

RESEARCH ARTICLE

DNFE: Directed network flow entropy for detecting tipping points during biological processes

Xueqing Peng^{1,2‡}, Rui Qiao^{1,2‡}, Peiluan Li^{1,2*}, Luonan Chen^{3,4,5*}

1 School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang, China, **2** Longmen Laboratory, Luoyang, Henan, China, **3** School of Mathematical Sciences and School of AI, Shanghai Jiao Tong University, Shanghai, China, **4** Guangdong Institute of Intelligence Science and Technology, Hengqin, China, **5** Key Laboratory of Systems Health Science of Zhejiang Province, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, China

‡ These authors share co-first authorship on this work.

* 9902639@haust.edu.cn (PL); lnchen@sjtu.edu.cn (LC)



OPEN ACCESS

Citation: Peng X, Qiao R, Li P, Chen L (2025) DNFE: Directed network flow entropy for detecting tipping points during biological processes. *PLoS Comput Biol* 21(7): e1013336. <https://doi.org/10.1371/journal.pcbi.1013336>

Editor: Xiao-Jun Tian, Arizona State University, UNITED STATES OF AMERICA

Received: February 18, 2025

Accepted: July 14, 2025

Published: July 29, 2025

Copyright: © 2025 Peng et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The code of DNFE is publicly available at GitHub (<https://github.com/Peng-xq/DNFE.git>) and BioCode (<https://ngdc.cncb.ac.cn/biocode/tool/BT7564>).

Funding: o This work was supported by National Key R&D Program of China (No. 2022YFA1004800 to LC), National Natural Science Foundation of China (Nos. 12131020,

Abstract

Typically, in dynamic biological processes, there is a critical state or tipping point that marks the transition from one stable state to another, surpassing which a considerable qualitative shift takes place. Identifying this tipping point and its driving network is essential to avert or delay disastrous outcomes. However, most traditional approaches built upon undirected networks still suffer from a lack of robustness and effectiveness when implemented based on high-dimensional small-sample data, especially for single-cell data. To address this challenge, we develop a directed network flow entropy (DNFE) method, which can transform measured omics data into a directed network. This method is applicable to both single-cell RNA-sequencing (scRNA-seq) and bulk data. Applying this algorithm to six real datasets, including three single-cell datasets, two bulk tumor datasets, and a blood dataset, the method is proved to be effective not only in identifying critical states, as well as their dynamic network biomarkers, but also in helping explore regulatory relationships between genes. Numerical simulation results demonstrate that the DNFE algorithm is robust across various noise levels and outperforms existing methods in detecting tipping points. Furthermore, the numerical simulations for 100-node and 1000-node gene regulatory networks illustrate the method's application for large-scale data. The DNFE method predicts active transcription factors, and further identified “dark genes”, which are usually overlooked with traditional methods.

Author summary

Numerous sophisticated systems undergo critical transitions, resulting in a sudden shift from one state to a markedly different state. Typically, in dynamic biological processes, there is a critical state or tipping point that marks the

31930022, T2350003, T2341007 to LC), Science and Technology Commission of Shanghai Municipality (No. 23JS1401300 to LC), Special Fund for Science and Technology Innovation Strategy of Guangdong Province (Nos. 2021B0909050004, 2021B0909060002 to LC), Key-Area Research and Development Program of Guangdong Province (No. 2021B0909060002 to LC), Natural Science Foundation of Henan Province (No.242300420242 to PL), R&D project of Pazhou Lab (Huangpu) (No.2023K0602), Key Science and Technology Research Project of Henan Province (No. 252102520024 to PL). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

transition from one stable state to another, surpassing which a considerable qualitative shift takes place. Identifying this tipping point and its driving network is essential to avert or delay disastrous outcomes. However, most traditional approaches built upon undirected networks still suffer from a lack of robustness and effectiveness when implemented based on high-dimensional small-sample data, especially for single-cell data. To address this challenge, we develop a directed network flow entropy (DNFE) method, which can transform measured omics data into a directed network. This method is applicable to both single-cell RNA-sequencing (scRNA-seq) and bulk data. Applying this algorithm to six real datasets, including three single-cell datasets, two bulk tumor datasets, and a blood dataset, the method is proved to be effective not only in identifying critical states, as well as their dynamic network biomarkers, but also in helping explore regulatory relationships between genes. Our results suggest that the DNFE method predicts active transcription factors, and further identified “dark genes”, which are usually overlooked with traditional methods.

Introduction

Numerous sophisticated systems undergo critical transitions, resulting in a sudden shift from one state to a markedly different state [1]. Single-cell RNA sequencing has revolutionized the investigation of cellular heterogeneity and functional diversity, enabling gene expression measurements within individual cells and allowing for inferences of cell population arrangements based on the trajectory topology and gene regulatory analysis [2–4]. A critical state exists in many biological processes, such as cell differentiation or embryonic development, where a marked transition or change in cell populations takes place [5–8]. Detecting a tipping point immediately prior to such critical transitions during biological processes is essential for shedding light on the intrinsic mechanisms involved in disease advancement [9,10]. However, owing to the similarity between the normal state and critical state, with respect to phenotype, along with average gene expression, traditional biomarkers often fail to detect these critical states.

To indicate the tipping point preceding the transition of biological systems, a theoretical framework referred to as the dynamic network biomarker (DNB) was introduced [11]. Based on the DNB theory, numerous methods have been proposed and utilized to identify the tipping points of complex diseases, such as influenza [12], breast cancer [13], and bladder urothelial carcinoma [14], as well as cell differentiation [15,16]. However, the use of this approach in single-cell RNA-sequencing (scRNA-seq) data analysis is limited because of obvious disturbances from amplification noise in transcripts and the occurrence of dropout events. Moreover, most of these computational approaches merely identify correlations among molecules/genes based on undirected networks, yet directed networks can reflect the interactions among genes and mine the underlying dynamic information in cell populations. In addition, directed networks can disclose the fundamental mechanisms underlying

molecular effects, which provides quantitative data to support further precise therapies. Directed networks can explicitly represent directional relationships between genes, which is crucial for describing causal relationships. Directed networks can also gain a better understanding of the regulatory relationships between genes, and identifying important regulatory factors and pathways. Numerous GRN inference methods based on directed networks have also been developed, as highlighted in recent reviews [17]. For instance, ANANSE [18] prioritizes static regulatory edges via a network-based method that uses properties of enhancers and their GRNs to predict key TFs in cell fate determination, and DeepMAPS [19] leverages building a heterogeneous graph containing both cells and genes. While these methods effectively identify gene-gene interactions, none yield similar instability scores. Thus, proposing a resilient and effective method is urgent for detecting critical states based on directed networks, especially for single-cell data from a complex biological process.

To address these issues, we developed a novel computational algorithm based on directed networks, called the directed network flow entropy (DNFE) method, to predict critical transitions in bulk and single-cell data. Unlike conventional GRN inference approaches, our method first constructed a time-specific directed network at a given time point. We utilize the DNFE scores as instability scores to describe the molecular collective fluctuation qualify the criticality of the biological process (The detailed differences are presented in [S1 Table](#)). Differing from conventional methods, we prioritize the first-order neighbors that have a direct connection to the core gene, in conjunction with the second-order neighbors that directly interact with any first-order neighbors. Specifically, we first construct a time-specific directed network at a given time point with a direction determination index. We then calculate the local DNFE for each local directed network according to the information of the directed network. Unlike traditional information entropy, DNFE fully utilizes local (core gene) network connectivity information rather than fluctuations in gene expression, thereby reliably quantifying network fluctuation. Finally, we utilize the DNFE score to describe the molecular collective fluctuation resulting from specific samples against the reference samples/cells and qualify the criticality of the biological process, i.e., critical collective fluctuation. DNFE is an effective means of identifying the tipping points during the course of complex diseases, and it has the following advantages: (i) The DNFE method decreases noise by leveraging dynamic and high-dimensional information from omics data through the reliable quantification of critical network fluctuations, thereby increasing its robustness. Moreover, we focus on the second-order neighbors, which better depict the structure of the network, thereby improving the efficiency of the method. The numerical simulation shows the algorithm's robustness, effectiveness, and applicability in handling vast amounts of data or large-scale datasets. (ii) Unlike most traditional methods, the DNFE method is proposed based on directed networks, which can help us gain a better understanding of the regulatory relationships between genes, identify important regulatory factors and pathways, and explore the structure and functionality of a gene regulatory network. (iii) The DNFE method can detect critical states before state transitions occur and recognize relevant DNB members with critical collective fluctuations and non-differential "dark genes", potentially serving as prognostic biomarkers of complex diseases and biological processes. (iv) The DNFE method can not only be applied to bulk and scRNA-seq data but can also be used to identify transcription factors (TFs) that serve as fundamental players influencing cell identity and guiding shifts in cell fate as well. One such identified TF is *ZNF888*, which is considered to hold great potential in early cell differentiation.

To illustrate the resilience and efficiency of DNFE, we conducted numerical simulations using gene expression data from an artificial gene regulatory network, varying the noise strength. With the increase in noise strength, DNFE outperformed existing methods [20–23] in detecting early-warning signals for impending critical states. Furthermore, the numerical simulation, conducted on a 1000-node network, unequivocally showed the algorithm's robustness and applicability in handling vast amounts of data or large-scale datasets. The DNFE method was applied to two bulk sequencing tumor datasets, including kidney renal papillary cell carcinoma (KIRP) and bladder cancer (BLCA) from The Cancer Genome Atlas (TCGA) database, effectively identifying critical states. Furthermore, cell fate commitment was accurately observed in three single-cell datasets, including the differentiation of human embryonic stem cells (hESCs) into definitive endoderm cells (DECs), mouse ESCs (mESCs) into mesoderm progenitor (MP) cells, and mouse embryonic fibroblasts (MEFs)

into neurons. DNFE identified two critical states (at 12 and 36 h of differentiation) in the differentiation process data for DEC. The tipping points at 12 and 36 h may represent the critical points of ES-to-ME and ME-to-DE transitions. Finally, the DNFE method was applied to a blood dataset of patients with a non-complicated SLE pregnancy (SLE-NC). Immunological changes in neutrophils and T cells at five time points were detected, which aligns with the transcriptional changes surrounding embryo implantation monitored during assisted reproductive technology, thereby demonstrating the efficacy of DNFE.

Results

Validation based on numerical simulation

To evaluate the efficacy and resiliency of the developed DNFE method, we used an eleven-node artificial network (Fig 1A) to illustrate the detection of early-warning signals when the system nears a critical state. Models based on Michaelis–Menten kinetics or Hill equations are widely used to describe the dynamics within gene regulatory networks [23–25],

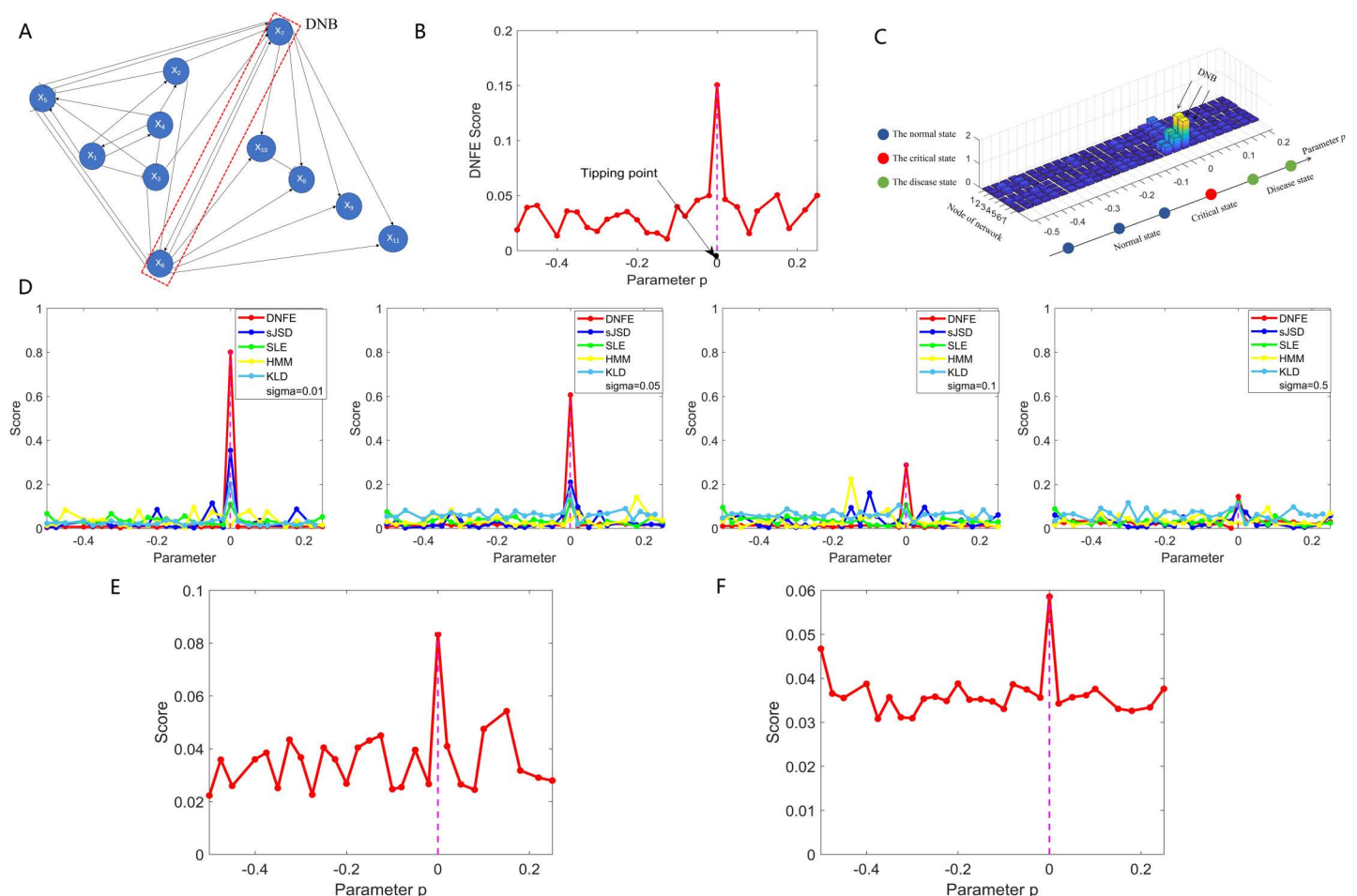


Fig 1. Validation of the directed network flow entropy (DNFE) method based on simulation datasets. (A) The 11-node network model depicts positive regulation with arrows and negative regulation with solid lines. (B) The curve of the DNFE score derived from the gene regulatory network. (C) The landscape of the DNFE score for 11 nodes. (D) Assessment of the stability of DNFE compared to alternative methods with varying noise intensities. (E) DNFE score curve based on the 100-node gene regulatory network. (F) DNFE curve score based on the 1000-node gene regulatory network.

<https://doi.org/10.1371/journal.pcbi.1013336.g001>

including processes, like transcription and translation [26,27], cyclic reactions [28], nonlinear biological processes [29,30], and other gene-regulatory activities [27,29]. A parameter varying between -0.5 and 0.23, with a bifurcation parameter value identified as the tipping point, was used to generate two simulated datasets.

Fig 1B depicts the dynamic variation in the DNFE score for the 11-node network. A marked increase in the DNFE score occurs at a specific parameter value, signaling an upcoming critical state. To effectively demonstrate the difference between the normal state and the pre-disease state, the local DNFE scores for the 11 individual networks are presented in Fig 1C, showcasing the overall landscape of the network flow entropy. Notably, as the system is distant from the critical state, the DNFE scores for all nodes remain stable and low. However, as the system nears the critical state, certain nodes exhibit abruptly altered expression variations and network connections (i.e., DNB), leading to a pronounced surge in the DNFE score, suggesting an impending critical state.

To ascertain the method's robustness, as depicted in Fig 1D, we compared the DNFE algorithm with established DNB methods across varying levels of noise. With an escalating intensity of noise, the DNFE method outperforms the existing methods, such as single-sample-based Jensen-Shannon Divergence (sJSD) method [20], single-sample landscape entropy (SLE) method [21], hidden Markov model (HMM) method [22], and Kullback-Leibler divergence (KLD) method [23] in stably delivering early-warning signals of critical transition, thereby verifying its superior efficacy and robustness.

In addition, two simulation datasets were generated using a 100-node and a 1000-node structure depicted by stochastic differential equations to detect early-warning signals of critical transitions in genetic regulatory networks. The dynamic changes in DNFE score for the 100-node network and the 1000-node network are illustrated Fig 1E and 1F, respectively. An abrupt increase in the DNFE score is observed at a specific parameter value $p = 0$, indicating the proximity to a critical state at a bifurcation point ($p = 0$). As the system nears the tipping point, an abrupt increase in the DNFE score is observed around the bifurcation point $p = 0$. Notably, the DNFE method demonstrates robustness when tested with the even larger 1000-node network. Additionally, the application of the algorithm to the simulation dataset generated by the 1000-node network resulted in quick calculations, underscoring the DNFE method's applicability for large-scale datasets. Detailed simulation and calculation information can be found in S2 File and S1 Fig.

Predicting critical states for hESC-to-DEC

To depict the efficiency of the developed method for single-cell datasets, we utilized DNFE on three scRNA-seq datasets of cell differentiation: hESC-to-DEC data, mESC-to-MP data, and MEF-to-neuron data. The application of the DNFE method to hESC-to-DEC data is illustrated within the primary text, and the analytical results using mESC-to-MP data and MEF-to-neuron data are provided in S3 File, S2 Fig for mESC-to-MP and S3 Fig for MEF-to-neuron data.

As depicted in Fig 2A, two critical states are identified by a sudden increase in the DNFE score: the first tipping point occurs at 12 h ($P=0.05$) and the second at 36 h ($P=0.03$). We have associated errorbars to Fig 2B to describe the changes in DNFE scores over time. As shown in this figure, the bar chart represents the average DNFE score at each time point, while the standard deviation (SD) reflects the dispersion of the data points. Calculations reveal that the standard deviation at each time point is predominantly within the range of 0.1 to 0.2. These transitions indicate two distinct differentiation processes following critical transitions, which are ES-to-ME and ME-to-DE. The top 5% of genes exhibiting the highest DNFE scores are designated as DNBs, and they are especially responsive to the tipping point that precedes disease deterioration. Fig 2C illustrates the landscape of the global DNFE score, where the score abruptly increases at 12 h and 36 h, suggesting identified tipping points or critical states. We visualized the score distribution for each gene, as illustrated in the Fig 2D, which shows the DNFE score abruptly increases at 12 h and 36 h. Fig 2E displays the dynamic development of the directed PPI network of DNFE signal biomarkers. There is an obvious and substantial shift in the DNFE score of the network following the critical point, suggesting that the network structure experiences a notable change. The nodes represent genes, and directed edges represent the interactions between genes. Moreover, as illustrated in Fig 2F and 2G, the gene expression levels of DNB genes alone do not effectively differentiate the critical state

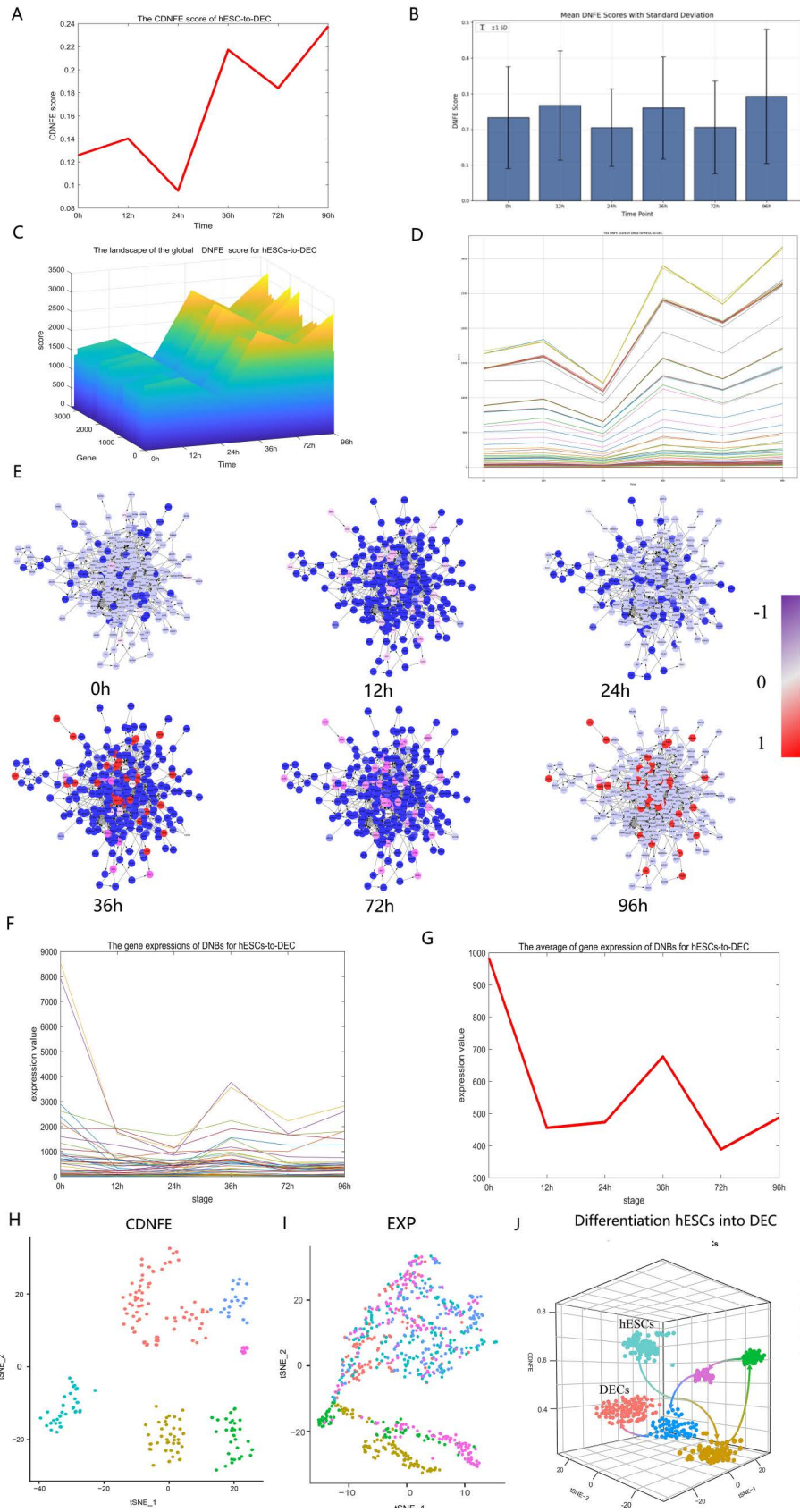


Fig 2. Critical states of the human embryonic stem cell (hESC) differentiation process revealed by directed network flow entropy (DNFE). (A) The DNFE scores at 12 and 36 h of hESC differentiation are elevated compared to those at the other time points, suggesting that these intervals represent critical states in the process of hESC differentiation. (B) The errorbars at each time point. (C) The landscape of the global DNFE score for hESC-to-definitive endoderm cell (DEC) transitions. The dynamic network biomarker (DNB) scores reach their peak at 12 h and 36 h, indicating the two tipping points. (D) The DNFE scores of DNBs for hESC-to-DEC transitions. (E) The gene-directed regulatory network's dynamic evolution is established through signaling genes. The nodes and directed edges represent genes and the interactions between genes, respectively. (F) The gene expression of DNBs for hESC-to-DEC transitions. (G) The mean gene expression of DNBs for hESC-to-DEC transitions. Clearly, gene expression levels of DNBs are insufficient to differentiate the critical state from other states, while the DNFE method can successfully distinguish between them. (H)-(J) The nodes, represented in various colors, correspond to cells at different time points. The DNFE method is superior to the EXP method at distinguishing temporal cell states. Additionally, DNFE scores can provide accurate predictions of differentiation trajectories.

<https://doi.org/10.1371/journal.pcbi.1013336.g002>

from other states. However, utilizing the DNFE scores of these DNB genes, the critical state can be identified successfully, as demonstrated in Fig 2A.

As our sample size met cluster analysis criteria (the sample size should be at least 5 times greater than the perplexity value, with a default value typically set at 30), we performed it to further demonstrate that DNFE can distinguish between different cellular states. The clustering analyses depicted in Fig 2H and 2I reveal that clustering analysis based on DNFE can effectively differentiate between cellular states at various time points, outperforming gene expression-based methods. The quantitative assessment of clustering performance is conventionally benchmarked using the Adjusted Rand Index (ARI). DNFE-based clustering demonstrated superior performance (ARI=0.78) compared to expression-based clustering (ARI=0.15). Network-based correlation works better than floating expression via gene expression. By employing the DNFE score, we can effectively mitigate the high noise and enhance the robustness of gene expression data. To further verify the capabilities of DNFE, pseudo-trajectory analysis was conducted on the hESC-to-DEC data. Performing temporal cell clustering via DNFE, three-dimensional representations of cell-lineage trajectories are presented in Fig 2H, where the z-axis signifies DNFE's potency estimation, while the x-axis and y-axis correspond to the t-SNE components. Fig 2J showcases the differentiation trajectories of cells transitioning from hESCs to DECs, with differentiation towards DECs observed after 36 h, aligning with the experimental findings [31]. These findings underscore the capability of DNFE-based potency estimation to monitor dynamic changes in cell potency and pinpoint the precise time points of cell fate commitment or differentiation into various cell types.

Identifying TFs in the endodermal differentiation of hESCs

The top 5% of genes with the highest DNFE scores serve as DNBs. And the DNB genes do not demonstrate differential expression but are sensitive to DNFE scores being termed "dark genes". As expected, the DNB genes at 12 h showed higher scores than those of the two surrounding time points (0 and 24 h) (Fig 3A and 3B), similarly to the DNB genes at 36 h (Fig 3C and 3D).

Next, we aimed to pinpoint the possible upstream transcriptional regulators of the DNB genes associated with the two critical states. We focused on TFs due to their crucial role in defining cell identity and driving cell fate transitions. We predicted the TFs on the CHEA3 website and identified 20 TFs for the first critical state (12 h) and 20 TFs for the second critical state (36 h) based on the chosen top 150 DNBs. These two groups of TFs could regulate 88% and 85% of the DNBs in each tipping point, respectively (Fig 3E and 3F). Among these DNB factors, *ZNF146* is associated with a poor cancer prognosis and functions as a pro-proliferative and growth-sustaining transcriptional regulator in cells, exerting an obvious effect on cellular dynamics [32]. *CEBPZ* is the key TF that regulates various aspects of cellular differentiation and functions in a variety of tissues [33]. The overexpression of *PRMT3* facilitates cell proliferation, migration, and invasion. Moreover, *PRMT3* stabilizes *C-MYC*, and the pro-proliferation role of *PRMT3* depends on *C-MYC* [34]. Within hESCs, *ZNF207* collaborates with key pluripotency TFs to regulate self-renewal and maintain pluripotency, while also directing cell commitment towards the ectoderm by directly regulating neuronal TFs, holding a crucial position in differentiation [35]. As

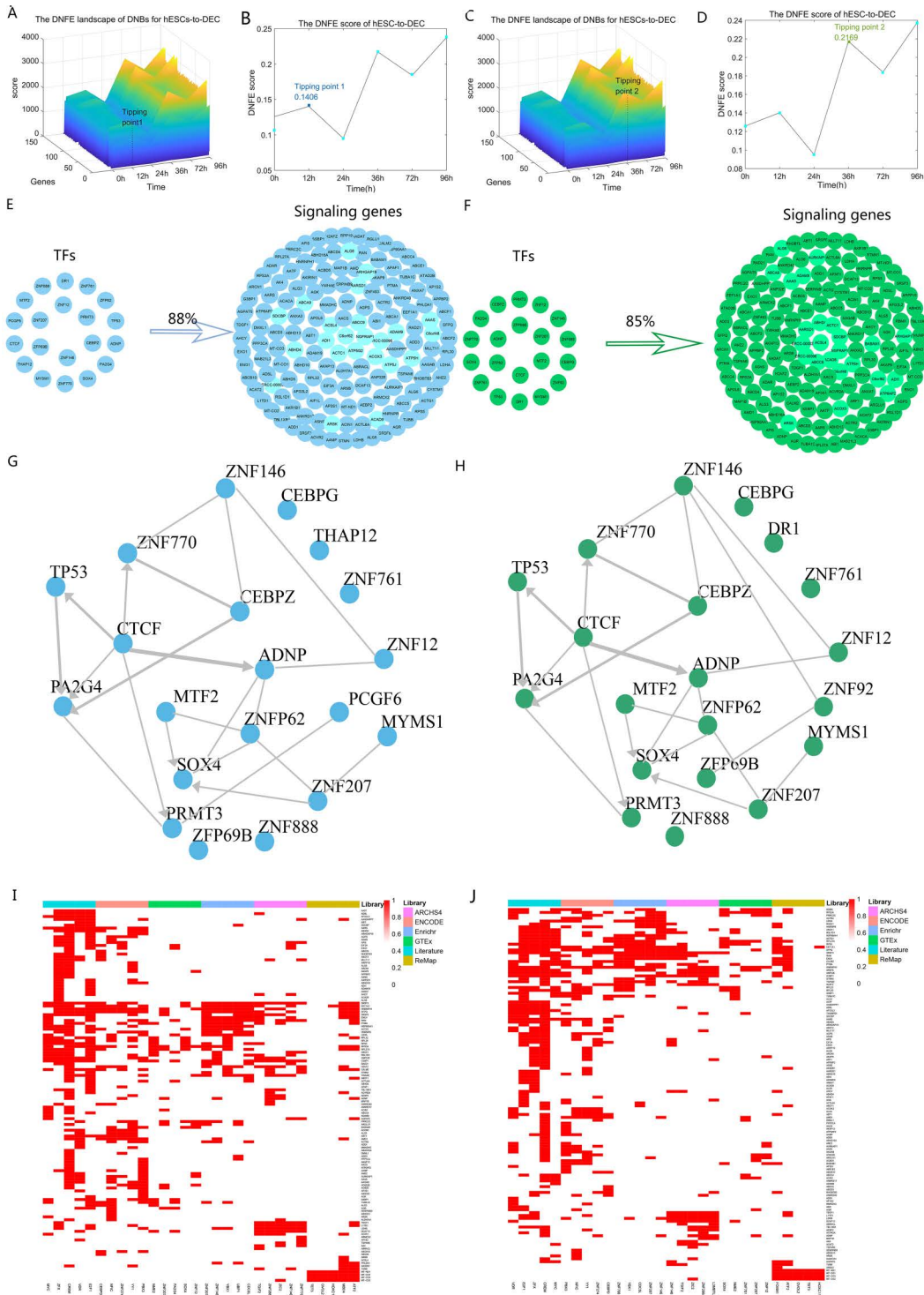


Fig 3. Critical states of human embryonic stem cell (hESC) differentiation process revealed using the directed network flow entropy (DNFE) method. (A) The expression dynamics of dynamic network biomarker (DNB) genes at the 12h mark during hESC differentiation highlight the initial stages of cellular commitment. (B) The DNFE scores at 12h provide a quantitative measure of the systemic perturbations within the cellular network, indicative of the first critical transition. (C) The behavior of DNB genes at the 36h mark further delineates the gene expression patterns as differentiation progresses. (D) The DNFE scores at 36h offer a subsequent quantitative assessment of network disruption, marking the second critical transition. (E)

It was discovered that a mere twenty pivotal upstream transcription factors are capable of regulating a substantial 88% of the DNB genes identified at the first critical state (12 h), underscoring the efficiency of the regulatory network. (F) Similarly, at the second critical state (36 h), the same set of twenty transcription factors can modulate 85% of the DNB genes, demonstrating their sustained influence. (G) The regulatory interactions among the top 20 transcription factors at the first tipping point were analyzed, revealing the complex interplay that governs early differentiation events. (H) At the second tipping point, the regulatory relationships among the top 20 transcription factors were also examined, providing a deeper understanding of the later stages of cellular commitment. (I) The matrix representing the top 5 transcription factors from each library, aligned with query genes at the first tipping point, was constructed to show the presence of query genes within the target gene sets of library transcription factors. (J) A similar matrix was created for the second tipping point, illustrating the targeted gene regulation by the top 5 transcription factors, as identified using each library.

<https://doi.org/10.1371/journal.pcbi.1013336.g003>

demonstrated in [Fig 3G](#), the local network shows the results of the mutual regulation between the predicted top 20 TFs in the regulatory network at the first tipping point, similarly to the second tipping point ([Fig 3H](#)). Particularly, among the forecasted top 20 TFs, *ZNF888* has been pinpointed as a methylation-influenced gene correlated with clear-cell renal cell carcinoma. It can enhance cellular proliferation through the suppression of *CDKN1A*, a gene that inhibits the cell cycle [\[36\]](#). Nonetheless, given the scant research on *ZNF888*, its precise role remains elusive, necessitating additional investigations to fully elucidate its function in cellular processes. The heatmaps in [Fig 3I](#) and [3J](#) display whether a gene is included in the target gene set of TFs from the library. The top five TFs are included in each library, and each column represents each TF. *SRSF3* and *HNRNPH1* appear most frequently in the target gene set of the TF library at the two tipping points. *SRSF3* is regarded as more important in the termination of transcription than in cleavage, potentially through its interaction with RNA sequences downstream of the cleavage site [\[37\]](#). Both in vivo and in vitro studies have indicated that silencing *HNRNPH1* hinders cell proliferation and enhances apoptosis in Chronic Myeloid Leukemia cells [\[38\]](#).

In summary, using the DNFE method, we have successfully detected two critical states that occur in the differentiation of hESCs into the endoderm. Additionally, the analysis has led to the prediction of twenty TFs that can serve as pivotal regulators in determining cell fate during this differentiation process.

Revealing functional roles of DNBs and “dark genes” for hESC-to-DEC

To delve deeper into the uncharted territory of regulatory elements in cancer-related pathways, as highlighted in [\[39\]](#), we conducted a comparative analysis between the DNB genes and the differentially expressed genes (DEGs). Our investigation revealed the presence of certain genes within the DNBs that did not exhibit differential expression at the molecular level but displayed elevated DNFE scores at the network level. These genes, termed “dark genes”, may have an apparent influence within the network context despite their lack of differential expression at the individual gene level.

To identify important pathways to characterize the potential biological mechanisms of “dark genes” for hESC-to-DEC data, we next performed gene set enrichment analysis (GSEA). Detailed information from the hallmark pathway enrichment analysis is shown in [S2 Table](#). As shown in [Fig 4A](#), the crucial pathways included the cAMP signaling pathway, the spliceosome, and pathogenic *Escherichia coli* infection. In addition, we conducted Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis to uncover the potential biological functions of the “dark genes” and their first-order DEGs within the DNB genes, exhibiting molecular-level differential expression. The enrichment analysis revealed that these “dark genes” and their DEG counterparts were evidently enriched in key signaling pathways, such as JAK–STAT signaling, the PI3K–Akt signaling pathway, the cell cycle, and pathways in cancer, along with other cancer-related pathways ([Fig 4B](#) and [4C](#)). Notably, the antagonism of JAK–STAT signaling has been shown to impede the progression of pre-neoplastic lesions to malignant tumors [\[40\]](#). The PI3K–Akt signaling pathway has a key function in tumor cell differentiation, proliferation, and apoptosis [\[41\]](#). Abnormal cell proliferation in cancer is associated with dysregulation of the cell cycle [\[42\]](#). The involvement of multiple signaling “pathways in cancer” underscores its complexity and close relationship with the disease [\[43\]](#). Aberrant regulation of the cell cycle is an important malignant characteristic of cancer, potentially influencing cancer progression and metastasis [\[44\]](#).

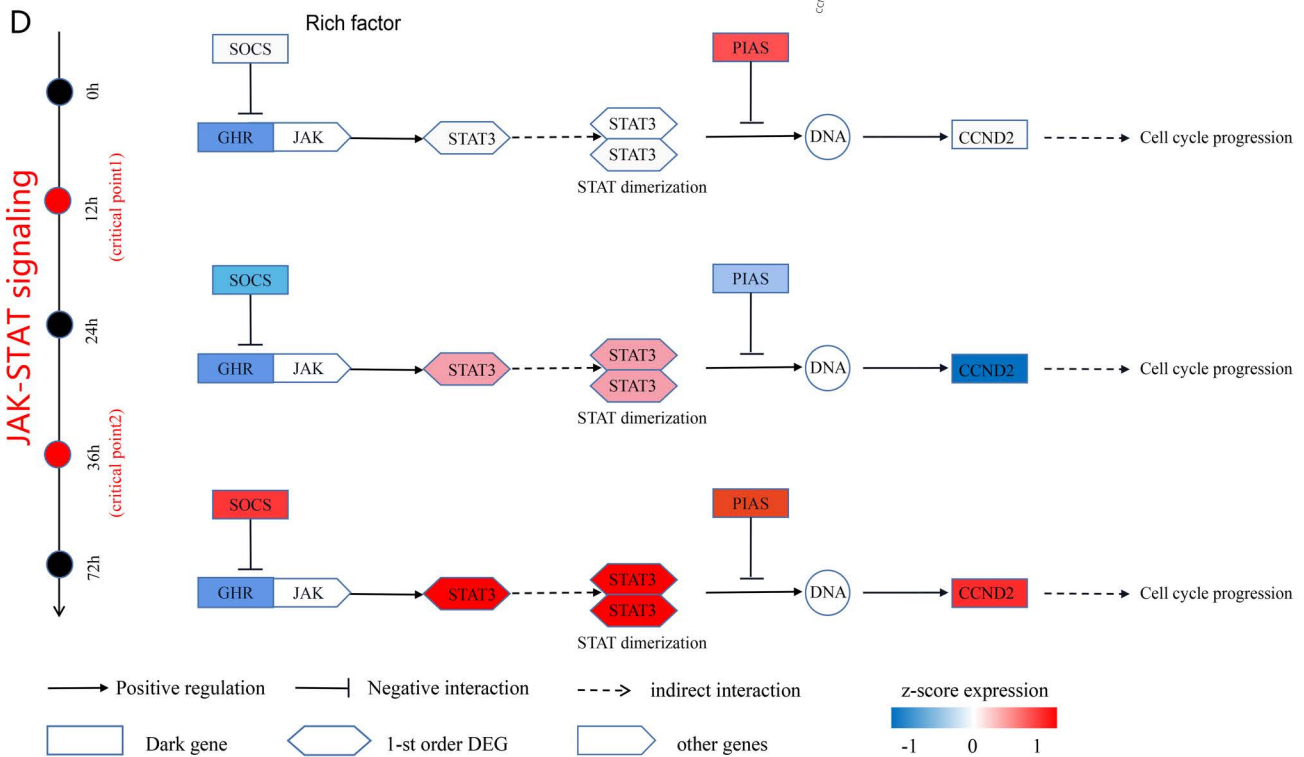
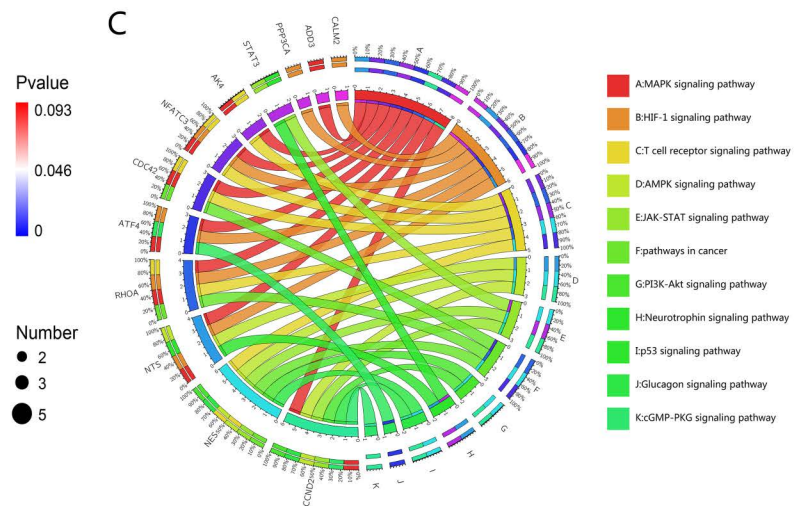
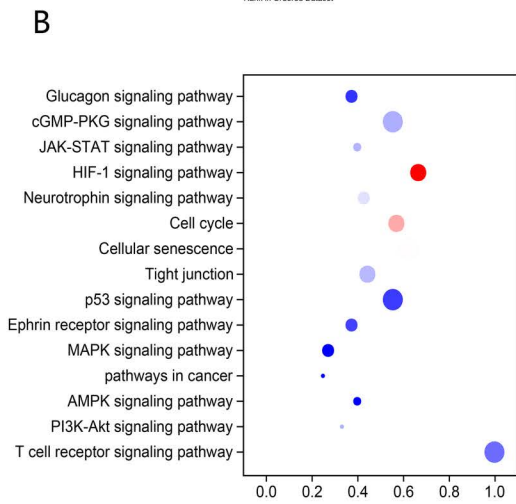
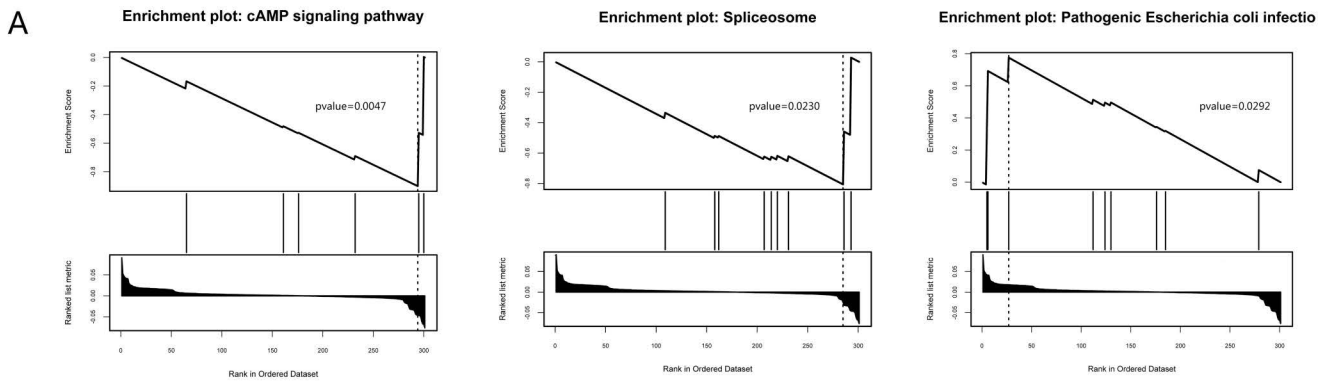


Fig 4. Regulatory mechanisms of embryo development revealed by the “dark genes” in the human embryonic stem cell (hESC)-to-definitive endoderm cell (DEC) process. (A) Gene set enrichment analysis of “dark genes”. (B) Enrichment analysis of Kyoto Encyclopedia of Genes and Genomes pathways for the “dark genes” in the hESC-to-DEC process. (C) The outer ring’s left segment delineates the dynamic network biomarkers (DNBs) recognized using the DNFE method, highlighting their presence in the differentiation process. The outer ring’s right segment corresponds to the array of biological processes in which these DNBs participate, showcasing their multifaceted roles. Within the inner ring, the color-coding and link thickness are utilized to represent the diversity of enriched pathways and the statistical significance of gene functions, respectively. (D) The “dark genes” and their first-order differentially expressed gene neighbors are further examined to uncover the intrinsic signaling mechanisms.

<https://doi.org/10.1371/journal.pcbi.1013336.g004>

Our findings indicate that the combined effect of the “dark genes” and their first-order neighboring DEGs regulates cell cycle progression near the critical transition via the JAK–STAT pathway (Fig 4D). The DNB gene *GHR* (growth hormone receptor), identified as the most upstream signaling molecule, does not exhibit a notable alteration in gene expression and is classified as a “dark gene” from a DNFE standpoint [5]. Suppressors of cytokine signaling (*SOCS*) genes inhibit the activation of *STAT3* mediated by *GHR* [45]. The first-order neighboring DEG, *STAT3*, shows obvious changes before and after the tipping point, which may cause dimerization of STAT, which in turn enters the nucleus and directly regulates the expression of downstream signaling genes [46]. Additionally, protein inhibitors of activated STATs have been detected as antagonists of cytokine-regulated signaling, inhibiting the activity of STAT TFs [47]. Moreover, our results indicate that expression of the *CCND2* gene changes notably after entering the critical state, which may facilitate cell cycle progression in cancer cells [48]. These observations support our hypothesis that signals for cell cycle progression are conveyed in a sequential manner across the cancer development process via the JAK–STAT pathway, with the ultimate effects becoming apparent only after the transition at the tipping point.

Detection of critical states for KIRP

To verify the computational efficacy of our developed method, we applied the DNFE method to two tumor datasets (KIRP, BLCA) from the TCGA database. The applications of the DNFE method to KIRP are depicted in the main text, and the results of the BLCA study are presented in S4 File and S4 Fig.

Fig 5A reveals a sudden increase in DNFE scores at Stage II, indicating the tipping point. This is followed by the distant metastasis of cancer, leading to an obvious deterioration in the patient’s tumor state. The top 5% of genes with the highest DNFE scores at the tipping point were detected as DNBs. Fig 5B illustrates the tipping point, where the DNFE score in DNBs shows an evident increase during Stage II. Cancer-related mortality is largely attributed to tumor progression, including metastasis, which entails the migration of carcinoma cells from the primary tumor to distant sites [49]. Consequently, identifying the critical state is essential for preventing or preparing for potential deterioration, allowing for timely and effective clinical interventions. A prognostic analysis was conducted using clinical data from KIRP samples, focusing on durations leading up to and subsequent to the critical state (Stage II). Fig 5C illustrates a notable difference in the survival curves for KIRP samples prior to and following Stage II. Additionally, the survival times of samples preceding the tipping point are considerably longer compared to those succeeding it. The abrupt decline in survival times among patients intensely indicates that Stage II represents the tipping point in KIRP. The development of KIRP is a multistage process, with the risk of deterioration increasing from one stage to the next. We further investigated the deterioration risk of KIRP by categorizing samples from each stage into two categories according to the median DNFE score: high-score samples and low-score samples. Subsequently, we conducted a comparison of the prognoses of these two categories. Fig 5D illustrates that the red curve denotes the survival probability of the high-score samples, while the light blue curve denotes the survival probability of the low-score samples. Statistically, the p-values obtained from the KIRP prognostic analysis were 0.00084 (Stage I-II), 0.019 (Stage III), and 0.036 (Stage IV). The p-values resulting from the log-rank test in the Kaplan–Meier analysis are below 0.05, indicating that the prognostic analysis is significant at a statistical level. Individuals within the high-score cohort at Stages I–II exhibit a longer survival duration than their counterparts in the low-score cohort. Conversely, in the post-critical state, the high-score group experiences a markedly reduced survival duration

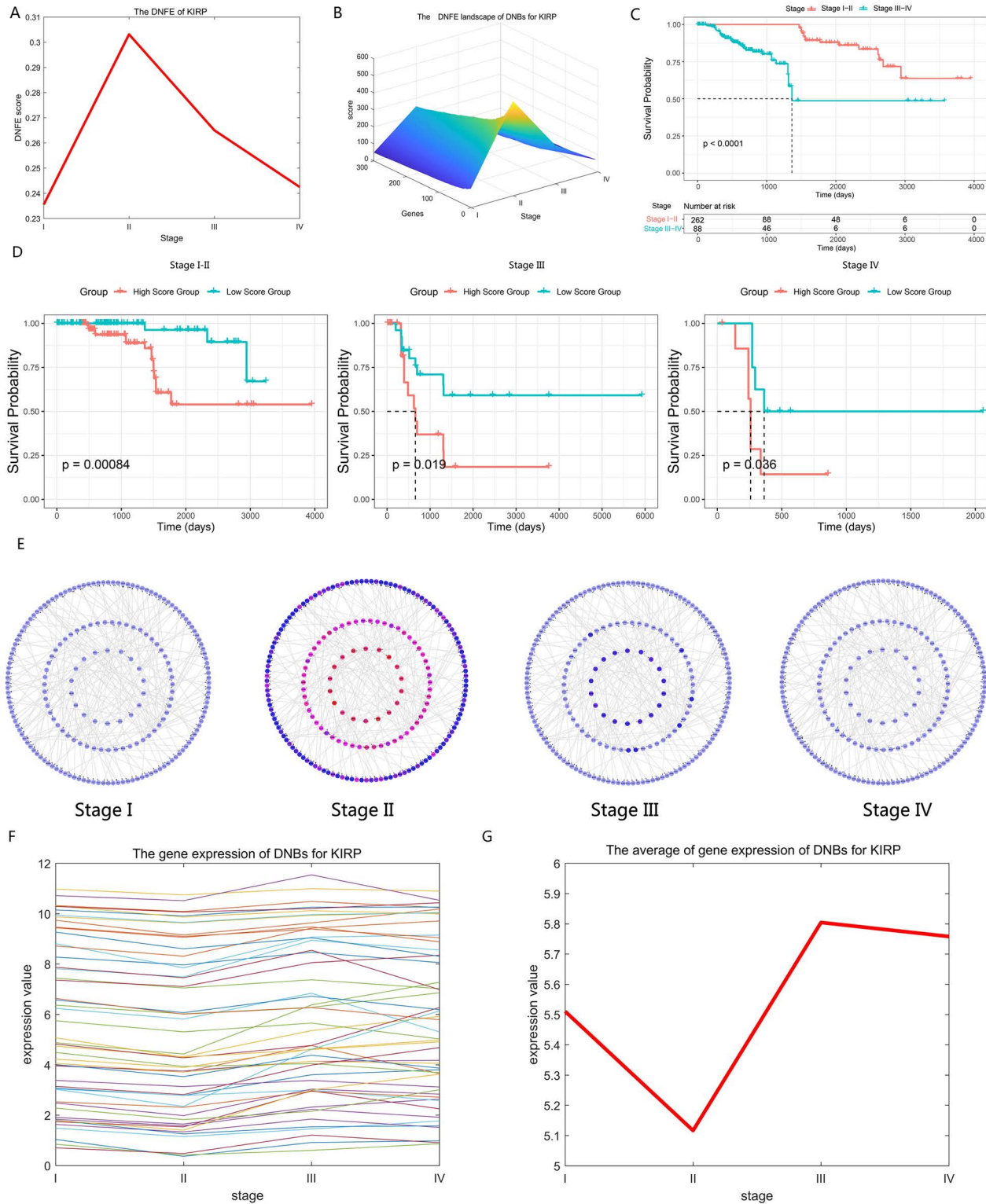


Fig 5. Detection of critical state for kidney renal papillary cell carcinoma (KIRP). (A) The directed network flow entropy (DNFE) score trajectory for KIRP reveals a notable increase at Stage II, signifying a critical transition point in the disease's progression. (B) The DNFE landscape for KIRP's dynamic network biomarkers (DNBs) exhibits a marked increase at Stage II, underscoring the importance of this phase as a critical state in the disease

trajectory. (C) A comparison of survival curves for KIRP delineates a stark difference, with patients experiencing an obviously longer survival time prior to Stage II compared to those post-Stage II, highlighting the prognostic significance of this stage. (D) The survival curves for KIRP, stratified by high-score and low-score samples, are contrasted to demonstrate the impact of DNFE scores on patient outcomes. High-score samples (red curves) suggest a poorer prognosis, while low-score samples (light blue curves) indicate a more favorable survival rate. (E) The dynamic development of the DNB directed network structure for KIRP indicates that early-warning signals before critical transitions can be identified at Stage II. (F) The gene expression analysis of DNBs is presented. (G) In KIRP, the average gene expression of DNBs fails to differentiate the tipping point from other points; however, the DNFE method has the capacity to accurately identify the tipping points.

<https://doi.org/10.1371/journal.pcbi.1013336.g005>

when juxtaposed with the low-score group. This post-critical state high-score group is characterized by an elevated risk of disease progression and an unfavorable prognosis. Fig 5E delineates the DNB’s molecular network, which is modulated by specific genes, revealing a distinct architecture for Stage II as opposed to other stages. Collectively, these findings underscore the pivotal nature of Stage II as a critical juncture in the disease trajectory. Furthermore, as shown in Fig 5F and 5G, utilizing DNFE scores derived from DNB genes, we were able to delineate the critical state from other stages with precision (Fig 5A), while merely using the gene expression levels of DNB genes fails.

Revealing functional roles of DNBs and “dark genes” in the progression of KIRP

Moreover, the enrichment analysis of DNBs has substantiated their pivotal role in the course of KIRP progression (Fig 6A). At the critical state, these biomarkers are notably enriched in signaling pathways that are intricately linked to the pathogenesis of the disease. Specifically, DNBs are found to be concentrated in pathways, such as the PI3K–Akt signaling pathway, the MAPK signaling pathway, etc. Fig 6B illustrates the fundamental mechanism elucidated through the functional analysis of “dark genes” and their first-order neighbors. It should be noted that there exists a signaling cascade responsive to “dark genes” in both the MAPK and PI3K/Akt signaling pathways that is crucial for cell proliferation. In the MAPK signaling pathway, the recognized “dark gene” *MAPK9* serves as a pivotal gene within the c-Jun N-terminal kinase subclass pathway, potentially triggering various upstream signals that promote cell reproduction and differentiation [50].

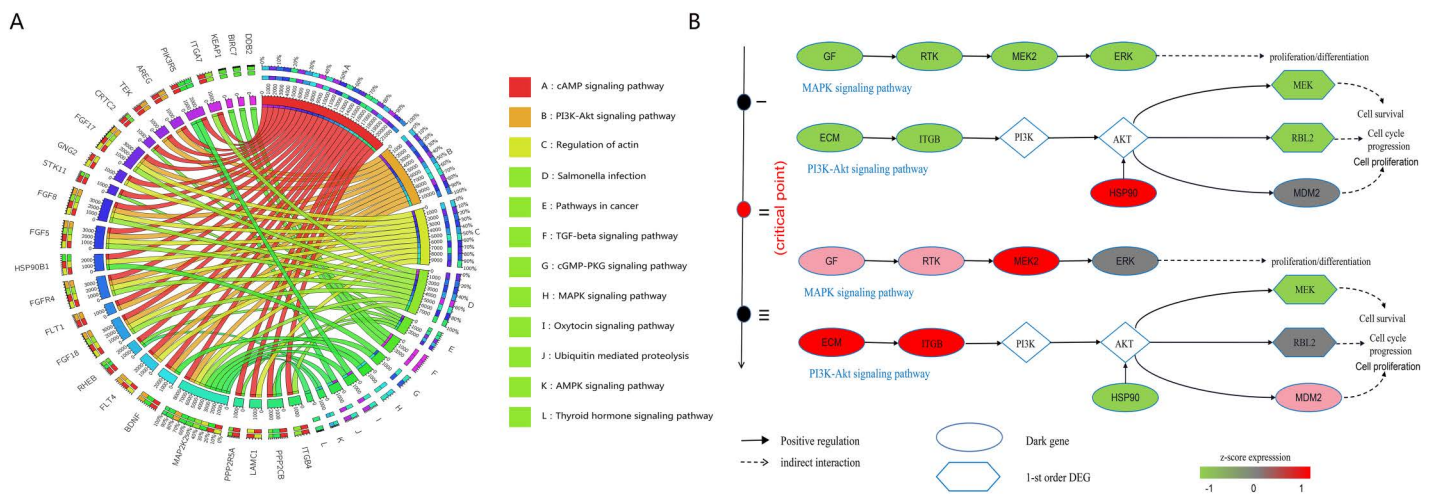


Fig 6. Directed network flow entropy (DNFE) signal biomarkers are involved in important biological processes for kidney renal papillary cell carcinoma (KIRP). (A) Dynamic network biomarkers (DNBs) in KIRP are integral to various biological processes and are implicated in key Kyoto Encyclopedia of Genes and Genomes pathways. (B) Transition dynamics of differentially expressed genes, particularly those before and after the critical state, are influenced by potential “dark genes”. “Dark genes” are a unique class of genes that, despite not showing differential expression at the traditional significance threshold ($P < 0.05$; t-test), exhibit an evident change in their DNFE values ($P < 0.05$).

<https://doi.org/10.1371/journal.pcbi.1013336.g006>

In the PIK3/Akt signaling pathway, an increase in *IGF1R*, *SOS1*, and *SOS2* levels leads to the activation of *RAS* and subsequently activates *RPS6KA3*, for which the cascade may contribute to mitogenic effects [51]. The downstream responses triggered by “dark genes” are closely linked to cell reproduction and differentiation processes. The persistent expression of relevant genes within these pathways from Stage I through Stage III is crucial for enhancing both proliferation and differentiation. This aligns with the literature findings [31] suggesting that the identified tipping point could serve as an important milestone in directing pluripotent stem cells towards the definite endoderm.

Predicating prognostic effects of “dark genes” for KIRP

As shown in Fig 7, *AKT1S1*, *NDRG1*, *VWF*, *ABCF3*, *TBC1D4*, and *PACSIN1* have been identified as “dark genes”. Upon reaching the tipping point, the DNFE scores of “dark genes” exhibit heightened sensitivity and evident upward trends leading up to the critical state of tumor deterioration, as opposed to the gene expression data. This trend is more evident than the changes observed in gene expression data. Furthermore, “dark genes” have been shown to be crucial for cancer prognoses. To assess the prognostic value of “dark genes”, we categorized all samples into high-score group and low-score group, which are established based on gene expression and DNFE scores. Fig 7 illustrates that

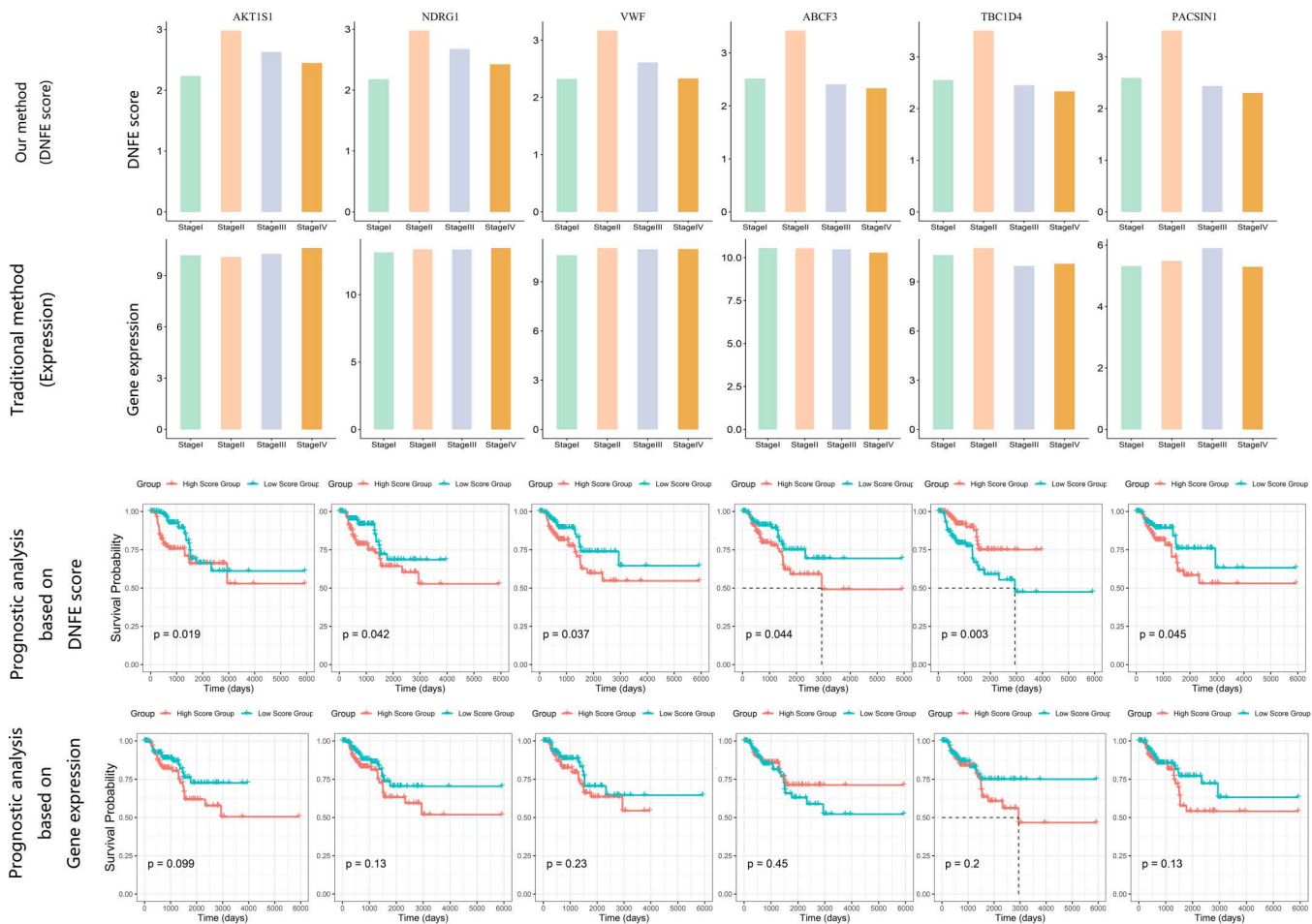


Fig 7. Prognostic comparative analysis between gene expression and directed network flow entropy (DNFE) scores of six “dark genes” for kidney renal papillary cell carcinoma.

<https://doi.org/10.1371/journal.pcbi.1013336.g007>

for the “dark gene” *VWF*, there is a noticeable elevation in the DNFE score at the tipping point of KIRP, in contrast to the relatively stable gene expression. Furthermore, a stratified analysis of the DNFE score and the gene expression of the “dark gene” *VWF* revealed prognostic significances of 0.037 and 0.23, respectively. Statistically, this indicates that the subgroups defined by the DNFE score can effectively segregate KIRP patients into groups with varying survival outcomes, a distinction that gene expression alone fails to achieve, underscoring the DNFE score’s utility in the early diagnosis of diseases. Similar findings were observed for other “dark genes”, further validating the DNFE score’s efficacy in disease prognosis.

Moreover, these “dark genes” hold a key position in the progression of cancer. For example, the downregulated expression of *VWF* in tumor tissues is related to *ERS* model genes, which implies that it may have a potential essential function in cancer progression [52]. The EGFR-mediated signaling pathways through *PTPN11* (*SHP2*) and *AKT1S1* (*PRAS40*), along with the enhanced anti-apoptotic signaling pathways resulting from acquired resistance to cetuximab, provide multiple opportunities for identifying and validating biomarkers, signaling pathways, and resistance mechanisms for clinical improvements in cancer treatment [53]. Cell proliferation assays, colony formation assays, and cell migration experiments indicate that overexpressing circRNA-*TBC1D4* aids in the migration of neuroblastoma (NB) cells but does not enhance in vitro proliferation or colony formation. These results suggest that circRNA-*TBC1D4* is generally downregulated in NB cells and may serve as a negative biomarker for this type of cancer [54].

Comparative analysis of prognostic outcomes based on gene expression and DNFE scores for the “dark genes” *AKT1S1*, *NDRG1*, *VWF*, *ABCF3*, *TBC1D4*, and *PACIN1* revealed that DNFE scores have greater sensitivity to tipping points before disease deterioration and provide superior prognostic predictions compared to gene expression levels. The prognosis associated with these “dark genes” indicates apparent differences in survival times between the two sample groups—those with high DNFE scores and those with low DNFE scores—highlighting the greater relevance of DNFE scores over gene expression in predicting outcomes.

Detecting changes in immune signatures for SLE-NC

DNFE not only predicts the critical state but also accurately detects changes in immune signatures in vivo. To demonstrate the efficiency of the proposed method based on a body fluid dataset, we utilized the DNFE method with a blood dataset of patients with a non-complicated SLE pregnancy (SLE-NC).

For neutrophils, the DNFE score for marker genes (Fig 8A) and the DNFE score for DNBs (Fig 8B) both increased from P1, reaching a maximum at P3, and subsequently decreased below P1 levels at PP, which is consistent with the alterations observed in neutrophils reported in the original manuscript (Fig 8C). For T cells, the DNFE score in marker genes (Fig 8D) and the DNFE score in DNBs (Fig 8E) decreased from P1, reached the minimum at P3, and subsequently increased below P1 levels at PP, which is consistent with the alterations observed in T cells reported in the original manuscript (Fig 8F). Immunological changes in neutrophils and T cells at five time points were detected by applying the algorithm to the SLE-NC data, which aligns with the transcriptional changes surrounding embryo implantation monitored during assisted reproductive technology, thereby demonstrating the efficacy of DNFE.

Revealing functional roles of DNBs and “dark genes” for SLE-NC

As shown in Fig 8A and 8C, the neutrophil cell signature steadily increased from P1, reaching the highest levels at P3, and decreased below the P1 baseline from P3 after delivery. In the early stages of pregnancy, the increase in estrogen and progesterone may stimulate the production and release of neutrophils in the bone marrow, resulting in an increase in the number of neutrophils. This helps to enhance the immune system’s response capability to protect the mother and the fetus from the risk of infection. In the later stages of pregnancy, as the fetus develops and the mother’s immune system gradually recovers, there may be a decrease in inflammation levels and a balance adjustment in the immune system, leading to a decrease in the number of neutrophils. This adjustment may help to prevent excessive immune responses

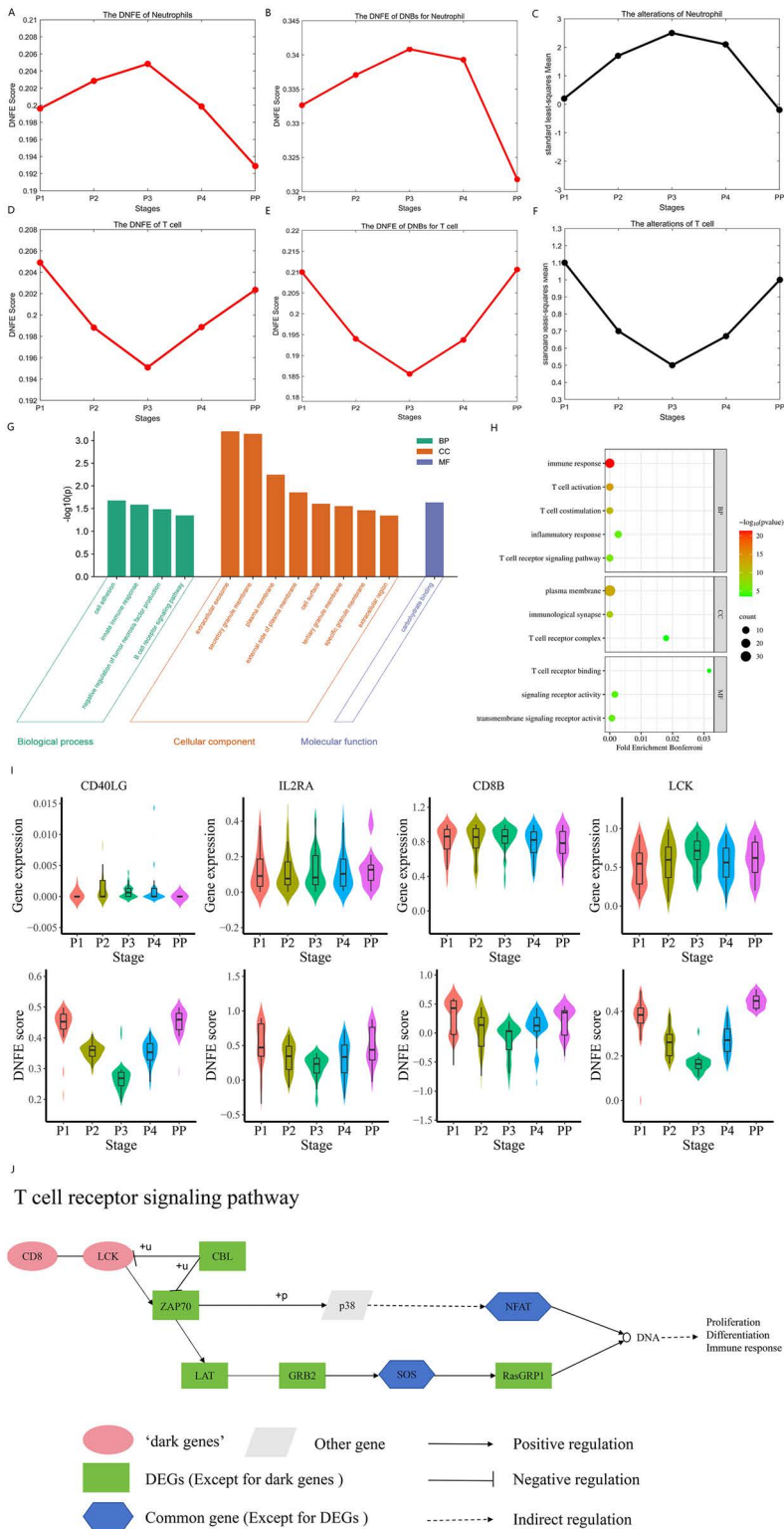


Fig 8. Identification of critical states and functional roles revealed by dynamic network biomarkers (DNBs) and non-differential “dark genes” for SLE-NC. (A) The directed network flow entropy (DNFE) score curve for neutrophils. (B) The DNFE score curve of DNBs for neutrophils. (C) The alterations curve for neutrophils in the original paper. (D) The DNFE score curve for T cells. (E) The DNFE score curve of DNBs for T cells. (F) The

alterations curve for T cells in the original paper. (G) A bar plot representing enriched gene ontology (GO) terms, with names and IDs indicated, sorted by the number of associated genes. (H) Dot plot for the enriched GO terms. (I) The comparison of changes in immune signatures in vivo between gene expression and DNFE scores of the “dark genes” *CD40LG*, *IL2RA*, *CD8B*, and *LCK*, for which DNFE scores are more sensitive to changes compared with gene expression data. (J) The regulation of related marker genes in the T cell receptor signaling pathway.

<https://doi.org/10.1371/journal.pcbi.1013336.g008>

while maintaining the immune system in an appropriate functional state to sustain the health of the mother and the fetus [8]. This result supports the notion that the regulatory role of the major transcriptional networks during healthy pregnancies is conserved in SLE pregnancies [55]. The T cell signature steadily decreased from P1, reaching the lowest levels at P3, and increased above the P1 baseline from P3 after delivery (Fig 8D and 8F). In the early stages of pregnancy, the number of T cells decreases to prevent the mother from having an immune response to the fetus. As pregnancy progresses, the fetus develops its own immune system, and the mother’s immune system gradually becomes more active, leading to an increase in T cell numbers to protect both the mother and the fetus from the risk of infection [56]. The findings suggest that the modulation of this pathway is essential for achieving successful pregnancies in SLE [55]. Immunological changes in neutrophils and T cells at five time points were detected by applying the algorithm to the SLE-NC data, which is not only consistent with the transcriptional alterations observed in the vicinity of embryo implantation monitored using assisted reproductive technology but also illustrates the changes in cells from a biological perspective, demonstrating the efficacy of DNFE.

Subsequently, we conducted GO enrichment analysis on the marker genes that may hold important biological relevance in SLE-NC. Some genes were found to be enriched in biological processes related to immunity or defense, e.g., the “innate immune response” (GO:0045087), “B cell receptor signaling pathway” (GO:0050853), “immune response” (GO:0006955), and others in the Gene Ontology analysis (Fig 8G and 8H). The onset of SLE is impacted by genetic susceptibility and dysregulation of the immune response [57]. The pathway involving T cell receptor signaling is vital for the immune system, regulating the activation, proliferation, and function of T cells. For pregnant women with SLE, the balance of the immune system is more fragile, which may lead to exacerbated autoimmune reactions and disease exacerbation. The abnormal activation or dysregulation of the T cell receptor signaling pathway is likely to be associated with the development and clinical manifestations of SLE during pregnancy.

To delve deeper into the elusive elements within the regulatory mechanisms of SLE-associated pathways, we conducted a comparative analysis between the DNB genes and DEGs. The investigation revealed the presence of certain genes within the DNBs that, while not exhibiting differential expression at the molecular level, have a notably high DNFE score at the network level. These genes, which elude traditional detection methods, are referred to as “dark genes”. For T cells, as shown in Fig 8I, *CD40LG*, *IL2RA*, *CD8B*, and *LCK* were found to be “dark genes”. DNFE scores reflect the trends in T cells better than those based on gene expression data.

To provide greater clarity on the pathway associations of T cell genes, we concentrated on analyzing the KEGG pathways related to the T cell receptor signaling pathway that are most relevant to SLE progression (Fig 8J). *CD8*, identified as a “dark gene”, is implicated in the T cell receptor signaling pathway, which was a key regulator. *CD8* contributes apparently to the immune response, T cell activation, and regulation of the immune response, etc. The elevated levels of *ICOS* and *CXCR5* on *CD8* T cells from SLE patients suggest a pathogenic *CD8* T cell subset exhibiting amplified cytotoxicity and potentially facilitating a stronger antibody response through B cell activation [58].

Discussion

The ability to discern pivotal transitional phases within biological phenomena, such as recognition of the prodromal phase preceding tumorigenesis and the critical junctures of cell lineage specification that occur during embryogenesis, holds great significance. The recognition of indicators that precede the onset of pathological conditions is vital for timely interventions. However, the characterization of biological system dynamics and accurate detection of tipping points or critical

states from datasets is challenging due to the resemblance between pre-transition states and critical states concerning the phenotype and average gene expression. In this paper, we introduce an innovative approach that utilizes directed network analysis to identify early-warning signals in biological processes, marking a departure from conventional techniques that focus solely on differential gene expression data. This method leverages the topological properties of networks to uncover subtle yet obvious changes that may precede critical transitions in biological systems. Our results indicate that the proposed DNFE method can achieve effective results in detecting the critical states of six real datasets, identifying DNBs, and revealing the intrinsic molecular mechanisms occurring during the progression of the disease. This method also enhances our understanding of the regulatory relationships between genes. Additionally, we found that “dark genes” from three embryonic differentiation datasets, two cancer datasets, and a blood dataset have an important influence on the progression of biological processes or pathways. Numerical simulations demonstrated the efficiency and resilience of the DNFE method compared to existing approaches, as well as its ability to handle vast amounts of data or large-scale datasets.

The methods [59,60] of criticality and tipping points used pseudotime analysis to simulate the trajectory of cell differentiation. While this approach achieves a continuous control variable, it does not provide actual time points. As a result, they could only make assumptions and use the pseudotime to trace the developmental trajectories of cells. In contrast, the datasets in our manuscript are associated with specific time points or stages, which are discrete-time data, including both time-series data and stage-course cancer data spanning multiple years. Based on dynamic system theory, a biological process—whether characterized by discrete-time data or continuous control variable data—can be described as a dynamical system. For discrete-time data, we employ difference equation models, whereas for continuous control variable data, we utilize differential equation models. And a tipping point is mathematically defined as a bifurcation point of the corresponding dynamical system. A tipping point corresponds a critical state, after which the system state shifts abruptly from one stable state to another attraction including another stable state or periodic oscillation.

In conclusion, we introduce an effective and robust computational method, DNFE, that is competent in identifying critical states from bulk and single-cell data and identifying corresponding DNBs. The DNFE method also shows promising potential for exploring potential molecular mechanisms of disease progression and discovering new network biomarkers and “dark genes”.

DNFE is a model-free and data-driven method. Nevertheless, DNFE also has limitations. To begin, every phase requires an adequate number of samples: a meager sample size can skew the method’s accuracy, while an excessively large sample size can dampen the computational efficacy and waste computational resources. Furthermore, while we primarily corroborated our analytical findings through comparisons with the literature and subsequent analyses, experimental confirmation was not sufficiently addressed in our work and remains an essential next step. TFs, for example, are crucial for understanding cellular functions, developmental biology, and the underpinnings of diseases. We forecasted TFs by pinpointing the “dark genes” with the aid of web-based instruments in this research. Nevertheless, this technique might not be completely precise, and further experimental corroboration is warranted.

Methods

Dataset

DNFE is utilized across six real datasets, including three single-cell expression datasets, a blood dataset, and two tumor datasets. mESC-to-MP, hESC-to-DEC, MEF-to-neuron, and SLE-NC data in this study are downloaded from the NCBI Gene Expression Omnibus (GEO) database (NCBI GEO: GSE79578 for mESC-to-MP data, GSE75748 for hESC-to-DEC data, GSE67310 for MEF-to-neuron data, and GSE108497 for SLE-NC data), which are publicly accessible at <http://www.ncbi.nlm.nih.gov/geo>. KIRP and BLCA data in this study are downloaded from TCGA database and are publicly accessible at <http://cancergenome.nih.gov>.

Data preprocessing and functional analysis

We analyzed the preprocessed gene expression datasets that were measured for SLE-NC. The datasets were obtained from the NCBI GEO database (access ID: GSE108497) and were not raw data offered by Hong et al. [21]. The datasets encompassed 31,621 probes.

Afterward, we processed the downloaded matrices to create the necessary gene expression matrices (GEMs) by converting IDs and removing duplicates and null values. When multiple probes were associated with a single gene, their values were averaged to yield GEMs containing 19,404 genes.

DNFE excels at forecasting critical states and precisely identifying in vivo shifts in immune signatures. Our analysis encompassed 512 samples in total. We utilized samples collected at designated gestational time frames (P1: < 16 weeks gestation [WG]; P2: 16–23 WG; P3: 24–31 WG; P4: 32–40 WG; and 8–20 weeks postpartum [PP]) for the microarray examinations. Pregnant patients with SLE were categorized based on the pregnancy outcomes: those without complications (NC), those with preeclampsia (PE), and those with other fetal-related complications without the presence of PE (OC). We performed immune signal probing assays based on SLE-NC data.

The enrichment analysis of DNBs is based on the Gene Ontology Consortium (<http://geneontology.org>) and DAVID Bioinformatics Resources (<https://david.ncifcrf.gov/>). The TFs are predicted through CHEA3 (<https://maayanlab.cloud/chea3/>). Protein–protein interaction (PPI) networks are drawn based on STRING (<https://string-db.org/>) and the client software Cytoscape (<https://cytoscape.org/>).

Outline of the DNFE algorithm

Given a collection of control cells/samples and a corresponding set of case cells/samples, we constructed a specific directed network. This construction is based on the rewiring of a weighted gene co-expression network analysis, guided by a direction determination index ω_{ij} (Fig 9A). Subsequently, the DNFE is calculated and utilized to detect the early warning signals of critical transitions during the development of complex biological processes (Fig 9B). Throughout the dynamic evolution of these processes, the DNFE score remains low while the system is stable. However, it shows a marked increase as the system nears a critical state. This sudden rise in the DNFE score acts as a signal for impending tipping points in the biological process (Fig 9C). The theoretical background of our model can be seen in S1 File.

Algorithm for disease prediction based on DNFE

DNFE is designed to identify tipping points during the development of complex diseases and identify DNBs. Given a set of reference/control samples that represent a relatively normal state, the subsequent algorithm is formulated to detect the tipping point based on a specific sample. A flow diagram of the DNFE is presented in Fig 1, with detailed procedures outlined in the following subsections.

Construction of the directed network

The WGCNA network is constructed based on reference samples and case samples. Based on the WGCNA network and gene expression data, the global directed network can be established using a direction determination index, which is defined as follows:

$$\omega_{ij} = \sum_{x \in \vec{X}} \sum_{y \in \vec{Y}} p(\hat{x}, y) \log \frac{p(\hat{x}, y)}{p(\hat{x})p(y)} - \sum_{x \in \vec{X}} \sum_{y \in \vec{Y}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (1)$$

Here, vectors \vec{U} and \vec{V} represent the expression profiles of genes g_i and g_j in all samples, respectively; \vec{X} and \vec{Y} are defined as $\vec{X} = (\vec{U} + \vec{V})/\sqrt{2}$, which ensures that information from \vec{U} and \vec{V} has equal weight and prevents the exponential growth of the feature values during the crossover process and improves the numerical stability of the algorithm;

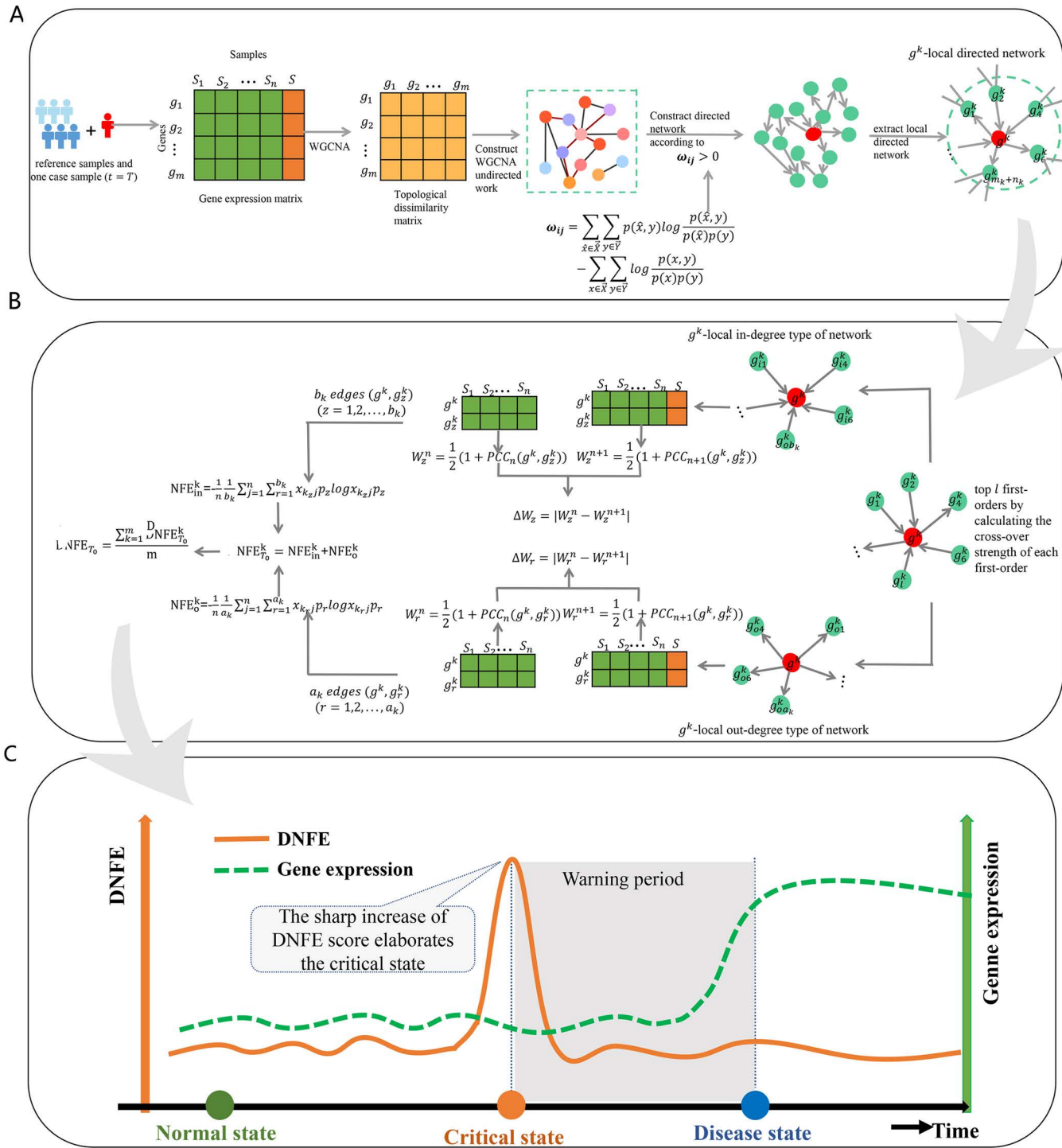


Fig 9. Schematic of the directed network flow entropy (DNFE) algorithm. (A) Given a set of control cells or samples and a corresponding set of case cells/samples at point T, we generated the specific directed network using rewired weighted gene co-expression network analysis with the direction determination index. (B) The global DNFE is calculated to pinpoint the tipping points of complex biological processes at point T. (C) In the dynamic biological progression, the DNFE score is at a low level as the system is in a normal state, but it surges dramatically as the system moves toward the critical state. This sudden rise in the DNFE score signifies the tipping point of the biological process.

<https://doi.org/10.1371/journal.pcbi.1013336.g009>

$\vec{X} = \vec{U}$, \vec{Y} is the phenotype representing the binary vector for each sample (0–1); $p(\hat{x}, y)$ is the joint probability density function (pdf) of \vec{X} and \vec{Y} ; $p(x, y)$ represents the pdf of \vec{X} and \vec{Y} ; and $p(\hat{x}), p(x), p(y)$ represent the edge pdf of $\vec{X}, \vec{X}, \vec{Y}$, respectively. The positive determination value indicates that the integration of the gene g_j is an improvement in the gene g_i 's mutual information; that is, there exists a directed edge (g_i, g_j) from g_i to g_j in the directional network. Explicitly, a directed edge (g_i, g_j) exists from gene g_i to g_j when ω_{ij} is greater than zero; otherwise, no directed edge (g_i, g_j) is present. We constructed the global directed network N^{global} using this approach, where each directed edge (g_i, g_j) from gene g_i to g_j is decided by direction determination index ω_{ij} .

Extraction of the gene module from the global network

Each local gene module $M^k (k = 1, \dots, m)$ is extracted from the global network. In addition, in the analysis of gene modules, the scope is limited to the immediate surroundings of each gene, encompassing only its first-order and second-order neighbors. The local gene module M^k centers on gene $g^k (k = 1, 2, \dots, m)$ and m_k first-order out-degree neighbors $\{g_{o1}^k, \dots, g_{om_k}^k\}$, n_k first-order in-degree neighbors $\{g_{i1}^k, \dots, g_{in_k}^k\}$.

Determination of the local directed network

There are p_j second-order out-degree neighbors $\{g_{jo1}^k, \dots, g_{jop_j}^k\}$ and q_j second-order in-degree neighbors $\{g_{ji1}^k, \dots, g_{jiq_j}^k\}$ for each first-order neighborhood gene $g_j^k (j = 1, 2, \dots, m_k + n_k)$. The crossover strength of each first-order neighborhood gene g_j^k was calculated as follows:

$$D_j = \lambda D_j^{in} + (1 - \lambda) D_j^{out} \quad (2)$$

where λ is a constant, $0 < \lambda < 1$. The choice of λ affects the final calculated gene importance ranking or the results of network structure analysis. In the process of determining the optimal value of λ , we iteratively modified it within our algorithm to observe if the resulting outcomes matched those observed clinically. In this study, the critical thresholds of λ was set to 1/5. The algorithms discussed in this article are implemented under the similarity weight principle, where a higher edge weight correlates with a shorter distance between points, suggesting a closer relationship. Therefore, the in-intensity of the node can reflect the importance of the node. $D_j^{in} = \sum_{i=1}^{q_j} \omega_{ij}$ is the in-strength of the first-order neighborhood genes, and $D_j^{out} = \sum_{i=1}^{p_j} \omega_{ij}$ is the out-strength of the first-order neighborhood genes. ω_{ij} represents the weight of the directed edge (g_i, g_j) , which is defined as follows:

$$\omega_{ij} = \frac{W_{ij}}{\sum_{i=1}^{p_i/q_i} W_{ij}} \quad (3)$$

Where

$$W_{ij} = \frac{(n-1) \Delta PCC(g_i^k, g_j^k)}{1 - (PCC^n(g_i^k, g_j^k))^2} \quad (4)$$

In this manner, the crossover strength of each first-order neighborhood gene g_j^k was ranked in descending order, and the top l first-order neighborhood genes were selected with the central gene g^k to form a local central network $N^k (k = 1, 2, \dots, m)$.

Calculating a local DNFE score for each local directed network

We can obtain the local directed network N^k ($k = 1, 2, \dots, m$), and each local directed network N^k is concentrated around gene g^k , including a_k first-order out-degree neighbors $\{g_{o1}^k, \dots, g_{oa_k}^k\}$ and b_k first-order in-degree neighbors $\{g_{i1}^k, \dots, g_{ib_k}^k\}$, and $a_k + b_k = l$.

For the local directed network (with a_k first-order out-degree neighbors and b_k first-order in-degree neighbors) centered on gene g^k , its relevant local DNFE score at $t = T$ is characterized below:

$$DNFE_T^k = NFE_o^k + NFE_{in}^k \quad (5)$$

where NFE_o^k and NFE_{in}^k are described below:

$$NFE_o^k = -\frac{1}{n} \frac{1}{a_k} \sum_{j=1}^n \sum_{r=1}^{a_k} x_{k,rj} p_r \log x_{k,rj} p_r \quad (6)$$

With

$$P_r = \frac{\Delta W_r}{\sum_{r=1}^{a_k} \Delta W_r} \quad (7)$$

$$\Delta W_r = |W_r^n - W_r^{n+1}| \quad (8)$$

$$W_r = \frac{1}{2}(1 + PCC_{out}(g^k, g_r^k)) \quad (9)$$

and

$$NFE_{in}^k = -\frac{1}{n} \frac{1}{b_k} \sum_{j=1}^n \sum_{z=1}^{b_k} x_{k,zj} p_z \log x_{k,zj} p_z \quad (10)$$

With

$$P_z = \frac{\Delta W_z}{\sum_{z=1}^{b_k} \Delta W_z} \quad (11)$$

$$\Delta W_z = |W_z^n - W_z^{n+1}| \quad (12)$$

$$W_z = \frac{1}{2}(1 + PCC_{in}(g^k, g_z^k)) \quad (13)$$

where W^n and W^{n+1} are the transformed Pearson correlation coefficients (PCCs) of gene expression regarding the center gene g^k and its first-order neighbor g_j^k using n reference samples and $n + 1$ mixed samples (composed of reference samples and one specific sample), respectively. $x_{k,rj}$ and $x_{k,zj}$ are the expression data of gene g_r^k and g_z^k in sample j .

Calculation of the DNFE score of the global directed network

At time point $t = T$, the DNFE score for the perturbed sample is defined as follows:

$$DNFE_T = \frac{\sum_{k=1}^m DNFE_T^k}{m} \quad (14)$$

The DNFE score, assessed at successive time points, serves as a metric of the overarching disruption induced by the case samples within the biological network. We repeat previous steps to calculate the other score at the different time $t = T$. If $DNFE_T$ sharply increases, then T is the tipping point, and the top 5% genes in terms of $DNFE_T$ are DNBs in this work. Otherwise, return to Fig 9A by taking next time point (i.e., $T + 1$) as a candidate critical state. To evaluate the capacity of the CDNFE score in capturing critical state, we applied one-sample t-test ($S = \frac{\sqrt{n}(\text{mean}(\hat{X}) - x)}{SD(\hat{X})}$), and the time point t can be conceived as critical state if the DNFE score $DNFE_T$ meets the following two conditions: (i) $DNFE_T > DNFE_{T-1}$, and (ii) $DNFE_t$ is statistically different ($P < 0.05$) from the prior information. Genes with the top 5% highest scores are classified as DNBs at the critical state. The DNFE algorithm primarily describes the fluctuations within the network, serving as a fundamental tool for quantifying the network's criticality or tipping point. This approach offers a valid and stable early warning signal before critical transitions.

Supporting information

S1 File. Theoretical background.

(PDF)

S2 File. Numerical simulation details.

(PDF)

S3 File. Pinpointing cell fate commitment during the cellular differentiation process.

(PDF)

S4 File. Identifying the critical transition for bladder cancer.

(PDF)

S1 Fig. Model of an eleven-molecule network. This illustration of a molecular network features eleven nodes, with their dynamic regulatory interactions defined based on the stochastic system in Eq. (S1) in S1 File. The positive or negative regulatory connections among the nodes are reflected by the edges.

(JPG)

S2 Fig. Analytical results for mouse embryonic stem cell (mESC)-to-mesoderm progenitor (MP) cell data.

(A) The temporal fluctuations in the directed network flow entropy (DNFE) scores during the mESC-to-MP transition reveal the impending critical state at the 24 h mark. (B) The DNFE landscape for the mESC-to-MP transition demonstrates an apparent surge in the DNFE score at 24 h, underscoring its utility in pinpointing critical developmental stages. (C) The top 20 upstream hub transcription factors exert regulatory control over 89% of the dynamic network biomarker (DNB) genes detected at 24 h, highlighting the pivotal role of these factors in the gene regulatory network. (D) Analysis of Kyoto Encyclopedia of Genes and Genomes pathway enrichment for "dark genes" involved in the mESC-to-MP process. (E) The signaling genes within the PI3K/Akt pathway, which are enriched at the tipping point, exhibit distinct regulatory patterns before and after lymph node metastasis. These genes are instrumental in the progression of cancer.

(JPG)

S3 Fig. Analysis results for mouse embryonic fibroblast (MEF)-to-neuron data. (A) The trajectory of the directed network flow entropy (DNFE) score throughout the MEF-to-neuron reprogramming timeline is marked by notable escalation at day 20, signifying a tipping point that heralds a critical phase in the cellular transition. (B) The DNFE landscape of dynamic network biomarkers (DNBs) during the MEF-to-neuron transition reveals a pronounced uptick in the DNFE score at day 20, which is indicative of a tipping point. (C) The temporal dynamics of DNBs throughout the MEF-to-neuron transition are characterized by a discernible pattern of change, with the critical state detectable at day 20. (D) Top 20 upstream hub transcription factors are able to regulate 89% of the identified DNB genes at day 20. (E) The results of the regulatory association between the top 20 transcription factors are shown as a network. (F) The top five transcription factors (TFs) identified using each library, with these TFs represented along the columns and the query genes along the rows. The entries within this matrix are populated based on whether a query gene is present or absent within the target gene set of a library TF at the second tipping point.

(JPG)

S4 Fig. Analysis results for bladder cancer (BLCA). (A) The fluctuation in the directed network flow entropy (DNFE) score describes a critical transition at Stage II, suggesting a pivotal moment in the disease's trajectory. (B) The DNFE landscape analysis delineates the score's behavior across different stages, providing a visual representation of disease progression. (C) A comparative analysis of survival curves for BLCA reveals a marked disparity in survival times, with an important drop post-stage II. (D) The dynamic development of dynamic network biomarker (DNB) genes is examined, highlighting the changes in these genes' behavior that coincide with the critical stage. (E) The gene expression of DNBs. (F) The mean gene expression of DNBs is assessed; while it is not sufficient to discern critical states from others, it is complemented by the DNFE method's ability to detect such stages with greater precision.

(JPG)

S1 Table. The comparison between the DNFE algorithm and other GRN methods.

(PDF)

S2 Table. Detailed hallmark pathway enrichment analysis information.

(PDF)

Author contributions

Conceptualization: Xueqing Peng, Rui Qiao, Peiluan Li, Luonan Chen.

Formal analysis: Xueqing Peng.

Funding acquisition: Peiluan Li, Luonan Chen.

Methodology: Xueqing Peng, Rui Qiao, Peiluan Li.

Software: Rui Qiao.

Supervision: Peiluan Li, Luonan Chen.

Validation: Xueqing Peng.

Visualization: Xueqing Peng, Rui Qiao.

Writing – original draft: Xueqing Peng, Rui Qiao, Peiluan Li.

Writing – review & editing: Xueqing Peng, Peiluan Li, Luonan Chen.

References

1. Chen L, Liu R, Liu Z-P, Li M, Aihara K. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci Rep.* 2012;2:342. <https://doi.org/10.1038/srep00342> PMID: [22461973](https://pubmed.ncbi.nlm.nih.gov/22461973/)

2. Saelens W, Cannoodt R, Todorov H, Saeyns Y. A comparison of single-cell trajectory inference methods. *Nat Biotechnol.* 2019;37(5):547–54. <https://doi.org/10.1038/s41587-019-0071-9> PMID: [30936559](https://pubmed.ncbi.nlm.nih.gov/30936559/)
3. Yuan Y, Bar-Joseph Z. Deep learning for inferring gene relationships from single-cell expression data. *Proc Natl Acad Sci U S A.* 2019;116(52):27151–8. <https://doi.org/10.1073/pnas.1911536116> PMID: [31822622](https://pubmed.ncbi.nlm.nih.gov/31822622/)
4. Abdelaal T, Michielsen L, Cats D, Hoogduin D, Mei H, Reinders MJT, et al. A comparison of automatic cell identification methods for single-cell RNA sequencing data. *Genome Biol.* 2019;20(1):194. <https://doi.org/10.1186/s13059-019-1795-z> PMID: [31500660](https://pubmed.ncbi.nlm.nih.gov/31500660/)
5. Zhong J, Han C, Zhang X, Chen P, Liu R. scGET: Predicting cell fate transition during early embryonic development by single-cell graph entropy. *Genomics Proteomics Bioinformatics.* 2021;19(3):461–74. <https://doi.org/10.1016/j.gpb.2020.11.008> PMID: [34954425](https://pubmed.ncbi.nlm.nih.gov/34954425/)
6. Hong R, Tong Y, Tang H, Zeng T, Liu R. eMCI: An Explainable multimodal correlation integration model for unveiling spatial transcriptomics and intercellular signaling. *Research (Wash D C).* 2024;7:0522. <https://doi.org/10.34133/research.0522> PMID: [39494219](https://pubmed.ncbi.nlm.nih.gov/39494219/)
7. Zhong J, Tang H, Huang Z, Chai H, Ling F, Chen P, et al. Uncovering the pre-deterioration state during disease progression based on Sample-Specific Causality Network Entropy (SCNE). *Research (Wash D C).* 2024;7:0368. <https://doi.org/10.34133/research.0368> PMID: [38716473](https://pubmed.ncbi.nlm.nih.gov/38716473/)
8. Zhong J, Han C, Chen P, Liu R. SGAE: single-cell gene association entropy for revealing critical states of cell transitions during embryonic development. *Brief Bioinform.* 2023;24(6):bbad366. <https://doi.org/10.1093/bib/bbad366> PMID: [37833841](https://pubmed.ncbi.nlm.nih.gov/37833841/)
9. Guo W-F, Yu X, Shi Q-Q, Liang J, Zhang S-W, Zeng T. Performance assessment of sample-specific network control methods for bulk and single-cell biological data analysis. *PLoS Comput Biol.* 2021;17(5):e1008962. <https://doi.org/10.1371/journal.pcbi.1008962> PMID: [33956788](https://pubmed.ncbi.nlm.nih.gov/33956788/)
10. Shi J, Aihara K, Chen L. Dynamics-based data science in biology. *Natl Sci Rev.* 2021;8(5):nwab029. <https://doi.org/10.1093/nsr/nwab029> PMID: [34691649](https://pubmed.ncbi.nlm.nih.gov/34691649/)
11. Liu R, Aihara K, Chen L. Dynamical network biomarkers for identifying critical transitions and their driving networks of biologic processes. *Quant Biol.* 2013;1(2):105–14. <https://doi.org/10.1007/s40484-013-0008-0>
12. Gao R, Yan J, Li P, Chen L. Detecting the critical states during disease development based on temporal network flow entropy. *Brief Bioinform.* 2022;23:bbac164.
13. Chen P, Liu R, Chen L, Aihara K. Identifying critical differentiation state of MCF-7 cells for breast cancer by dynamical network biomarkers. *Front Genet.* 2015;6:252. <https://doi.org/10.3389/fgene.2015.00252> PMID: [26284108](https://pubmed.ncbi.nlm.nih.gov/26284108/)
14. Yan JL, Li PL, Li Y, Gao R, Bi C, Chen LN. Disease Prediction by Network Information Gain on a Single Sample Basis. *Fundamental Research.* 2023;n.pag.
15. Zhong J, Han C, Wang Y, Chen P, Liu R. Identifying the critical state of complex biological systems by the directed-network rank score method. *Bioinformatics.* 2022;38(24):5398–405. <https://doi.org/10.1093/bioinformatics/btac707> PMID: [36282843](https://pubmed.ncbi.nlm.nih.gov/36282843/)
16. Liu X, Chang X, Leng S, Tang H, Aihara K, Chen L. Detection for disease tipping points by landscape dynamic network biomarkers. *Natl Sci Rev.* 2019;6(4):775–85. <https://doi.org/10.1093/nsr/nwy162> PMID: [34691933](https://pubmed.ncbi.nlm.nih.gov/34691933/)
17. Badia-i-Mompel P, Wessels L, Müller-Dott S. Gene regulatory network inference in the era of single-cell multi-omics. *Nat Rev Genet.* 2023;24:739–54.
18. Xu Q, Georgiou G, Frölich S, van der Sande M, Veenstra GJC, Zhou H, et al. ANANSE: an enhancer network-based computational approach for predicting key transcription factors in cell fate determination. *Nucleic Acids Res.* 2021;49(14):7966–85. <https://doi.org/10.1093/nar/gkab598> PMID: [34244796](https://pubmed.ncbi.nlm.nih.gov/34244796/)
19. Ma A, Wang X, Li J, Wang C, Xiao T, Liu Y, et al. Single-cell biological network inference using a heterogeneous graph transformer. *Nat Commun.* 2023;14(1):964. <https://doi.org/10.1038/s41467-023-36559-0> PMID: [36810839](https://pubmed.ncbi.nlm.nih.gov/36810839/)
20. Yan J, Li P, Gao R, Li Y, Chen L. Identifying Critical States of Complex Diseases by Single-Sample Jensen-Shannon Divergence. *Front Oncol.* 2021;11:684781. <https://doi.org/10.3389/fonc.2021.684781> PMID: [34150649](https://pubmed.ncbi.nlm.nih.gov/34150649/)
21. Liu R, Chen P, Chen L. Single-sample landscape entropy reveals the imminent phase transition during disease progression. *Bioinformatics.* 2020;36(5):1522–32. <https://doi.org/10.1093/bioinformatics/btz758> PMID: [31598632](https://pubmed.ncbi.nlm.nih.gov/31598632/)
22. Zhong J, Liu R, Chen P. Identifying critical state of complex diseases by single-sample Kullback-Leibler divergence. *BMC Genomics.* 2020;21(1):87. <https://doi.org/10.1186/s12864-020-6490-7> PMID: [31992202](https://pubmed.ncbi.nlm.nih.gov/31992202/)
23. Liu R, Zhong J, Yu X, Li Y, Chen P. Identifying critical state of complex diseases by single-sample-based hidden markov model. *Front Genet.* 2019;10:285. <https://doi.org/10.3389/fgene.2019.00285> PMID: [31019526](https://pubmed.ncbi.nlm.nih.gov/31019526/)
24. Chen P, Li Y, Liu X, Liu R, Chen L. Detecting the tipping points in a three-state model of complex diseases by temporal differential networks. *J Transl Med.* 2017;15(1):217. <https://doi.org/10.1186/s12967-017-1320-7> PMID: [29073904](https://pubmed.ncbi.nlm.nih.gov/29073904/)
25. Foo M, Kim J, Bates DG. Modelling and control of gene regulatory networks for perturbation mitigation. *IEEE/ACM Trans Comput Biol Bioinform.* 2019;16(2):583–95. <https://doi.org/10.1109/TCBB.2017.2771775> PMID: [29994499](https://pubmed.ncbi.nlm.nih.gov/29994499/)
26. Khanin R, Vinciotti V, Mersinias V, Smith CP, Wit E. Statistical reconstruction of transcription factor activity using Michaelis-Menten kinetics. *Biometrics.* 2007;63(3):816–23. <https://doi.org/10.1111/j.1541-0420.2007.00757.x> PMID: [17825013](https://pubmed.ncbi.nlm.nih.gov/17825013/)
27. Ronen M, Rosenberg R, Shraiman BI, Alon U. Assigning numbers to the arrows: parameterizing a gene regulation network by using accurate expression kinetics. *Proc Natl Acad Sci U S A.* 2002;99(16):10555–60. <https://doi.org/10.1073/pnas.152046799> PMID: [12145321](https://pubmed.ncbi.nlm.nih.gov/12145321/)

28. Sueyoshi C, Naka T. Stability analysis for the cellular signaling systems composed of two phosphorylation-dephosphorylation cyclic reactions. *Mathematics*. 2017;7:33–45.
29. Chen LN, Wang R, Li C, Aihara K. *Modeling biomolecular networks in cells: Structures and dynamics*. 1st ed. London: Springer; 2010.
30. Chen LN, Wang R, Zhang X. *Biomolecular networks: methods and applications in systems biology*. 1st ed. New Jersey: Wiley, John and Sons Ltd. 2009.
31. Chu L-F, Leng N, Zhang J, Hou Z, Mamott D, Vereide DT, et al. Single-cell RNA-seq reveals novel regulators of human embryonic stem cell differentiation to definitive endoderm. *Genome Biol*. 2016;17(1):173. <https://doi.org/10.1186/s13059-016-1033-x> PMID: 27534536
32. Milazzo G, Perini G, Giorgi FM. Single-Cell Sequencing Identifies Master Regulators Affected by Panobinostat in Neuroblastoma Cells. *Genes (Basel)*. 2022;13(12):2240. <https://doi.org/10.3390/genes13122240> PMID: 36553506
33. Xia Y, Qadota H, Wang Z-H, Liu P, Liu X, Ye KX, et al. Neuronal C/EBP β /AEP pathway shortens life span via selective GABAergic neuronal degeneration by FOXO repression. *Sci Adv*. 2022;8(13):eabj8658. <https://doi.org/10.1126/sciadv.abj8658> PMID: 35353567
34. Hu Y, Su Y, He Y, Liu W, Xiao B. Arginine methyltransferase PRMT3 promote tumorigenesis through regulating c-MYC stabilization in colorectal cancer. *Gene*. 2021;791:145718.
35. Fang F, Xia N, Angulo B, Carey J, Cady Z, Durruthy-Durruthy J, et al. A distinct isoform of ZNF207 controls self-renewal and pluripotency of human embryonic stem cells. *Nat Commun*. 2018;9(1):4384. <https://doi.org/10.1038/s41467-018-06908-5> PMID: 30349051
36. Wang J, Zhang Q, Zhu Q, Liu C, Nan X, Wang F, et al. Identification of methylation-driven genes related to prognosis in clear-cell renal cell carcinoma. *J Cell Physiol*. 2020;235(2):1296–308. <https://doi.org/10.1002/jcp.29046> PMID: 31273792
37. Zhou Z, Gong Q, Lin Z, Wang Y, Li M, Wang L, et al. Emerging Roles of SRSF3 as a Therapeutic Target for Cancer. *Front Oncol*. 2020;10:577636. <https://doi.org/10.3389/fonc.2020.577636> PMID: 33072610
38. Liu M, Yang L, Liu X, Nie Z, Zhang X, Lu Y, et al. HNRNPH1 Is a Novel Regulator Of Cellular Proliferation and Disease Progression in Chronic Myeloid Leukemia. *Front Oncol*. 2021;11:682859. <https://doi.org/10.3389/fonc.2021.682859> PMID: 34295818
39. Johnson JM, Edwards S, Shoemaker D, Schadt EE. Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. *Trends Genet*. 2005;21(2):93–102. <https://doi.org/10.1016/j.tig.2004.12.009> PMID: 15661355
40. Bose S, Banerjee S, Mondal A, Chakraborty U, Pumarol J, Croley CR, et al. Targeting the JAK/STAT Signaling Pathway Using Phytocompounds for Cancer Prevention and Therapy. *Cells* 2020;9:1451.
41. Martini M, De Santis MC, Braccini L, Gulluni F, Hirsch E. PI3K/AKT signaling pathway and cancer: an updated review. *Ann Med*. 2014;46(6):372–83. <https://doi.org/10.3109/07853890.2014.912836> PMID: 24897931
42. Nakayama KI, Nakayama K. Regulation of the cell cycle by SCF-type ubiquitin ligases. *Semin Cell Dev Biol*. 2005;16(3):323–33. <https://doi.org/10.1016/j.semcdb.2005.02.010> PMID: 15840441
43. Dai Y, Qiang W, Yu X, Cai S, Lin K, Xie L, et al. Guizhi Fuling Decoction inhibiting the PI3K and MAPK pathways in breast cancer cells revealed by HTS2 technology and systems pharmacology. *Comput Struct Biotechnol J*. 2020;18:1121–36. <https://doi.org/10.1016/j.csbj.2020.05.004> PMID: 32489526
44. Hanahan D. Hallmarks of cancer: new dimensions. *Cancer discovery*. 2022;12:31–46.
45. Schaefer F, Chen Y, Tsao T, Nouri P, Rabkin R. Impaired JAK-STAT signal transduction contributes to growth hormone resistance in chronic uremia. *J Clin Invest*. 2001;108:467–75.
46. Darnell JE Jr, Kerr IM, Stark GR. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science*. 1994;264(5164):1415–21. <https://doi.org/10.1126/science.8197455> PMID: 8197455
47. Schmidt D, Müller S. PIAS/SUMO: new partners in transcriptional regulation. *Cell Mol Life Sci*. 2003;60(12):2561–74. <https://doi.org/10.1007/s00018-003-3129-1> PMID: 14685683
48. Meyerson M, Harlow E. Identification of G1 kinase activity for cdk6, a novel cyclin D partner. *Mol Cell Biol*. 1994;14(3):2077–86. <https://doi.org/10.1128/mcb.14.3.2077-2086.1994> PMID: 8114739
49. Zhong J, Ding D, Liu J, Liu R, Chen P. SPNE: sample-perturbed network entropy for revealing critical states of complex biological systems. *Brief Bioinform*. 2023;24(2):bbad028. <https://doi.org/10.1093/bib/bbad028> PMID: 36705581
50. Zhang W, Liu HT. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Res*. 2002;12(1):9–18. <https://doi.org/10.1038/sj.cr.7290105> PMID: 11942415
51. Richardson CJ, Schalm SS, Blenis J. PI3-kinase and TOR: PIKTORing cell growth. *Semin Cell Dev Biol*. 2004;15(2):147–59. <https://doi.org/10.1016/j.semcdb.2003.12.023> PMID: 15209374
52. Sadler B, Christopherson PA, Haller G, Montgomery RR, Di Paola J. von Willebrand factor antigen levels are associated with burden of rare non-synonymous variants in the VWF gene. *Blood*. 2021;137(23):3277–83. <https://doi.org/10.1182/blood.2020099999> PMID: 33556167
53. Wang C, Chen Y, Chen K, Zhang L. Long noncoding RNA LINC01134 promotes hepatocellular carcinoma metastasis via activating AKT1S1 and NF- κ B signaling. *Front Cell Dev Biol*. 2020;8:429.
54. Eickelschulte S, Hartwig S, Leiser B, Lehr S, Joschko V, Chokkalingam M, et al. AKT/AMPK-mediated phosphorylation of TBC1D4 disrupts the interaction with insulin-regulated aminopeptidase. *J Biol Chem*. 2021;296:100637. <https://doi.org/10.1016/j.jbc.2021.100637> PMID: 33872597

55. Piccinni M-P. T cells in pregnancy. *Chem Immunol Allergy*. 2005;89:3–9. <https://doi.org/10.1159/000087904> PMID: [16129948](https://pubmed.ncbi.nlm.nih.gov/16129948/)
56. Kamil Alhassbalawi N, Zare Ebrahimabad M, Seyedhosseini FS, Bagheri Y, Abdollahi N, Nazari A, et al. Circulating miR-21 Overexpression Correlates with PDCD4 and IL-10 in Systemic Lupus Erythematosus (SLE): A Promising Diagnostic and Prognostic Biomarker. *Rep Biochem Mol Biol*. 2023;12(2):220–32. <https://doi.org/10.61186/rbmb.12.2.220> PMID: [38317820](https://pubmed.ncbi.nlm.nih.gov/38317820/)
57. Sengupta S, Bhattacharya G, Mohanty S, Shaw SK, Jogdand GM, Jha R. IL-21, Inflammatory Cytokines and Hyperpolarized CD8 T Cells Are Central Players in Lupus Immune Pathology. *Antioxidants*. 2023;12:181.
58. Hong S, Banchereau R, Maslow B-SL, Guerra MM, Cardenas J, Baisch J, et al. Longitudinal profiling of human blood transcriptome in healthy and lupus pregnancy. *J Exp Med*. 2019;216(5):1154–69. <https://doi.org/10.1084/jem.20190185> PMID: [30962246](https://pubmed.ncbi.nlm.nih.gov/30962246/)
59. Yang XH, Goldstein A, Sun Y, Wang Z, Wei M, Moskowitz IP, et al. Detecting critical transition signals from single-cell transcriptomes to infer lineage-determining transcription factors. *Nucleic Acids Res*. 2022;50(16):e91. <https://doi.org/10.1093/nar/gkac452> PMID: [35640613](https://pubmed.ncbi.nlm.nih.gov/35640613/)
60. Barcenas M, Bocci F, Nie Q. Tipping points in epithelial-mesenchymal lineages from single-cell transcriptomics data. *Biophys J*. 2024;123(17):2849–59. <https://doi.org/10.1016/j.bpj.2024.03.021> PMID: [38504523](https://pubmed.ncbi.nlm.nih.gov/38504523/)