  $n$  genes can be represented at time  $t$  by the state vector  $x(t) = (x_1(t), x_2(t), \dots, x_n(t))$ . In this vector,  $x_i(t)$  for  $i \in \{1, 2, \dots, n\}$  denotes the state of the  $i$ th gene at time  $t$ , and thus,  $x(t)$  corresponds to a point in an  $n$ -dimensional gene expression state space. The state of each molecule is influenced by the GRN, where the future state of each molecule is determined by nonlinear interactions with its upstream molecules. Consequently, a cellular state evolves over time as a nonlinear function of all molecules in the network,  $x(t + 1) = F(x_1(t), x_2(t), \dots, x_n(t))$ , converging towards a particular state (or a set of states) in the state space, which is called the attractor state. Empirical estimation of the network structure and nonlinear function parameters from experimental data is essential for a reasonably accurate model of cell-fate determination. However, obtaining the dynamic data needed for the estimation has been experimentally challenging. Fortunately, recent advances in single-cell omics and related analytical technologies can now provide the necessary temporal data required for the construction of comprehensive mathematical models.

### Methodologies available to construct mathematical models using single-cell omics data

Single-cell sequencing technologies are advancing rapidly and are now capable of analyzing datasets, including transcriptomes, epigenomes, and proteomes from hundreds of thousands of cells. In particular, RNA sequencing has emerged as the leading technique in single-cell omics<sup>13</sup>. Sequencing RNA at the single-cell level allows for a detailed examination of



**Fig. 1 | Cell-fate changes through the lens of the complex dynamics of molecular networks.** **a** Cell-fate changes in the epigenetic landscape. Waddington’s landscape is widely recognized as a conceptual framework to comprehend the transitions of pluripotent cellular states toward specific valleys that represent distinct cell fates. In the landscape metaphor, these cell fates are separated by an epigenetic barrier that restricts transitions, yet the perturbation of a specific master regulator can overcome this barrier to induce a dramatic transition between cell fates (top). Each cell fate can be regarded as the most stable state (also referred to an attractor) within the attractor landscape, which represents their transition trajectories (middle). Since each cell fate is characterized by the gene activities of the underlying molecular regulatory network, altering these complex molecular interactions by regulating a specific master regulator can reshape the attractor landscape, causing it to converge into a

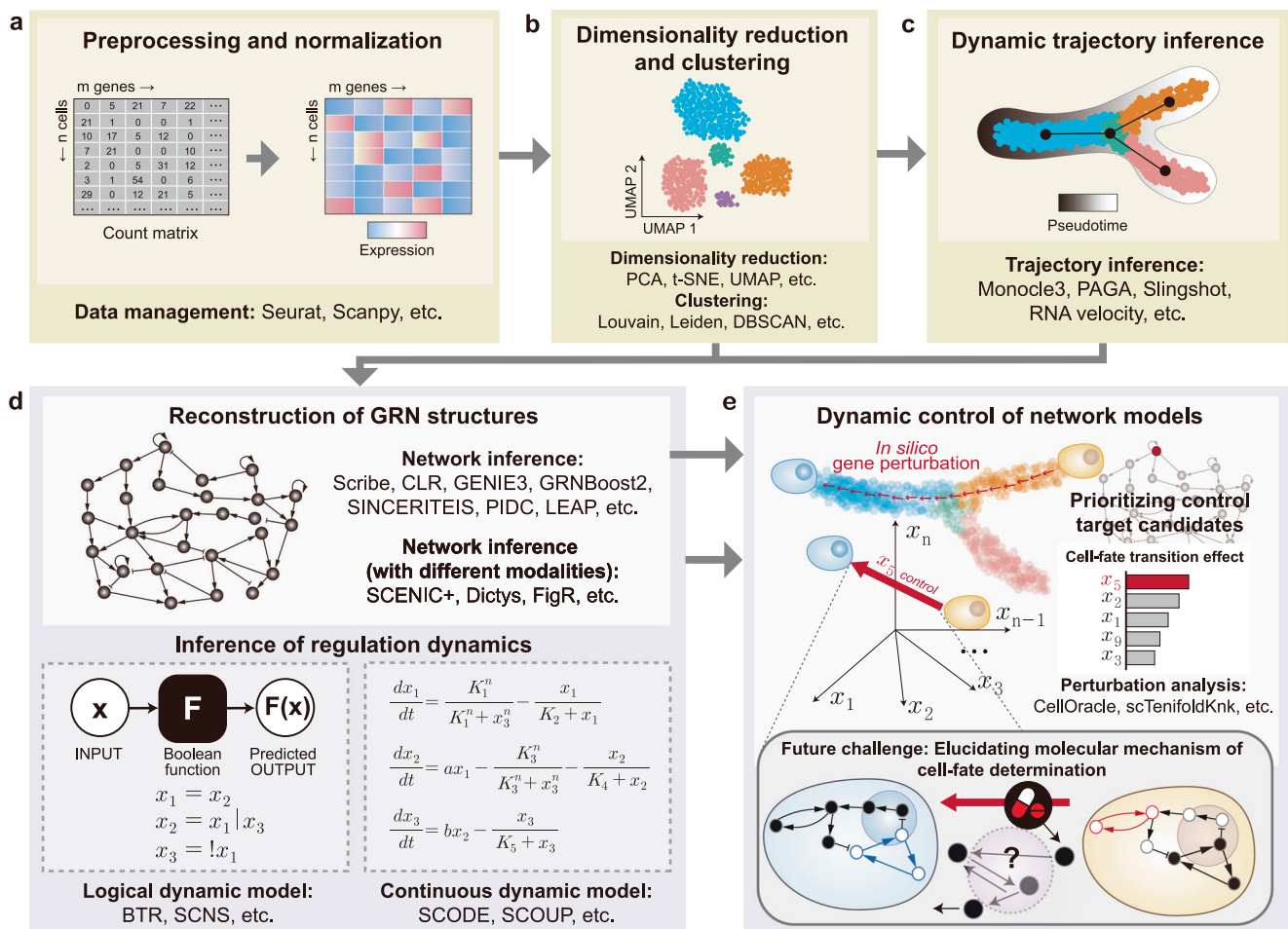
specific cell fate (bottom). **b** A mathematical model to unravel the hidden molecular regulation logic of cellular systems. A mathematical model can describe the dynamics of cellular systems in terms of gene regulatory interactions. The model can formalize cellular phenotypes such as proliferation, cell cycle arrest, and apoptosis, by assembling molecular components into the network. For example, a logical dynamic model is composed of  $n$  genes, whose state at time  $t$ , defined by  $x(t) = (x_1(t), \dots, x_n(t))$ , corresponds to a point in a  $n$ -dimensional gene expression state space represented by the attractor landscape. The state of each molecule is influenced by its dynamic regulations, where the future state of each molecule can be determined by a logical equation (e.g., AND gate) of its upstream molecules. A cellular state evolves through the logical equations of all molecules in the network, eventually converging on a specific attractor at time  $t = M$ .

gene transcription, providing a high-dimensional fingerprint that identifies unique cellular characteristics. Consequently, scRNA-seq has become an invaluable tool for investigating cell identity and state transitions at the level of individual cells. In this perspective, we focus on scRNA-seq. We illustrate current methodologies for utilizing scRNA-seq data in studying cell-fate determination, which encompass several critical stages as summarized below.

The initial stage involves preprocessing raw count data (Fig. 2a). This crucial first step, using tools such as Seurat<sup>14</sup>, Scanpy<sup>15</sup>, and Bioconductor-based SingleCellExperiment<sup>16</sup>, involves rigorous quality control, normalization, and feature selection, to establish a robust foundation for subsequent analyses. Following preprocessing, the complexity of data is dealt with through dimensionality reduction techniques like PCA<sup>17</sup>, t-SNE<sup>18</sup>, and UMAP<sup>19</sup>, effectively simplifying the data while preserving its essential characteristics (Fig. 2b). Following dimensionality reduction, clustering algorithms such as Louvain<sup>20</sup>, Leiden<sup>20</sup>, DBSCAN<sup>21</sup>, and SINCERA<sup>22</sup> are used to group cells with shared identities. This step is followed by the annotation of each clustered cell type using established biological knowledge, and the identification of specific cell states that correlate with the cell fates of interest.

To re-order cellular trajectories according to gene expression changes, pseudotime analysis tools like Monocle3<sup>23</sup> are employed to map cell trajectories from scRNA-seq snapshots (Fig. 2c). RNA velocity<sup>24,25</sup>, based on vectors derived from mRNA splicing dynamics, also indicates the direction and likelihood of cell state transitions. Additional methods like Slingshot<sup>26</sup>, PAGA<sup>27</sup>, and FateID<sup>28</sup> are also useful for inferring cellular dynamic trajectories and quantifying cell-fate probabilities.

The next phase involves constructing a molecular regulatory network (Fig. 2d, top), which infers molecular interactions from previously processed data<sup>29</sup>, often utilizing mutual information to understand nonlinear relationships within transcriptomic data. Tools such as Scribe<sup>30</sup> and CLR<sup>31</sup>, utilize mutual information to gauge nonlinear transcriptomic relationships, and are instrumental in building the molecular regulatory network. GENIE3<sup>32</sup>, GRNBoost2<sup>33</sup>, PIDC<sup>34</sup>, LEAP<sup>35</sup>, and SCENIC<sup>36</sup> are also widely used for this purpose. Furthermore, by integrating different single-cell data modalities (e.g., gene expression, chromatin accessibility), SCENIC+<sup>37</sup>, Pando<sup>38</sup>, Dictys<sup>39</sup>, and CellOracle<sup>40</sup> can estimate the regulatory effect of each transcription factor (TF) on each gene mediated by specific regions of DNA, and then infer the more specific GRN structure.



**Fig. 2 | A process for constructing mathematical models using scRNA-seq data.** **a** Utilization of scRNA-seq data in studying cell-fate determination. Initially, transcriptomic data is initially preprocessed to build a count matrix of cells by genes derived from raw data processing steps. Next, the count matrix is preprocessed using tools such as Seurat and Scanpy. **b** Identification of specific cell fates. Dimensionality reduction methods are applied to handle the complexity of the data, and clustering algorithms group cells with shared identities. Subsequently, specific cellular states related to the cell fates of interest are identified. **c** Inference of dynamic trajectories to reveal cell-fate changes. Cellular trajectories are derived from gene expression changes. In particular, Monocle3 and RNA velocity are generally used to infer

cellular dynamic trajectories from scRNA-seq datasets. **d** Inference of GRN structures and regulation dynamics. Molecular interactions are obtained from the processed data. Furthermore, a network model can be constructed by inferring the critical regulation dynamics involved in cell-fate determination by discrete or continuous mathematical formalisms. **e** Dynamic control of network models. In silico gene perturbation analysis with tools like CellOracle and scTenifoldKnk can predict the effects of control target candidates. However, they lack insight into regulatory dynamics represented by attractor landscapes. To tackle this limitation, the future direction of research should progress beyond applying control theory to elucidate the molecular mechanisms underlying cell-fate changes.

In addition to the constructed GRN structure, mathematical models are constructed by inferring and encapsulating critical regulatory dynamics from scRNA-seq data (Fig. 2d, bottom). To construct a logical dynamic model, methods such as BTR (BoolTraineR)<sup>41</sup> and SCNS (single cell network synthesis)<sup>42</sup> can be employed. Building a Boolean network model requires optimizing Boolean regulation logic. This process is tailored to each algorithm used, such as the Z3 solver and the Quine-McCluskey (QM) algorithm, and is performed on binarized data to ensure optimal precision. Parallel to this, continuous dynamic modeling adopts a different approach. This method often involves deriving ordinary differential equations (ODEs) or stochastic differential equations, with a particular emphasis on pseudo-time ordering of the data<sup>43–45</sup>. For example, SCODE<sup>44</sup> employs linear regression techniques, whereas SCOU<sup>45</sup> utilizes a continuous diffusion process and is designed to analyze single-cell expression data during differentiation processes. This latter model, based on the Ornstein-Uhlenbeck process, is especially adept at determining the interdependencies between gene expression levels at distinct temporal points. This method offers a more detailed understanding of cellular dynamics across various time points.

To summarize, we provide a brief overview of useful methods for constructing mathematical models (Table 1). These models are invaluable for identifying molecular targets, particularly master regulators, through extensive systematic perturbation analysis using tools like CellOracle<sup>40</sup> and scTenifoldKnk<sup>46</sup>, which predict gene perturbation effects and identify key cell-fate regulators (Fig. 2e). However, despite their advanced capabilities, these methods require laborious processes that involve individually

controlling molecules and analyzing complex systems. Furthermore, these tools do not inherently incorporate an understanding of system dynamics, such as the attractor landscape, into their analyses. This poses significant limitations on revealing the specific molecular regulatory mechanisms that dictate cell-fate decisions. To address these challenges, control theory may offer a new path forward.

### Emerging significance of applying control theory to explore core regulatory circuits and their master regulators

Control theory is a field of study that focuses on system characterization and manipulation. It involves understanding how systems behave and devising methods to achieve desired outcomes through control actions. This discipline has evolved over time to address vital challenges in complex systems. In the early 20th century, Black's development of the negative feedback amplifier laid the groundwork for feedback control of simple systems<sup>47</sup>. In the 21st century, Wolkenhauer, Kitano, and Cho introduced the interdisciplinary principles of systems biology, merging control theory with high-throughput technologies for cellular research<sup>48</sup>. Aligned with this evolution, in 2011, Barabási and Slotine made a significant advancement in complex network control by integrating network science with control theory<sup>49</sup>. Their works highlighted the future focus on addressing the intricate interplay of elements in a nonlinear, networked system.

Current developments in complex network control theories have evolved in two principal directions: one that focuses on the structural

**Table 1 | A brief overview of available methodologies to construct network models**

Methods	Possible input data types		Type of network interactions		Analysis of regulation dynamics	Default motif database	Implementation	URL
	scRNA-seq	scATAC-seq	Signed	Weighted				
GENIE3 <sup>32</sup> / GRNBoost2 <sup>33</sup>	O	X	X	O	X	X	Python and R	<a href="https://github.com/aertslab">https://github.com/aertslab</a>
SINCERITIES <sup>79</sup>	O	X	O	X	X	X	R and MATLAB	<a href="https://github.com/CABSEL/SINCERITIES">https://github.com/CABSEL/SINCERITIES</a>
PIDC <sup>34</sup>	O	X	X	X	X	X	Julia	<a href="https://github.com/Tchanders/NetworkInference.jl">https://github.com/Tchanders/NetworkInference.jl</a>
LEAP <sup>35</sup>	O	X	O	O	X	X	R	R package LEAP available on CRAN
SCENIC <sup>36</sup>	O	X	O	O	X	cisTarget	Python and R	<a href="https://scenic.aertslab.org/">https://scenic.aertslab.org/</a>
SCENIC+ <sup>37</sup>	O	O	O	O	X	cisTarget	Python and R	<a href="https://github.com/aertslab/scenicplus">https://github.com/aertslab/scenicplus</a>
scMTNI <sup>80</sup>	O	O	X	O	X	CIS-BP	C++	<a href="https://github.com/Roy-lab/scMTNI">https://github.com/Roy-lab/scMTNI</a>
Pando <sup>38</sup>	O	O	O	X	X	CIS-BP	R	<a href="https://github.com/quadbio/Pando">https://github.com/quadbio/Pando</a>
CellOracle <sup>40</sup>	O	O	O	O	X	CIS-BP	Python	<a href="https://github.com/morris-lab/CellOracle">https://github.com/morris-lab/CellOracle</a>
FigR <sup>81</sup>	O	O	O	X	X	CIS-BP	R	<a href="https://github.com/buenrostrolab/FigR">https://github.com/buenrostrolab/FigR</a>
Dictys <sup>39</sup>	O	O	O	O	X	HOCOMOCO	Python	<a href="https://github.com/pinellolab/dictys">https://github.com/pinellolab/dictys</a>
scTenifoldKnk <sup>46</sup>	O	X	O	O	X	X	R and MATLAB	<a href="https://github.com/cailab-tamu/scTenifoldKnk">https://github.com/cailab-tamu/scTenifoldKnk</a>
BTR <sup>41</sup>	O	X	O	X	O	X	R	<a href="https://github.com/cheeyeelim/btr">https://github.com/cheeyeelim/btr</a>
SCNS <sup>42</sup>	O	X	O	X	O	X	F# and Javascript	<a href="https://github.com/swoodhouse/SCNS-GUI">https://github.com/swoodhouse/SCNS-GUI</a>
SCODE <sup>44</sup>	O	X	O	O	O	X	R and Julia	<a href="https://github.com/hmatsu1226/SCODE">https://github.com/hmatsu1226/SCODE</a>
SCOU <sup>45</sup>	O	X	O	X	O	X	C++	<a href="https://github.com/hmatsu1226/SCOU">https://github.com/hmatsu1226/SCOU</a>



characteristics of networks and another that considers the inherent non-linear dynamics within these networks. Controlling network centrality determines the most influential regulators in network interactions using metrics such as degree centrality, betweenness centrality, and eigenvector centrality<sup>50</sup>. This approach often aims to identify a minimal subset of driver nodes required to direct a network from any given state to a specific desired state<sup>51</sup>. In contrast, strategies for controlling network dynamics, like logical domain of influence (LDOI)<sup>52</sup>, feedback vertex set (FVS)<sup>53</sup>, stable motif<sup>54</sup>, searching for differentially expressed positive circuits (DEPCs)<sup>55</sup>, and global stabilization analysis<sup>56</sup>, were suggested to manage state transitions caused by the network dynamics. These methods collectively aim not only to identify ‘control targets’, defined as a specific set of nodes capable of steering the system towards a set of desired states, but also offer tools for analyzing molecular regulatory mechanisms. For instance, stable motif analysis has been instrumental in uncovering crucial feedback loops that govern processes like the epithelial-to-mesenchymal transition<sup>57</sup>, leukemia cell-fate decisions<sup>54</sup>, and the differentiation of helper T cells<sup>54</sup>. In addition, the LDOI approach has identified key control targets and their influences, and the DEPC method has discovered related positive circuits to stabilize a certain attractor<sup>55</sup>. While these methods provide solutions for controlling theoretical model systems, some result in suboptimal solutions with unnecessary targets, and yet others face scalability issues, rendering them limited to very small-scale models. Furthermore, to our knowledge, there has been no attempt to systematically identify both control targets and the resulting fate-determining paths (or circuits) revealing specific molecular regulatory mechanisms. Thus, identifying optimal control targets (i.e., master regulators) and the resulting core regulatory circuits for real-world cell-fate control remains a significant challenge.

To address the aforementioned challenge, leveraging the key biological features that play a role in cell-fate determination could be an essential strategy. Many previous experimental studies have already shown that while cell systems are comprised of a complex large network, a core regulatory circuit, composed of only a few key molecules, plays a crucial role in determining cell fates<sup>58–61</sup>. In particular, such core regulatory circuits are often composed of multiple nested feedback loops, with at least one being a positive feedback loop. For example, in quorum sensing, molecules like N-acyl homoserine lactone (AHL) in bacteria facilitate population-wide communication<sup>62</sup>; the regulatory feedback between TFs OCT4, SOX2, and NANOG maintains pluripotency in embryonic stem cells<sup>63</sup>; the double negative feedback loop between PU.1 and GATA1 controls the erythroid versus myeloid lineage commitment<sup>64</sup>; and the interconnected SNAIL/miR-34 and ZEB/miR-200 feedback loops are key regulatory components of the epithelial–mesenchymal transition (EMT) process in cancer metastasis<sup>65</sup> (Fig. 3a). Furthermore, combinations of feedback and feedforward controls can generate biological transitions reliably in noisy environments where the activity of individual components can vary over a range of parameters<sup>66</sup>. Those previous studies indicate the significance of specific feedback mechanisms within core regulatory circuits in governing crucial cellular processes. This leads to the following question: how can we identify these core regulatory circuits and their master regulators?

To computationally resolve this problem, the concept of a network kernel has been introduced, encompassing the ‘kernel’ method for simplifying networks<sup>67</sup> and the ‘control kernel’ approach for altering attractor landscapes with minimal regulators<sup>68,69</sup>. These kernel-based approaches are crucial for focusing on and controlling the core regulatory circuits that govern vital biological phenomena. Similarly, since cellular circuits with positive feedback loops induce multistationary behavior, identifying and analyzing the properties of feedback loops can offer critical clues to determine core regulatory circuits within large networks<sup>69–71</sup>. Another algorithm, ‘OpitCon’, can identify combination targets using a subgraph based on structural controllability theory, and then describe specific core downstream subnetworks and their crosstalk links that contribute to therapy resistance<sup>72</sup>. More recently, Rukhlenko et al.<sup>73</sup> developed a novel framework called ‘cSTAR’. This method uses omics data to classify cell states and transforms them into mechanistic models. These models consist of a key molecular

regulatory network that is instrumental in controlling the attractor landscape that governs cell-fate determination. Moreover, systematic perturbation to the model helps to identify control targets for the desired change in cell fate.

Based on previous studies, we postulate that core regulatory circuits consist of subnetworks interconnected by feedback loops, and that these circuits are the primary drivers of cell fate. Further advancing this concept, we can propose a detailed procedure as follows. First, we can investigate all feedback loops associated with different cellular states of interest (Fig. 3b). Subsequently, through an analysis of network dynamics influenced by feedback loop states, we can prioritize the positive feedback loops that are instrumental in controlling the attractor landscape governing cell-fate determination (Fig. 3c). For this, we can employ previous control methods such as LDOI<sup>52</sup>, stable motif<sup>54</sup>, and DEPC<sup>55</sup>, or improve these methods for further exploration of key subnetworks. As a result, the prioritized feedback loops can be refined to form a core regulatory circuit. Lastly, through systematic perturbation to the model (or core regulatory circuit) or by employing the improved kernel-based methods, we can identify (minimal) master regulators for the desired cell-fate change (Fig. 3d). Together, this control theory-based approach would be a highly promising way to understand and manipulate cell-fate processes, centered around core regulatory circuits, of a dynamic system.

From this, we integrate aforementioned progresses and suggest a comprehensive framework for systematically identifying core regulatory circuits and their master regulators for cell-fate change (Fig. 4). This framework consists of three interconnected components: capturing dynamic information from single-cell omics data, constructing and analyzing a mathematical model of molecular mechanisms in cell-fate determination, and identifying the candidates of master regulators for a desired cell-fate control (Fig. 4a–c). To explicitly introduce our framework, we provide an illustrative example as follows.

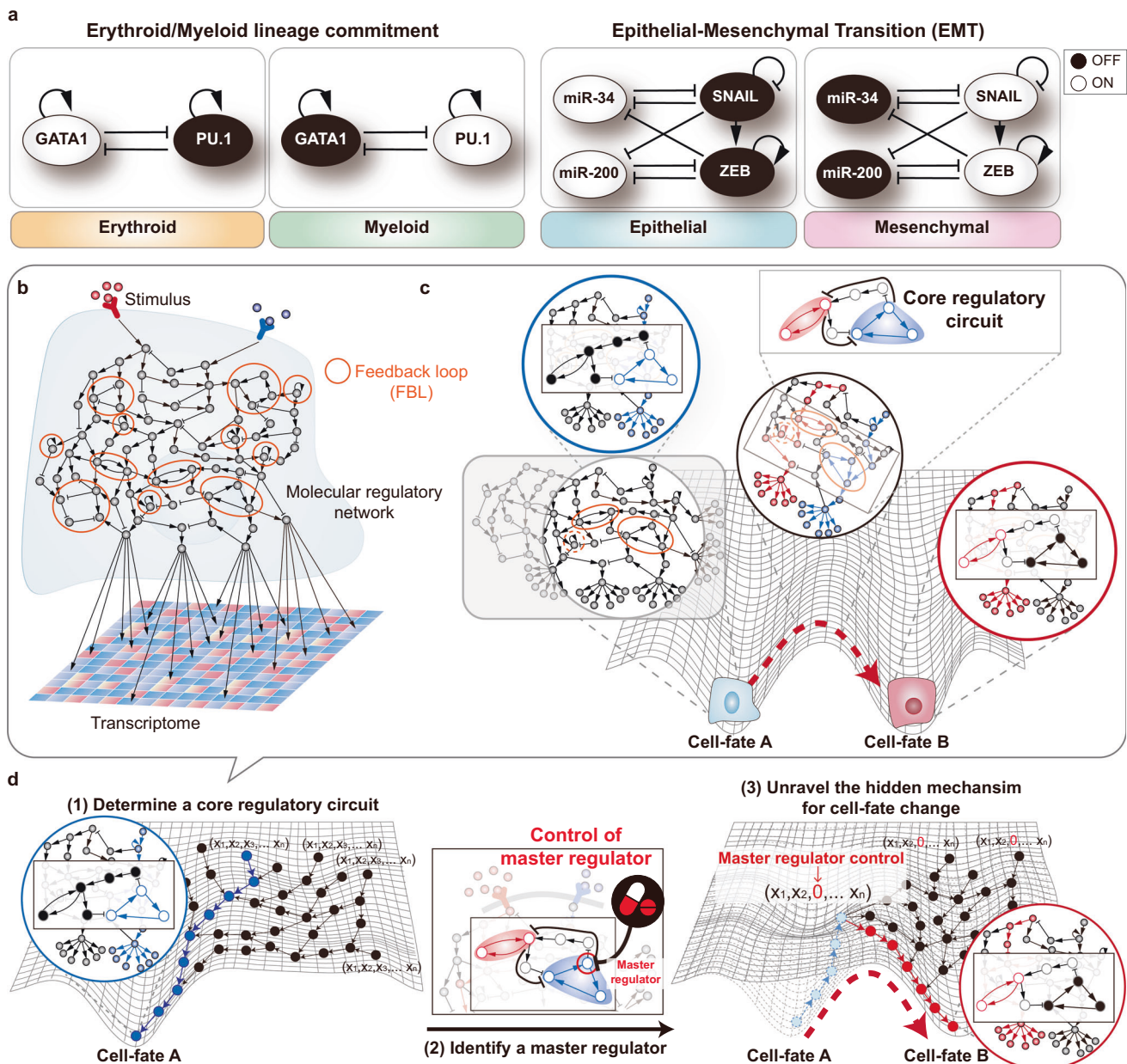
### Illustrative example: identifying control targets for cancer reversion by using single-cell RNA-sequencing data from lung cancer samples

Traditional anticancer therapies have focused on removing cancer cells, but their effectiveness is limited due to the inevitable emergence of resistance, which arises as a consequence of cancer cell plasticity. Cancer plasticity is an emerging hallmark of cancer cells, and it plays a crucial role in cancer initiation and progression, as well as adaptation to therapy and intratumoral heterogeneity<sup>74–76</sup>. Hence, recent research has shifted focus towards targeting highly plastic cancer cells emerging during tumorigenesis, aiming to inhibit their development.

This illustrative example shows the emerging concept of ‘cancer reversion’, which seeks to transform highly plastic lung cancer cells back into normal cells, instead of merely eliminating cancer cells. Recent studies have indicated that cells with high plasticity emerge during mouse lung tumorigenesis upon introduction of KRAS mutations into AT2 cells, which are typically considered the origin of lung cancer. These cells lose their original characteristics and transition through AT1/AT2-like states, eventually diversifying into various cancer cell types such as EMT-type, embryonic liver-like, GI-like, and high-cycling cells, thereby increasing tumor heterogeneity (Fig. 5a).

In this illustrative example, we begin with obtaining time-series scRNA-seq data collected during lung tumorigenesis from AT2 cell state to high-plasticity cell state (HPCS) from a public repository<sup>77</sup>. This dataset provides single-cell transcriptomic information for ~2200 cells. The proposed framework entails preprocessing, dimensionality reduction, and clustering to annotate cell identities at different stages during tumorigenesis. Then, it conducts pseudotime ordering to delineate the trajectory from AT2 cells, through intermediate states, to HPCS cells, providing a comprehensive map of cellular evolution during lung tumorigenesis (Fig. 5b).

The proposed framework identifies critical TFs that can regulate gene expression changes along the lung tumorigenesis trajectory by comparing the activities of TFs among the distinct cell clusters. Then, the interactions



**Fig. 3 | Core regulatory circuits in a dynamic network model.** **a** Examples of key regulatory components for cell-fate determination. **b** Representation of regulation dynamics at the network-level. The molecular regulatory network is useful to describe the regulation dynamics inferred from transcriptome data in response to a stimulus. From these molecular regulations, all feedback loops are investigated to find out an essential subnetwork. **c** Core regulatory circuit for cell-fate determination. As each cell fate is defined by the gene activities within the molecular regulatory network, alteration in a few critical molecular interactions can lead to a shift towards a specific cell fate. In other words, a specific subnetwork composed of interconnected positive feedback loops (referred to as a ‘core regulatory circuit’) is considered the

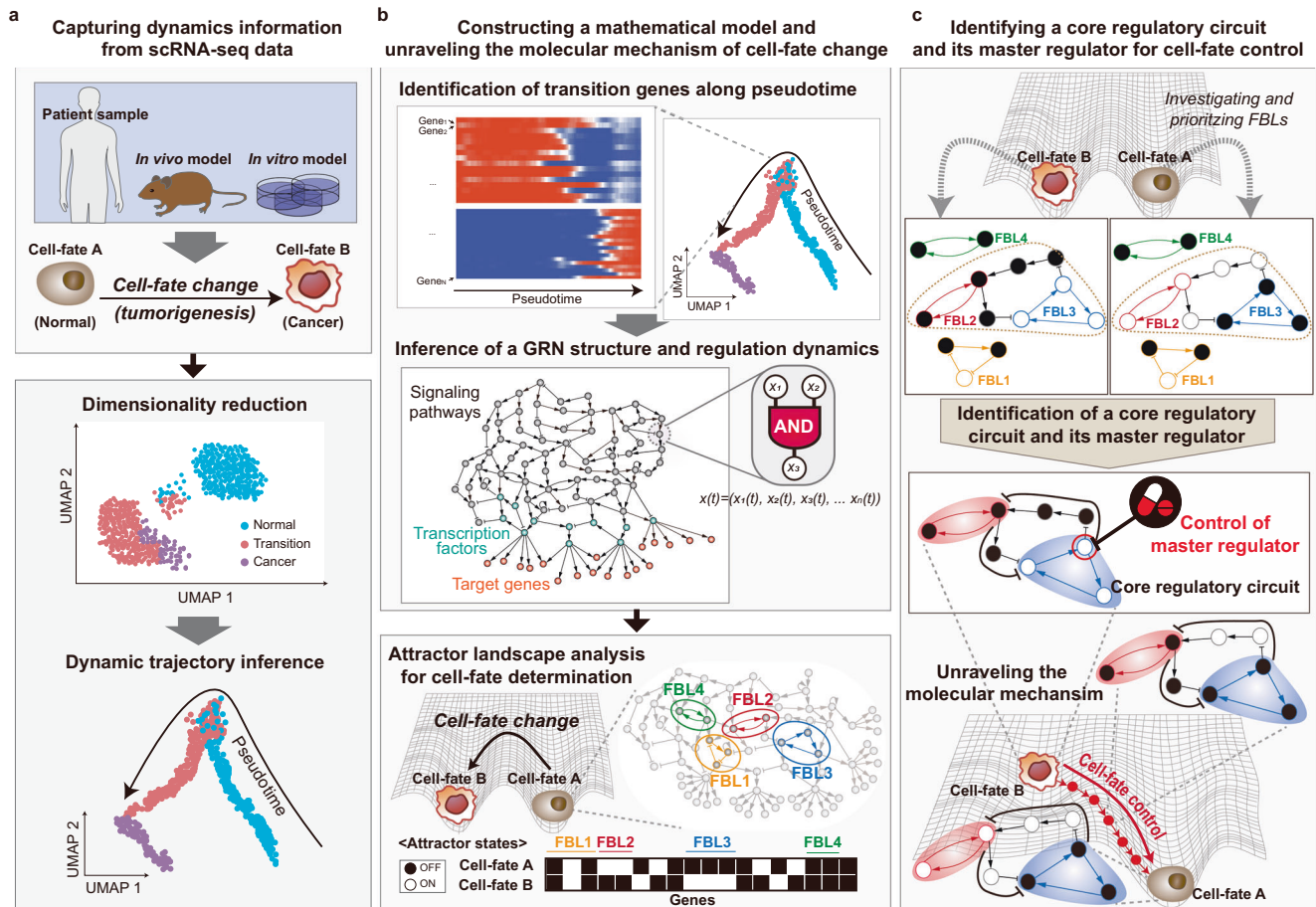
primary driver of cell-fate changes. For example, it could be a toggle switch, a well-known designed circuit in cell-fate decision-making. **d** Elucidating the molecular mechanism through a core regulatory circuit and its master regulator. Each cell fate can be regarded as an attractor within the attractor landscape that includes the transition trajectory. First, based on detailed analysis as outlined in **b**, **c**, a core regulatory circuit essential for cell-fate determination can be extracted. Next, a master regulator that governs the change of cellular states described by the core regulatory circuit can be identified. Finally, the hidden mechanism underlying the transition when the master regulator is controlled can be unraveled.

among these TFs can be inferred, resulting in a GRN structure consisting of 105 nodes and 304 links (Fig. 5c). Next, to explore essential regulatory mechanisms associated with specific phenotypes, the framework can investigate all positive feedback loops showing differential activation between AT2 and HPCS states by evaluating their TF activities. By aggregating those feedback loops, the framework generates a reduced GRN, consisting of 23 nodes and 74 links, which contains key dynamical information (Fig. 5d). Intriguingly, TFs within the same state (e.g., AT2 or HPCS) are primarily interconnected through mutually activating positive feedback,

whereas TFs across distinct states (e.g., AT2 versus HPCS) engage predominantly through mutually inhibitory interactions.

With the GRN structure and trajectory information, a Boolean logic model can be constructed to simulate the tumorigenesis process. This process involves discretizing the continuous expression values of each TF along the pseudotime trajectory (employing clustering methods like k-means). Subsequently, the influence of each TF in the GRN structure can be determined through Boolean logic functions, utilizing the QM algorithm (Fig. 5e). Using the Boolean network model, the framework can pinpoint the





**Fig. 4 | A framework for identifying core regulatory circuits and master regulators for cell-fate change.** **a** Capturing critical information on dynamics from scRNA-seq data. ScRNA-seq data can be obtained from cancer samples, in vivo models, or in vitro models. From this collected dataset, the dynamics of cellular states during cell-fate changes (e.g., tumorigenesis) can be captured by the following steps (top). First, dimensionality reduction and clustering methods are generally employed to clarify normal, transition, and tumor states. Then we utilize dynamic trajectory inference methods to reveal the trajectories during cell-fate change (bottom). **b** Constructing a mathematical model to analyze the molecular mechanism. We identify transition genes that play a significant role in changing cellular states during cell-fate determination. A GRN structure and regulation

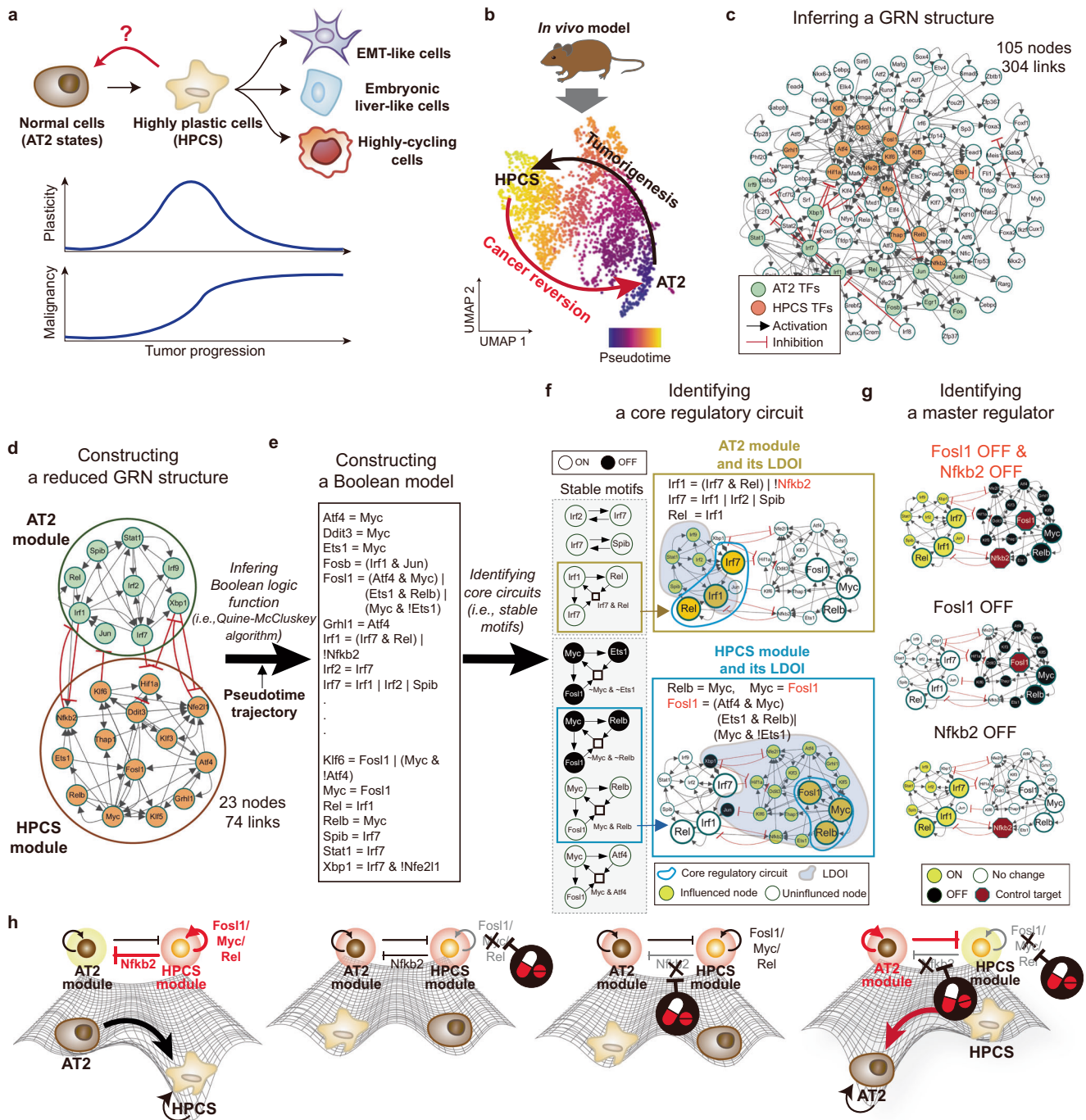
dynamics are inferred by integrating these identified genes and the molecular dynamic interactions that connect them (top). The model is further investigated through attractor landscape analysis to elucidate the underlying mechanism for cell-fate determination. Based on the corresponding attractors, feedback loops can be selected from the GRN structure (bottom). **c** Identifying a core regulatory circuit and its master regulator for cell-fate control. Through evaluating and prioritizing feedback loops, we can identify a crucial network of interconnected feedback loops as a core regulatory circuit, pinpointing the most influential regulator as the master regulator. Analyzing how the states of feedback loops within this circuit change when the master regulator is controlled allows us to reveal the molecular mechanisms underlying cell-fate control.

feedback loops of the most dynamic significance (in this example, stable motifs). These feedback loops and their corresponding LDOIs have the potential to stabilize either phenotype when modulated. Then the framework can assess the influence of the LDOIs on AT2 and HPCS modules, and prioritize key feedback loops based on their relevance to either phenotype (Fig. 5f). The most influential feedback loops can form a core regulatory circuit, consisting of Rel, Irf1, Irf7, Fos1, Myc, and Relb. Once the core regulatory circuit is stabilized, it can fix the values of most TFs within the respective AT2 and HPCS modules.

The proposed framework then leverages information on core regulatory circuits to find control targets that can inhibit the HPCS module and activate the AT2 module. The LDOIs for individual nodes and pairs of nodes are calculated, then each LDOI is checked whether it can fix the influential feedback loops in the desired state to identify control targets. For example, a combination of Fos1 (or Myc) and Nfkb2 can be identified as a master regulator, thereby ensuring the stable activation of the AT2 module and the deactivation of the HPCS module (Fig. 5g). Inhibition of Fos1 can deactivate HPCS TF activities while having no impact on AT2 TFs. On the other hand, disruption of Nfkb2 can de-repress AT2 TF activities in the HPCS state, activating downstream regulators and leading to the re-emergence of

the AT2 state. Combined inhibition of E2f4 and Nfkb2 can destabilize the HPCS state, and enable a stable activation of the AT2 module. This combinatorial inhibition strategy can, therefore, effectively disrupt the positive feedback within the HPCS module while also neutralizing its negative impact on the AT2 module. The result shows a stable shift from a highly plastic cancerous state back to a normal-like AT2 state. Moreover, the role of key molecules in plastic cancer cells can be illustrated using the landscape concept (Fig. 5h). In these cells, Fos1 and Nfkb2 are highly active, stabilizing the HPCS while repressing the normal AT2 module. This ensures the persistence of the HPCS. However, to revert HPCS back to normal, merely inhibiting either Fos1 or Nfkb2 is insufficient, as each alone does not fully shift the landscape towards the AT2 phenotype. The positive feedback of Fos1 maintains the HPCS state, or the remaining activity of Nfkb2 suppresses the AT2 module. Only the simultaneous inhibition of both Fos1 and Nfkb2 can significantly alter the landscape, favoring a shift towards the AT2 phenotype.

In summary, this example illustrates how systems biology can be applied to understand and manipulate cell-fate transitions in cancer. By integrating mathematical modeling with experimental data, we can identify crucial regulatory mechanisms that reverse cell states, offering new insights



and potential therapeutic strategies in cancer treatment. Although this example illustrates the whole process proposed for inducing cancer reversion from public single-cell data, limitations are noted. Cell-fate determination is a complex phenomenon that encompasses changes not just at the transcriptomic level but also across various molecular levels. The continuous development of technologies that utilize single-cell multiomics data, which integrate transcriptomic, proteomic, and epigenetic data, has been pivotal. The use of such data can significantly enhance the proposed framework. In particular, the recent increase in single-cell multiomics data, combining scRNA-seq and scATAC-seq, has been noteworthy. scATAC-seq allows for the pruning of indirect regulation information from functional regulatory relationships obtained from transcriptomic data by providing information on the open chromatin regions at promoter sites. This enables the construction of more accurate GRN structures and,

consequently, the development of more precise mathematical models. Furthermore, factors like the physiological state of a cell, not directly included in omics data, can significantly influence cell fate<sup>78</sup>. Although our framework does not directly incorporate such state information, it can be indirectly reflected through TF activity within the GRN structure and its mathematical model. The future availability of technologies for simultaneous measurement of the physiological and molecular states of a cell promises the development of models that integrate these dimensions for a comprehensive understanding.

### Conclusions

Among the various molecules within a cell, only a few key molecules have a significant impact on cell fate. What distinguishes these master regulators from other molecules? How can we identify those master regulators, and



**Fig. 5 | Identifying control targets for reversing lung cancer cell state.** **a** Emerging role of cancer plasticity in tumorigenesis. When normal cells (e.g., AT2 states) are exposed to KRAS mutations alone or combined with TP53 mutations, these cells evolve into highly plastic cells (e.g., HPCS). This progression results in increased plasticity, leading to diversification into various cancer cell types, such as EMT-like, embryonic liver-like, and highly-cycling cells, thereby enhancing malignancy. **b** Low dimensional representation of tumorigenesis and cancer reversion. After collecting and preprocessing the dataset obtained from the in vivo model, dimensionality reduction and clustering are employed to annotate cellular identities at different stages during tumorigenesis. UMAP representation visualizes the two cellular states, AT2 and HPCS, colored by pseudotime. **c** The resulting GRN structure. A gene regulatory network is composed of critical transcription factors (TFs) that regulate gene expression changes along the lung tumorigenesis trajectory. **d** Identifying differential modules between AT2 and HPCS states. The network comprises selected feedback loops with differential activation between AT2 and HPCS. It reveals mutually activating interactions within each state and inhibitory interactions between distinct states. The structure is reminiscent of a toggle switch, a well-known design principle in cell-fate determination, suggesting a fundamental regulatory

structure in tumorigenesis. **e** Inferring Boolean network model. A Boolean network model is inferred from the network structure and pseudotime trajectory by applying Quine-McCluskey algorithm. **f** Identifying a core regulatory circuit. By analyzing the Boolean network model, we can identify crucial feedback loops essential for phenotype stabilization. Assessing the impact of their LDOIs on AT2 and HPCS modules reveals a key circuit critical for phenotype stabilization. This core regulatory circuit, including *Rel*, *Irf1*, *Irf7*, *Fosl1*, *Myc*, and *Relb*, can play a crucial role in controlling cell fates. **g** Identifying a master regulator. Inhibition of only *Fosl1* decreases TF activities of HPCS (top). *Nfkb2* acts as a major negative regulator of the positive feedback loop among AT2 TFs (middle). Disrupting only *Nfkb2* leads to the re-emergence of the AT2 states. The combined inhibition of *Fosl1* and *Nfkb2* effectively destabilizes HPCS, thereby facilitating a stable transition from a cancer state back to a normal AT2 state (bottom). **h** Conceptual illustration of the reversion dynamic landscape through combinatorial inhibition strategy. Targeting either molecule alone proves inadequate due to their integral feedback mechanisms. Conversely, simultaneous inhibition of both *Fosl1* and *Nfkb2* is necessary and effective for a significant cellular landscape shift, driving the transition toward the AT2 phenotype.

through what molecular regulatory mechanisms do they induce cell-fate changes? To answer these questions, it is necessary to analyze and understand the behavior of a huge molecular regulatory network within a cell. In the past, such attempts were limited due to experimental constraints. Yet, recent advances in single-cell omics technologies, along with a decade of progress in network control technologies, enable us to answer these fundamental questions and usher in a renaissance for systems biology. The conceptual framework we introduce can offer an unprecedented opportunity for cell fate control by integrating the latest technological innovations into a comprehensive, novel strategy.

Received: 11 December 2023; Accepted: 16 April 2024;  
Published online: 06 May 2024

## References

- Moris, N., Pina, C. & Arias, A. M. Transition states and cell fate decisions in epigenetic landscapes. *Nat. Rev. Genet.* **17**, 693–703 (2016).
- Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
- Pancieria, T. et al. Induction of expandable tissue-specific stem/progenitor cells through transient expression of YAP/TAZ. *Cell Stem Cell* **19**, 725–737 (2016).
- Dow, L. E. et al. Apc restoration promotes cellular differentiation and reestablishes crypt homeostasis in colorectal cancer. *Cell* **161**, 1539–1552 (2015).
- Shin, S. Y. et al. Functional roles of multiple feedback loops in extracellular signal-regulated kinase and Wnt signaling pathways that regulate epithelial-mesenchymal transition. *Cancer Res.* **70**, 6715–6724 (2010).
- Hong, J. Y. et al. Computational modeling of apoptotic signaling pathways induced by cisplatin. *BMC Syst. Biol.* **6**, 122 (2012).
- Park, S. G. et al. The influence of the signal dynamics of activated form of IKK on NF- $\kappa$ B and anti-apoptotic gene expressions: a systems biology approach. *FEBS Lett.* **580**, 822–830 (2006).
- Lee, H. S., Hwang, C. Y., Shin, S. Y., Kwon, K. S. & Cho, K. H. MLK3 is part of a feedback mechanism that regulates different cellular responses to reactive oxygen species. *Sci. Signal.* **7**, ra52 (2014).
- Waddington, C. H. *The strategy of the genes.* (Routledge, 2014).
- Wang, J., Zhang, K., Xu, L. & Wang, E. Quantifying the Waddington landscape and biological paths for development and differentiation. *Proc. Natl. Acad. Sci. USA* **108**, 8257–8262 (2011).
- Zhang, J., Nie, Q. & Zhou, T. Revealing dynamic mechanisms of cell fate decisions from single-cell transcriptomic data. *Front. Genet.* **10**, 1280 (2019).
- Joo, J. I., Zhou, J. X., Huang, S. & Cho, K.-H. Determining relative dynamic stability of cell states using boolean network model. *Sci. Rep.* **8**, 12077 (2018).
- Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp. Mol. Med.* **50**, 1–14 (2018).
- Satiya, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
- Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
- Amezquita, R. A. et al. Orchestrating single-cell analysis with Bioconductor. *Nat. Methods* **17**, 137–145 (2020).
- Jolliffe, I. T. & Cadima, J. Principal component analysis: a review and recent developments. *Philos. Trans. A Math. Phys. Eng. Sci.* **374**, 20150202 (2016).
- Van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
- McInnes, L., Healy, J. & Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv* <https://arxiv.org/abs/1802.03426> (2018).
- Traag, V. A., Waltman, L. & Van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
- Ester, M., Krieger, H.-P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. *In Kdd* **96**, 226–231 (1996).
- Guo, M., Wang, H., Potter, S. S., Whitsett, J. A. & Xu, Y. SINCERA: a pipeline for single-cell RNA-seq profiling analysis. *PLoS Comput. Biol.* **11**, e1004575 (2015).
- Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
- La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
- Gayoso, A. et al. Deep generative modeling of transcriptional dynamics for RNA velocity analysis in single cells. *Nat. Methods* **21**, 50–59 (2024).
- Street, K. et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 1–16 (2018).
- Wolf, F. A. et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 9 (2019).
- Herman, J. S., Sagar, N. & Gruen, D. FateID infers cell fate bias in multipotent progenitors from single-cell RNA-seq data. *Nat. Methods* **15**, 379–386 (2018).
- Badia, I. M. P. et al. Gene regulatory network inference in the era of single-cell multi-omics. *Nat. Rev. Genet.* **24**, 739–754 (2023).

30. Qiu, X. et al. Inferring causal gene regulatory networks from coupled single-cell expression dynamics using scribe. *Cell Syst.* **10**, 265–274. e211 (2020).
31. Greenfield, A., Madar, A., Ostrer, H. & Bonneau, R. DREAM4: combining genetic and dynamic information to identify biological networks and dynamical models. *PLoS One* **5**, e13397 (2010).
32. Huynh-Thu, V. A., Irthum, A., Wehenkel, L. & Geurts, P. Inferring regulatory networks from expression data using tree-based methods. *PLoS One* **5**, e12776 (2010).
33. Moerman, T. et al. GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics* **35**, 2159–2161 (2019).
34. Chan, T. E., Stumpf, M. P. H. & Babbie, A. C. Gene regulatory network inference from single-cell data using multivariate information measures. *Cell Syst.* **5**, 251–267. e253 (2017).
35. Specht, A. T. & Li, J. LEAP: constructing gene co-expression networks for single-cell RNA-sequencing data using pseudotime ordering. *Bioinformatics* **33**, 764–766 (2017).
36. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
37. Bravo Gonzalez-Blas, C. et al. SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nat. Methods* **20**, 1355–1367 (2023).
38. Fleck, J. S. et al. Inferring and perturbing cell fate regulomes in human brain organoids. *Nature* **621**, 365–372 (2023).
39. Wang, L. et al. Dictys: dynamic gene regulatory network dissects developmental continuum with single-cell multiomics. *Nat. Methods* **20**, 1368–1378 (2023).
40. Kamimoto, K. et al. Dissecting cell identity via network inference and in silico gene perturbation. *Nature* **614**, 742–751 (2023).
41. Lim, C. Y. et al. BTR: training asynchronous Boolean models using single-cell expression data. *BMC Bioinformatics* **17**, 355 (2016).
42. Woodhouse, S., Piterman, N., Wintersteiger, C. M., Gottgens, B. & Fisher, J. SCNS: a graphical tool for reconstructing executable regulatory networks from single-cell genomic data. *BMC Syst. Biol.* **12**, 59 (2018).
43. Luo, S., Wang, Z., Zhang, Z., Zhou, T. & Zhang, J. Genome-wide inference reveals that feedback regulations constrain promoter-dependent transcriptional burst kinetics. *Nucleic Acids Res.* **51**, 68–83 (2023).
44. Matsumoto, H. et al. SCODE: an efficient regulatory network inference algorithm from single-cell RNA-Seq during differentiation. *Bioinformatics* **33**, 2314–2321 (2017).
45. Matsumoto, H. & Kiryu, H. SCROUP: a probabilistic model based on the Ornstein-Uhlenbeck process to analyze single-cell expression data during differentiation. *BMC Bioinformatics* **17**, 232 (2016).
46. Osorio, D. et al. scTenifoldKnk: an efficient virtual knockout tool for gene function predictions via single-cell gene regulatory network perturbation. *Patterns* **3**, 100434 (2022).
47. Black, H. S. Stabilized feedback amplifiers. *Bell Syst. Tech. J.* **13**, 1–18 (1934).
48. Wolkenhauer, O., Kitano, H. & Kwang-Hyun, C. Systems biology. *IEEE Control Syst. Mag.* **23**, 38–48 (2003).
49. Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. Controllability of complex networks. *nature* **473**, 167–173 (2011).
50. Valente, T. W., Coronges, K., Lakon, C. & Costenbader, E. How correlated are network centrality measures? *Connect (Tor.)* **28**, 16–26 (2008).
51. Liu, Y. Y., Slotine, J. J. & Barabasi, A. L. Control centrality and hierarchical structure in complex networks. *PLoS One* **7**, e44459 (2012).
52. Yang, G., Gomez Tejeda Zanudo, J. & Albert, R. Target control in logical models using the domain of influence of nodes. *Front. Physiol.* **9**, 454 (2018).
53. Mochizuki, A., Fiedler, B., Kurosawa, G. & Saito, D. Dynamics and control at feedback vertex sets. II: a faithful monitor to determine the diversity of molecular activities in regulatory networks. *J. Theor. Biol.* **335**, 130–146 (2013).
54. Zanudo, J. G. & Albert, R. Cell fate reprogramming by control of intracellular network dynamics. *PLoS Comput. Biol.* **11**, e1004193 (2015).
55. Crespo, I., Perumal, T. M., Jurkowski, W. & del Sol, A. Detecting cellular reprogramming determinants by differential stability analysis of gene regulatory networks. *BMC Syst. Biol.* **7**, 140 (2013).
56. Yang, J.-M., Lee, C.-K. & Cho, K.-H. Robust stabilizing control of perturbed biological networks via coordinate transformation and algebraic analysis. *In: IEEE Transactions on Neural Networks and Learning Systems* (2022).
57. Steinway, S. N. et al. Network modeling of TGFbeta signaling in hepatocellular carcinoma epithelial-to-mesenchymal transition reveals joint sonic hedgehog and Wnt pathway activation. *Cancer Res.* **74**, 5963–5977 (2014).
58. Choi, S. R., Hwang, C. Y., Lee, J. & Cho, K. H. Network analysis identifies regulators of basal-like breast cancer reprogramming and endocrine therapy vulnerability. *Cancer Res.* **82**, 320–333 (2022).
59. Kim, N., Hwang, C. Y., Kim, T., Kim, H. & Cho, K. H. A cell-fate reprogramming strategy reverses epithelial-to-mesenchymal transition of lung cancer cells while avoiding hybrid states. *Cancer Res.* **83**, 956–970 (2023).
60. An, S. et al. Inhibition of 3-phosphoinositide-dependent protein kinase 1 (PDK1) can revert cellular senescence in human dermal fibroblasts. *Proc. Natl Acad. Sci. USA* **117**, 31535–31546 (2020).
61. Seo, C. H., Kim, J. R., Kim, M. S. & Cho, K. H. Hub genes with positive feedbacks function as master switches in developmental gene regulatory networks. *Bioinformatics* **25**, 1898–1904 (2009).
62. Kumar, L. et al. Molecular mechanisms and applications of N-acyl homoserine lactone-mediated quorum sensing in bacteria. *Molecules* **27**, <https://doi.org/10.3390/molecules27217584> (2022).
63. Chickarmane, V., Troein, C., Nuber, U. A., Sauro, H. M. & Peterson, C. Transcriptional dynamics of the embryonic stem cell switch. *PLoS Comput. Biol.* **2**, e123 (2006).
64. Xu, J. & Orkin, S. H. The erythroid/myeloid lineage fate paradigm takes a new player. *Embo J.* **30**, 983–985 (2011).
65. Lu, M., Jolly, M. K., Levine, H., Onuchic, J. N. & Ben-Jacob, E. MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination. *Proc. Natl Acad. Sci. USA* **110**, 18144–18149 (2013).
66. Hornung, G. & Barkai, N. Noise propagation and signaling sensitivity in biological networks: a role for positive feedback. *PLoS Comput. Biol.* **4**, e8 (2008).
67. Kim, J. R. et al. Reduction of complex signaling networks to a representative kernel. *Sci. Signal* **4**, ra35 (2011).
68. Kim, J., Park, S. M. & Cho, K. H. Discovery of a kernel for controlling biomolecular regulatory networks. *Sci. Rep.* **3**, 2223 (2013).
69. An, S. et al. Global stabilizing control of large-scale biomolecular regulatory networks. *Bioinformatics* **39**, btad045 (2023).
70. Kwon, Y. K. & Cho, K. H. Boolean dynamics of biological networks with multiple coupled feedback loops. *Biophys. J.* **92**, 2975–2981 (2007).
71. Deritei, D., Rozum, J., Ravasz Regan, E. & Albert, R. A feedback loop of conditionally stable circuits drives the cell cycle from checkpoint to checkpoint. *Sci. Rep.* **9**, 16430 (2019).
72. Hu, Y. et al. Optimal control nodes in disease-perturbed networks as targets for combination therapy. *Nat. Commun.* **10**, 2180 (2019).
73. Rukhlenko, O. S. et al. Control of cell state transitions. *Nature* **609**, 975–985 (2022).
74. Torborg, S. R., Li, Z. X., Chan, J. E. & Tammela, T. Cellular and molecular mechanisms of plasticity in cancer. *Trends Cancer* **8**, 735–746 (2022).
75. Barkley, D., Rao, A., Pour, M., Franca, G. S. & Yanai, I. Cancer cell states and emergent properties of the dynamic tumor system. *Genome Res.* **31**, 1719–1727 (2021).

76. Hanahan, D. Hallmarks of cancer: new dimensions. *Cancer Discov.* **12**, 31–46 (2022).
77. Marjanovic, N. D. et al. Emergence of a high-plasticity cell state during lung cancer evolution. *Cancer Cell* **38**, 229–246.e213 (2020).
78. Zhu, J., Chu, P. & Fu, X. Unbalanced response to growth variations reshapes the cell fate decision landscape. *Nat. Chem. Biol.* **19**, 1097–1104 (2023).
79. Papili Gao, N., Ud-Dean, S. M. M., Gandrillon, O. & Gunawan, R. SINCERITIES: inferring gene regulatory networks from time-stamped single cell transcriptional expression profiles. *Bioinformatics* **34**, 258–266 (2018).
80. Zhang, S. et al. Inference of cell type-specific gene regulatory networks on cell lineages from single cell omic datasets. *Nat. Commun.* **14**, 3064 (2023).
81. Kartha, V. K. et al. Functional inference of gene regulation using single-cell multi-omics. *Cell Genom* **2**, 100166 (2022).

### Acknowledgements

The authors thank Corbin Hopper for his critical reading and comments. This work was supported by the National Research Foundation of Korea (NRF) grants funded by the Korean government, the Ministry of Science and ICT (2023R1A2C3002619 and 2021M3A9I4024447 (Bio & Medical Technology Development Program)).

### Author contributions

All authors contributed to the discussion of the content, writing, reviewing, and editing the manuscript. K.-H.C. conceived the idea, designed the project, and supervised the study.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to Kwang-Hyun Cho.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024