Quantifying landscape-flux via single-cell transcriptomics uncovers the underlying mechanism of cell cycle

Ligang Zhu[1,2], Jin Wang[3*]

[1]College of Physics, Jilin University, Changchun 130021, China.
[2]State Key Laboratory of Electroanalytical Chemistry, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun 130022, China.
[3]Department of Chemistry, Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA.

*Corresponding author

Correspondence to: Jin Wang. Email: jin.wang.1@stonybrook.edu.

Summary

Recent developments of single-cell sequencing technology enabling acquisition of the whole transcriptome data. To uncover the underlying mechanism of cell cycle function from such data, we reconstruct a continuous vector field based on the discrete single-cell RNA velocity to quantify the global non-equilibrium dynamical landscape and flux. We reveal that biological noise can make the global landscape more complex and less predictable. Genetic perturbations alter landscape-flux, thus identify key genes in maintaining cell cycle dynamics and predict the associated effects on cell cycle behaviour. Cell cycle initiation costs energy and sustaining cell cycle requires dissipation to increase oscillatory phase coherence. This approach enables the inference of cell cycle gene regulatory networks directly from single-cell transcriptomic data, including feedback mechanisms. Our study provides a new framework with insights into cell cycle regulation from single-cell transcriptome data and can be extended to other biological processes, such as differentiation-development and disease pathogenesis.

Introduction

The cell cycle is a complex process that controls cellular growth and division, playing a crucial role in the development and maintenance of living organisms[1,2]. An improved understanding of the regulation of the cell cycle is essential for comprehending various biological processes, including cellular differentiation, tissue development, and disease pathogenesis. Integrating statistical mechanics and single-cell biology has provided novel opportunities to study the cell cycle and its regulation[3]. Statistical mechanics offers a powerful framework for investigating complex systems, molecular interactions, and driving forces underlying cellular processes[4]. Contrarily, single-cell biology provides a method to explore cellular processes at the single-cell level, which facilitates the comprehension of the heterogeneity of the cell populations and the mechanisms of cell differentiation[5]. The precise study of the cell cycle is vital for understanding the life of a single cell, which is the fundamental unit of living systems. Conventionally, researchers explore the dynamics of the trajectories or perform local stability analysis around fixed points for the nonlinear system. However, this does not give a global picture or global stability, so people search further for landscapes to characterize the global stability of the system. However, the landscape alone can only quantify the dynamics of the equilibrium system with detailed balance (no net input to or output from), but it cannot describe the whole dynamics of the non-equilibrium system, an additional rotational flux as the driving force is necessary as we pointed out[6]. The transitions between the gene expression states, which determine the driving forces, arise from a non-equilibrium effective potential landscape caused by the steady-state probability and the rotational steady-state probability flux between system states[7]. The steady-state probability flux measures the degree of the non-equilibrium and the irreversible behavior while links dynamic and thermodynamic aspects of the system[8]. However, prior studies modeled the dynamics of cellular signaling networks based on the existing knowledge of gene regulatory networks, necessitating the estimation of several model parameters and network model simplification, deviating significantly from the real biological systems[9,10]. With advancements in single-cell sequencing technology, complete transcriptomic data can be acquired. Several studies constructed the cell fate landscape from the state manifold of the single cell, which did not obtain the underlying mechanism of the cell signaling regulation[3,11-16]. Consequently, studying the underlying mechanism via non-equilibrium dynamics and thermodynamics of transcriptional regulation of cell function directly from single-cell transcriptomic data demands urgent attention.

Our work aims to contribute to the understanding of the cell cycle mechanism via nonequilibrium dynamics and thermodynamics by investigating the landscape-flux quantified from the single-cell transcriptomic data. Single-cell RNA sequencing (scRNA-seq) provides information on the gene expressions in the individual cells and enables us to analyze the gene expressions during the cell cycle. Through assessing the single-cell RNA velocity, reconstructing vector fields, reducing dimensions, mapping these patterns to lower-dimensional spaces, we can visualize the landscape-flux and understand the complex interactions between the different signaling processes that regulate the cell cycle, as well as the dynamic flux and thermodynamic dissipation that drive and maintain the cell cycle processes. We then performed in silico perturbations that enable us to identify key genes for cell cycle regulation and predict the effects of their genetic perturbations on the cell cycle

dynamics and thermodynamics. Calculating the phase diffusion, coherence, and coherence time of the cell cycle oscillations, we quantified their relationships with dynamic flux and thermodynamic consumption. We found that noise mediates phase diffusion and reduces coherence and coherence time, resulting in shorter or even impossible cell cycle maintenance times. However, living cells continuously fight to maintain cell cycle oscillation by consuming energy, leading to a non-equilibrium process where entropy increases. We also found that the dynamic flux plays a role in maintaining the continuous cell cycle, which dynamically drives the cell continuously into the next cell cycle phase and promotes the coherent oscillation (cell cycle forward) to maintain the cell cycle oscillation. Lastly, we constructed a transcription factor regulatory network by identifying differentially-expressed genes between different cell cycle phases and validated the key gene circuit behind cell cycle oscillation dynamics by calculating the Jacobian.

The cell cycle is crucial for cell replication and division, governing cellular proliferation and development, which are the foundation of life. Our study employs a unique combination of experimental high throughout sequencing and theoretical landscape-flux approaches to understand the origin and driving forces of the cell cycle. Computational analysis identifies key genes and regulators critical for normal cellular function and compares them with the experimental data. Our work enhances our understanding of the biological functions of the cell cycle and has implications for maintaining normal cell function, which is crucial in disease prevention and human health. Moreover, the global principles from non-equilibrium dynamics and thermodynamics can be used to pin down the mechanisms in regulating cell cycle processes across organisms and underlying regulatory networks. In summary, our research represents an essential step forward towards comprehending the cell cycle and its regulation, potentially having a significant impact on the field of single-cell biology and beyond.

Results

Quantifying cell cycle non-equilibrium landscape-flux from single-cell transcriptomics data

Traditional systems biology studies have relied on a priori knowledge of biological regulatory networks to model and study cell dynamics; however, this approach is associated with several limitations. One major limitation lies in the simplification of networks, which may lead to an incomplete understanding of complex biological systems[17]. Additionally, obtaining reliable results from these models can be challenging due to the estimation of a large number of parameters, such that they may not accurately reflect the real biological system[10]. Although some studies have constructed a cell fate landscape from the state manifold of single cells, they are unable to obtain the underlying physical mechanism related to cell signaling regulation[13,15,16]. To overcome these limitations, we present a novel approach to studying the underlying mechanism of the cell cycle via non-equilibrium dynamics and thermodynamics by bypassing the prior known regulatory networks and quantifying the cell cycle regulatory dynamics directly from the single-cell data.

We quantified the cell cycle dynamics landscape-flux from the single-cell transcriptomics data, following four primary steps (**Figure 1A**). Firstly, we downloaded and processed the cell cycle transcriptomics data, which provided information on gene expression patterns of individual cells over time for subsequent analysis. Secondly, we estimated RNA velocity from transcriptomics data. Thirdly, we reconstructed the vector field of cell cycle dynamics from the RNA velocity information to visualize the movements of cells throughout the cell cycle. This captured the overall flow of cells through the cell cycle and provided us with a way to visualize the movements of cells through the cell cycle. In the final step, the landscape-flux of global cell cycle dynamics and thermodynamics was quantified. This method included UMAP-based identification of cell cycle phase clusters, calculation of RNA velocity, reconstruction of cell cycle dynamics' vector field, computation of the cell cycle dynamic potential landscape, curl flux of the cell cycle dynamic landscape, and gradient force assessment of the cell cycle dynamic landscape.

To gain a deeper understanding of the underlying structure and flow of cells through the cell cycle, we performed a landscape-flux analysis of the cell cycle dynamics using scRNA-seq data from U2OS cells[18] and scEU-seq data from RPE1 cells[19]. We firstly identified cell cycle phase clusters in U2OS cells by applying the UMAP algorithm[20] to generate a low-dimensional representation of the cells (**Figure 1B**). Cells were then colored based on their cell cycle phase, allowing us to visualize the distribution of the cells in the cell cycle and to observe the heterogeneity of the cell cycle phases. The expressions of the transcription factors in the cells are different in different cell cycle phases, and the gene regulations are also different, so the cells have different cycle states (**Figure S1**). RNA velocity[21,22] was then calculated for each cell to capture the direction and magnitude of changes in the gene expressions (**Figure 1C**), which facilitated visualization of phase portraits and insight into cell cycle progression (**Figure S2**). The vector field of cell cycle dynamics was reconstructed using Dynamo[23], which allows us to reconstruct the function of the continuous and analytic velocity

vector field from sparse, noisy single-cell velocity measurements (**Figure 1D**). At the same time, we performed a differential geometry analysis of the cell cycle based on this vector field. Speed defines the length of the velocity vector in the vector field, that is, the magnitude of the velocity, which represents the magnitude of the force in the Langevin equation (**Figure S3B**). Divergence characterizes the local flux exiting versus entering an infinitesimal region in the expression space - the "outgoingness". The sources (sinks) often have strong positive (negative) divergence. The divergence in the cell cycle vector field represents the degree of convergence, reflecting the distribution of different cell cycle phases (**Figure S3C**). Curl characterizes the infinitesimal rotation of a cell state in the vector field, which is gradient-free and represents the cell cycle state. The U2OS cell cycle velocity vector field is noise-free, so the limit cycle of the cell cycle oscillation dynamics is small, and the curl is larger at the small limit cycle (**Figure S3D**). The acceleration field reveals hotspots of cell states where the velocities change dramatically. When a cell leaves one cycle phase and enters another cycle phase, the acceleration is high. Acceleration represents an early driver of cell cycle regulation (**Figure S3E**). The curvature field reveals hotspots where the velocity abruptly changes direction, such as regions around unstable fixed points where the gene expression changes from activation to repression or vice versa. Genes that contribute strongly to the curvature are drivers that control the cell fate. The curvature reflects the cell cycle phase bifurcation, which is larger in the G0 phase, indicating the switch between the oscillation-limited cycle of the maintenance cycle and the G0 monostable state of the arrest cell cycle (**Figure S3F**). Collectively, these results complementary revealed hotspots of the cell states where velocities change dramatically, highlighting key transition points and driving forces of the cell cycle, including stability and instability. Using the reconstructed function, we also quantified the potential landscape $\mathrm{U}\ (\mathrm{U} = -\ln P_{ss})$ of the global cell cycle dynamics in UMAP, representing the steady state probability distribution P_{ss} of cells at different locations in different cell cycle phases with the shape of an irregular Mexican hat (**Figure 1E**). The Mexican hat-shaped landscape features two ring valleys, and the small ring valley in the inner blue area is a virtual loop, and the real cell cycle oscillation trajectory corresponds to the outer ring valley[24]. The emergence of such two different limit sets is a common phenomenon when performing dimensionality reduction with limit cycles to UMAP plane, which is reminiscent of the Poincar´e-Bendixson theorem in planar dynamics theory[25]. The different phases of the cell cycle correspond to different basins of attractions (Purple balls) along the outer close oscillation ring valley path in the Mexcian landscape. Inside or outside the Mexican hat outer ring, the potential is higher and the probability of the system reaching these areas is low. The basin depth quantifies the duration of the stay in the corresponding phases and the local barrier between these phases or basins corresponds to the check point of the cell cycle[9]. While the landscape attracts the system down to the cell cycle ring valley from the gradient part of the driving force, it is the curl flux from the rotational part of the driving force which drives the coherent oscillation of the cell cycle. Therefore, the Mexican hat shape landscape and the curl flux together guarantee the stability and the robustness of cell cycle oscillation dynamics. Additionally, we quantified the curl flux and negative gradient force of the cell cycle nonequilibrium dynamics landscape in UMAP from the reconstructed function (**Figure 1F** and **1G**). The curl flux indicates the amount of rotational flow in the cell cycle dynamics landscape that drives the cell cycle oscillations along the closed ring valley path, while the negative

gradient force indicates the direction and magnitude of the forces stabilizing the cell cycle phase. If the driving force from the flux is greater than the force from the negative potential gradient, the cell will go across the G0 basin barrier into the G1-S phase and the cell will enter the cycle phase. Otherwise, the cell will return to the G0 basin and arrest in the G0 state. Biologically, this means that the nutrient supply is not large enough, or the DNA is damaged, as in the case of senescent cells[26].

Similarly, we evaluated the total RNA velocity from the scEU-seq of the RPE1 cell cycle, reconstructed the vector field for further differential geometric analysis (**Figure S5**), and quantified the non-equilibrium landscape-flux (**Figure S4**). These findings reveal the presence of the curl flux as another component of the driving force for the nonequilibrium dynamics in additional to the gradient of the cell cycle landscape, providing deeper insights into the mechanisms behind the stability and behavior of intrinsic nonequilibrium cell cycle progression that cannot be discerned by the landscape quantification alone. In fact, the cell cycle is one of the best examples of demonstrating the nonequilibrium dynamics being dictated by both the landscape and the flux. What's more, the creation of the cell cycle non-equilibrium landscape-flux through single-cell transcriptomics data presents a new and powerful tool for the study of cell cycle dynamics, allowing for insights into underlying mechanisms that drive cell cycle progression.

Noise alters the global cell cycle dynamics and thermodynamics

Variability in biomolecular noise or changes in molecular states can affect intracellular gene regulations and networks, which are crucial for the uncertain, multi-destiny behavior of the system[27-29]. To investigate the effects of noise intensity on the underlying mechanism of the cell cycle via nonequilibrium dynamics and thermodynamics, we performed simulations at varying noise intensities and analyzed the resulting global changes in the cell cycle dynamics and thermodynamics. Firstly, we examined changes to the global cell cycle dynamic landscape and show that as the noise intensity increases, the global dynamics landscape of the cell cycle becomes more complex and less predictable. Specifically, with zero noise, the cell cycle progresses very smoothly and predictably (**Figure 2A**), while the ring valley of cell cycle becomes wider as noise intensity increases (**Figure 2B** and **2C**), whereas the ring valley is broken when the noise is high, showing the discrete attraction basins, at which time the cell cycle is broken and the cells arrest in the different phases (**Figure 2D**). Moreover, we found that the barrier height decreased in the G0 phase and oscillation center in the landscape as noise intensity increased (**Figure 2I**). Barrier height of G0 phase is defined as the barrier from the G0 basin to the saddle point (between G0 and G1-S), which quantifies how easily cells can switch from quiescent G0 phase to the cyclic state. The barrier height of the oscillation center is defined as the potential difference between the minimum potential of the landscape and the local maximum inside the limit cycle, which characterizes the global stability of the oscillation system[7]. We also investigated and found that the noise increases the cell cycle period, indicative of the rate at which the cell cycle oscillates and how fast it grows (**Figure 2J** and **S7**). Oscillation amplitude of UMAP1 increases with the noise intensity, resulting in a wider range of the oscillation limit cycle (**Figure 2j, S6** and **S7**). Increased variability or noise effectively makes potential landscape basins shallower, thereby broadening the distribution of

biological variables[30]. By quantifying the trajectory autocorrelation function and the corresponding power spectrum density of the oscillating trajectory for different diffusion coefficient D, we can see that as the noise increases (the stability of the oscillation system decreases), the peak height of the power spectrum density decreases and the distribution of the power spectrum becomes more spread out (**Figure S6**). This indicates that the peak height and the width of the power spectrum density can be a measure of the stability of the oscillation system.

Furthermore, we quantified the curl flux as a major component of the driving force of the cell cycle and the entropy production rate (EPR) as the measure of the thermodynamic cost association to the nonequilibrium curl flux of the cell cycle dynamic landscape to investigate how noise intensity affected the underlying mechanism of the cell cycle via nonequilibrium dynamics and thermodynamics. As shown in Figure 2E-H, increasing noise intensity resulted in a more widely distributed flux, with cells flowing between different phases less predictably. We observed that in the small noise range, the flux decreases as the noise increases, while the cell cycle is still oscillating (**Figure 2K**). Larger flux derived from the energy input or nutrient supply implies a larger cell cycle driving force and subsequently results in a shorter time to complete the cell cycle oscillation (period is smaller). Indeed, the speed of the cell cycle is a hallmark of cancer. However, when the noise is very large, although the flux becomes larger with the increase of noise, the cell cycle has been destroyed. At this point, the flux drive on the cell cycle can no longer resist the destruction of noise. In addition to global dynamics landscape alteration, noise intensity also affects the thermodynamics and energetics of the cell cycle. We found that the EPR of the cell cycle decreases with increasing noise intensity (**Figure 2K**). Furthermore, the energy per period of the cell cycle decreases with increasing noise intensity (**Figure 2L**), indicating a change in the energy requirements of the system. These results imply that the cell cycle becomes less efficient at higher noise levels, suggesting that the system requires energy to maintain the coherent cell cycle oscillations.

Taken together, these results demonstrate that noise intensity alters the underlying mechanism of the cell cycle via global nonequilibrium dynamics and thermodynamics, suggesting that the feedback of intracellular gene interactions plays an important role in maintaining the robustness and stability of the cell cycle in the face of environmental fluctuations. These findings have crucial implications for our understanding of the mechanisms that control cell cycle progression and may have important applications in the development of novel therapeutic strategies for the treatment of diseases associated with cell cycle dysregulation. These changes suggest that noise intensity may play a critical role in regulating the onset and termination of cell cycle phases.

In silico perturbation analysis cell cycle dynamics and thermodynamics under genetic perturbation

Genetic perturbation refers to the alteration of the function of a biological system by external or internal factors, such as environmental stimuli, drug inhibition, and gene knockdown[31]. It is used to study how genes interact with each other and affect cellular processes and phenotypes[32-34]. To investigate the effects of the genetic perturbations on the underlying

mechanism of the cell cycle via nonequilibrium dynamics and thermodynamics, we performed in silico perturbations of different genes in a controlled and systematic manner. Key genes involved in cell cycle regulation, such as cyclin-dependent kinases (CDKs), cyclins, and tumor suppressor genes, were perturbed in this study. Similar to Figure 1, we first estimated the RNA velocity of U2OS cells under different genetic backgrounds by in silico perturbation (**Figure S10**), and then reconstructed the different vector fields based on the perturbed RNA velocity (**Figure S11**). Finally, we quantified the global cell cycle dynamic landscapes under different genetic perturbations by adding noise in the dynamic simulation (**Figure 3A**). We found that perturbing distinct genes resulted in specific changes in the cell cycle landscape and alterations in EPR and flux (**Figure 3B** and **3C**).

Specifically, the negative perturbation of *CDC20* led to an increase in the barrier height of the G2-M phase with higher chance of arrest in this phase, while positive perturbation of *CDC20* led to a deeper basin of the G1 phase with higher chance of arrest in this phase (**Figure 3A**). As the experiment showed, *CDC20* plays a crucial role in cell division, being the target protein in the spindle checkpoint during mitosis[35]. We also found that the negative perturbation of *CDC25C* caused disruption of cell cycle oscillation and led to a deeper well of the G2-M phase leading to possible arrest, albeit less obviously than the effect caused by perturbing CDC20 (**Figure 3A**). Biologically, the downregulation of *CDC25C* induces cell cycle arrest in G2/M phase in response to DNA damage via p53-mediated signal transduction, and its abnormal expression is associated with cancer initiation, development, metastasis, occurrence and poor prognosis[36]. Similarly, genetic perturbations in RPE1 cells are manifested that the negative perturbation of *KPNA2* led to a higher barrier height of the G2-M phase, whereas the positive perturbation of *CDC20* led to a higher barrier height of the M phase ((**Figure S12A**). This finding aligns with experimental evidence: *KPNA2* knockdown downregulated *CCNB2* and *CDK1*, inhibited cell proliferation, and induced cell cycle arrest in the G2-M phase[37]. *CCNE2* is a regulatory subunit of cyclin-dependent kinase 2 (Cdk2) and is thought to control the transition of quiescent cells into the cell cycle[38]. Our genetic perturbations also confirm the cell cycle initiation function of *CCNE2*, as reflected in the deepening of the G0 well under negative perturbation of *CCNE2*, and disappearance of the G0 basin with the appearance of other periodic phases following positive perturbation of *CCNE2* (**Figure S12A**).

Figure 3B shows the EPR of cell cycle nonequilibrium thermodynamics with different genetic perturbations. We observed that different genetic perturbations can cause significant changes in EPR, with some resulting in increased energy consumption and others in decreased consumption. However, the EPR is lower under genetic perturbations than in WT, strong perturbations disrupt cell oscillatory dynamics (**Figure 3B** and **S12B**). Thus, maintaining continuous cell oscillation requires greater energetic costs than arresting in different periodic phases. We also quantified the average flux of cell cycle nonequilibrium dynamics with various genetic perturbations and found that WT cells exhibit the largest curl flux, where the flux as a driving force of the cell cycle is greater than the negative gradient force, and the flux drives the cell cycle to continue moving forward to maintain the coherent oscillation (**Figure 3C** and **S12C**). In contrast, specific genetic perturbations result in a flux less than the negative gradient force, insufficient to drive the cell cycle forward, and causing cells to arrest at certain cell cycle phases.

Overall, the results demonstrate the efficacy of using the reconstructed vector field based on RNA velocity derived from single-cell omics data of the cell cycle to predict how genetic perturbations impact cell cycle behavior. This methodology has the potential for identifying the critical regulatory genes involved in maintaining cell cycle coherent oscillation, such as *CDC25C*, *CCNE2*, *KPNA2*, etc., which may be the new targets in cancer.

The underlying mechanism of cell cycle initiation and termination via non-equilibrium dynamics and thermodynamics

The ability to drive cellular state transitions has gained attention as a promising strategy for disease modelling[39]. The quantified Waddington landscape provides a way of describing and revealing biological pathways of developmental processes[40]. A previous study confirmed that gradient and curl forces control the dynamics of these biological paths on the landscape, rather than following the steepest descent path expected[17]. To investigate the cell cycle progression and transitions between different phases, we computed the Least Action Paths (LAPs) and Mean First Passage Time (MFPT) between different cell cycle phases. LAPs are optimal phase space trajectories minimizing the action's cost of moving from one state to another, while MFPT is the average transition time between states. Specifically, the optimal path between any two cell states (e.g., the fixed points of G0 phase and G1/S phase) is searched by changing the continuous path connecting the source and target states while minimizing its effect and updating the associated transition time[9]. The resulting LAP has the highest transition probability and is correlated with a specific transition time.

We first map the LAPs between different cell cycle phases in the velocity vector field, where the nodes along the paths represent the discrete points, with color indicating LAP transition time and arrows depicting path direction, and the black curves connecting different phases represent LAPs (**Figure 4A** and **S14A**). We also calculated and illustrated the LAPs on the landscape, which reveals the optimal path (least action) that cells take to overcome potential barriers during the cell cycle (**Figure 4B** and **S14B**). Green curves connect G0 to G1-S and M to G0, representing the cell cycle's initiation and termination, respectively. The cell cycle path is non-uniform and sandwiched between flux (black arrows) and negative potential gradient (white arrows), demonstrating that progression arises from both curl flux and negative potential gradient force. To enter the cycle from the G0 basin to the G1-S basin, the cells must overcome a barrier on the path (saddle point on the landscape) located between the G0 and G1-S basins, where the flux acts as a driving force.

In addition to calculating the LAPs, we also quantified the time of the cell cycle initiation and termination (**Figure 4C-4E**). We indicated that noise reduces the least actions of cell cycle initiation and termination, resulting in a decrease in MFPT, suggesting that strong fluctuations make it easier for cells to enter the cycle, but also easier to terminate the cell cycle when the cycle coherence is weakened. Furthermore, we found that the logarithm of MFPT is directly proportional to the barrier height on the landscape, i.e. the higher the barrier, the more stable the cell state, and the more difficult it is for the cell to leave this cycle phase (the larger the least action and the longer the escape time) (**Figure 4F** and **S13A**). The barrier heights between the attraction basins correlate with the escape time, which characterizes the stability

of the cell cycle phase. Additionally, the logarithm of MFPT is negatively correlated with the logarithm of non-equilibrium dynamic flux, suggesting that the larger the flux, the stronger the driving force, and the easier it is for cells to escape a given steady state (**Figure 4G**). However, thermodynamic energy cost-wise, the logarithm of MFPT is positively correlated with the logarithm of EPR (**Figure 4H**). The longer the cell takes to escape, the more energy it consumes. To overcome this barrier, cells must expend a certain amount of energy, known as cell cycle initiation energy[41]. In summary, these results provide a quantitative assessment of the initial energy required for the progression of the cell cycle (e.g., nutrient supply capacity and metabolic activity[42]) and provide graphical, physical, and quantitative explanations for the checkpoint mechanisms of the cell cycle.

Maintenance of the cell cycle requires energy-dissipation and flux to enhance the coherence of the oscillation phase

Biological systems need to function accurately in the presence of strong noise, and various regulatory mechanisms have evolved to suppress the fluctuation and environmental effects in processing vital life processes such as cell cycle and development[43]. In a noisy environment, cell cycle oscillations can be highly inaccurate due to phase fluctuation, the phase of oscillation follows diffusive dynamics[44,45]. A previous study found that phase coherence is a property of waves that describes how well-aligned the phases of the waves are. When waves are perfectly coherent, their peaks and troughs align precisely, creating a stable and predictable pattern[7,46]. We quantified the relationship between noise intensity and cell cycle oscillation properties (phase diffusion, phase coherence, and coherent time) and found that noise fluctuations increase phase diffusion, decrease phase coherence, and make the coherence time (cell cycle maintenance time) to become lower (**Figure S15** and **S16**). In the system of cell cycle regulation, phase diffusion refers to the random fluctuation of the phase of oscillate waves. After certain times, the oscillations can lose the memory or track of the previous dynamics. This makes it harder to maintain the periodicity. The phase coherence is important for maintaining accurate and reliable oscillations in noisy environments. Coherent time refers to the duration over which a system exhibits coherence, or the ability to maintain a stable and predictable pattern[47]. In other words, it is the length of time during which oscillating waves in the cell cycle system remain perfectly aligned. There are significant correlations between these factors: increasing phase diffusion leads to a corresponding decrease in phase coherence, and phase diffusion can reduce the coherent time of the cell cycle system, especially in large fluctuation systems (**Figure 5A** and **5B**). Decreased phase coherence leads to decreased coherent time because less well-aligned waves become less stable and predictable over time (**Figure 5C**).

To identify the non-equilibrium mechanism of the cell maintenance cycle progression, we further investigated the effects of dynamic flux and thermodynamic dissipation on the characteristic features of cell cycle oscillations. Our results indicate that the logarithm of phase diffusion is negatively correlated with the logarithm of EPR (energy dissipation rate), while phase coherence and coherent time are positively correlated with the logarithm of EPR (**Figure 5D-5F**). This suggests that thermodynamic dissipation plays an essential role in

regulating the overall dynamics of the cell cycle oscillation. As cells enter the cycle, metabolic processes generate ATP and other energy molecules to provide thermodynamic energy dissipation to maintain the cell cycle oscillations. This added energy can help counteracting the effects of phase diffusion, maintaining coherence, and resulting in more reliable and effective signaling between different cell cycle phases. We also found that changes in the logarithm of phase diffusion were negatively correlated with changes in the logarithm of dynamic flux (**Figure 5G)**. Furthermore, we observed that higher flux levels were associated with increased phase coherence and coherence time (**Figure 5H** and **5I)**. These findings highlight the critical role of non-equilibrium dynamic flux force in increasing the stability of the cell cycle flow by enhancing phase coherence and coherent time.

Overall, these results suggest that there is a cost-performance trade-off for noisy biochemical oscillations, where increasing phase coherence comes at a higher thermodynamic cost. While the energy cost of an oscillating system can counteract the effects of phase diffusion and maintain coherence over longer periods of time, continuous energy input is required to sustain the system's oscillations, which can be energetically costly in certain circumstances. The energy cost of cell cycle progression is essential for ensuring that cells divide in a controlled and regulated manner, but excessive energy cost or disruptions in energy metabolism can result in cell cycle defects and potentially serious health problems such as cancer. Therefore, regulating energy metabolism and maintaining cell cycle oscillations are critical for preserving the health and proper function of cells and tissues.

Cell cycle gene regulatory network inference from single-cell transcriptomics data

Gene expression does not appear in isolation, which is influenced by a complex interplay of regulatory interactions with other genes and small molecules. The discovery of these regulatory interactions is the aim of gene regulatory network (GRN) inference methods[48-54]. The use of transcription factor (TF) regulatory network modeling to analyze biological dynamic processes is an interdisciplinary field that has been widely studied in systems biology. To decode the regulatory mechanism of cell cycle phases, we first identified 27 and 16 differentially expressed TFs of U2OS cell and RPE1 cell, separately (**Table S1** and **S2**), which were used to construct the genetic network that underlies each cell cycle stage. The analytic results demonstrate that the enriched expression of cycle phase-specific TFs regulates and maintains cells into different cycle stages via corresponding niche components (**Figure 6A** and **S18A**). The nodes in the network represent TFs colored according to their associated cell cycle phase, while the edge thickness represents the relative scores of interactions between connecting states. We also identified key TFs for each cell cycle phase and constructed a wiring diagram for the cell cycle network model with these key TFs (**Figure 6B** and **S18B**). The GRN identifies key TFs for each cell cycle phase known to play important roles in different cell cycle phases and biological processes such as DNA replication and mitosis. For instance, *CDC6*, *MCM3*, and *CCNE2* are key regulators of G1/S transition and cell cycle initiation that regulate multiple GRNs (*CyclinE* and *CDK2*) in the G1-S phases[38,55]. While the S phase is regulated by *MCM6*, *AURKB*, and *CDK1*, which are related to DNA replication and

phosphorylate several key proteins to control the transition from G1 to S phase of the cell cycle[56-58]. Other important TFs identified include *CDC25C*, *CENPE*, and *KIF23*, which regulate the G2/M checkpoint by affecting mitotic cell cycle[59-61]. Additionally, the GRN reveals that the M phase is regulated by *CDC20*[35], *PIF1*[62] and *TPX2*[63], *PDE4B*[64], and *KIF20A*[65], highlighting the crucial role of these factors in controlling the M-G1 phase.

The GRN inferred above only indicates the presence or absence of the interaction between TFs, but does not provide quantitative information on positive and negative regulation relationships. Based on the differential geometry analysis of the reconstructed cell cycle vector field from the single-cell transcriptomics data, we reveal the quantitative information about gene regulation and further uncover the molecular mechanism underlying the maintenance of the cell cycle and the interactions between the key TFs. Taking U2OS cells as an example, we demonstrate via our inferred regulatory network that *CCNE2* and *CENPE* regulate the G1-S and G2-M phase, respectively, showing peak expression in their respective regulatory cycle phases (**Figure 6C**). Jacobian analysis of vector fields can help us study cellular state-dependent gene interactions, with positive and negative values of the Jacobian matrix reflecting activation and inhibition regulation, respectively[23]. Across all cells, Jacobian analyses demonstrate the activation of *CCNE2* by *CENPE*, repression of *CENPE* by *CCNE2*, self-activation of *CENPE*, and self-activation of *CCNE2* (**Figure 6D**). These results collectively suggest that the interlinked positive and negative feedback loops between *CCNE2*, which regulates the G1-S checkpoint, and *CENPE*, which regulates the G2-M checkpoint, underlies the dynamics of cell cycle oscillations, revealing the critical role of two transcription factors in maintaining cell cycle progression (**Figure 6F**). Although only one negative feedback loop is required for biological oscillation, we found that *CCNE2* and *CENPE* have additional self-activating positive feedback loops that promote modularity of cell cycle frequency while essentially unchanged amplitude (**Figure 2J** and **S7**). Previous computational studies have shown that it is often difficult to adjust the frequency of a negative feedback oscillator without affecting the amplitude, whereas with positive plus negative feedback, a widely adjustable frequency and a near-constant amplitude can be achieved[66]. This also reveals that keeping the cell cycle more robust and easier to evolve requires a lot of positive and negative feedback coupling. To further quantitatively understand the feedback regulation mechanism between regulators, we first plotted distributions of the four Jacobian elements versus expression of each gene (**Figure S17**). Then, we estimated kinetic parameters with derivatives of a simplistic inhibitory or activation hill equation by fitting the Jacobian vs. expression curve in the response heatmap (**Figure 6E**). Finally, we theoretically and numerically simulate the velocity dynamics curve of the hill equation as a function of gene expression (**Figure 6G**).

Similarly, based on the REP1 cell cycle regulation network inferred from scEU-seq data, we further revealed the feedback mechanism between *CCNE1*(a G1-S phase regulator) and *KIF23* (a G2-M phase regulator), and quantified the intensity of feedback between each other (**Figure S18** and **S19**). By inferring the feedback mechanism between different cell cycle phase regulators, we confirmed that the cell cycle regulatory network is composed of motifs coupled with many negative feedback loops, which together constitute a topology that maintains periodic oscillation. These analyses not only infer the regulatory network of cell cycle function from single-cell omics data, so as to find out the topological basis of cell function, but also further quantify the interaction strength between the regulatory factors, which can provide

newly quantitative and accurate guidance for the on-demand design of functional elements in synthetic biology[67].

Discussion

The cell cycle is a complex process that is tightly regulated to ensure proper cell growth, differentiation, and proliferation. Dysregulation of the cell cycle can lead to various diseases, including cancer. Therefore, understanding the fundamental mechanisms that govern the cell cycle is critical for advancing our knowledge of cellular physiology and developing therapeutic strategies for treating diseases. Experimental advances in single-cell multi-omics have provided a unique and rich perspective on this process, but we lack an appropriate analytical framework to use these data to study the underlying mechanisms through the non-equilibrium dynamics and thermodynamic behavior of cells. In this study, we utilized an integrative approach combining single-cell transcriptomics with nonequilibrium statistical physics to better understand the complex regulatory networks and analyze the molecular events involved in the cell cycle at the single-cell level to fill this unmet gap. This process includes processing the upstream dimension reduction clustering of single-cell transcriptome data of the cell cycle, estimation of RNA velocity, reconstruction of cell cycle vector fields, and quantification of non-equilibrium dynamic and thermodynamic landscapes-flux.

Conventionally, researchers have focused on examining the dynamics of trajectories or analyzing the local stabilities of the nonlinear dynamical systems. However, these approaches often offer only a local view of the system's stability and behavior. To obtain a more comprehensive understanding of the overall stability and behavior, people have turned to the concept of landscape, which can describe the global stability of the system. However, it has become apparent that relying solely on the landscape for the dynamics is true only for equilibrium systems with detailed balance (no net input to or output from the system) is insufficient for capturing the entire dynamics for the nonequilibrium systems[7-9,17,40,68,69]. Our findings in this study of the cell cycle illustrate this point well, as its complex behavior requires the presence of an additional rotational flux. While the landscape contributes to attracting the system into the close oscillation ring valley, it is the flux that drives the coherent oscillation along the cycle path. By combining both concepts, one can gain deeper insights into the mechanisms behind the stability and behavior of the complex systems.

One of the key findings of our research is the significant impact of the biological noise on the underlying mechanism of the cell cycle via nonequilibrium dynamics and thermodynamics. In the small noise range, when the cell cycle oscillation is not disrupted, the noise increases the cell cycle amplitude while conversely decreases the cell cycle period, accompanied by the decrease in the curl flux and EPR. Larger flux means more energy input and larger cell cycle driving force and therefore lead to less time to complete the cell cycle oscillation (the period is smaller). Increased the noise effectively makes potential landscape basins shallower, thereby broadening the distribution of biological variables. After the noise increases to a certain extent, the potential energy landscape shows completely different basins, the cell cycle is broken, and the cells arrest in different cycle stages (attractor-basins). At this time, as the noise increases, although the curl flux increases, it appears that the dominant role of the noise masks the energy supply which is not enough to maintain the coherent oscillation. Our study also

demonstrated that the key genetic perturbations can significantly alter the underlying mechanism of the cell cycle via nonequilibrium dynamics and thermodynamics. Specifically, perturbations of cell cycle gene related to cell cycle initiation, DNA replication or mitosis all lead to landscape-flux changes that are consistent with biological effects. These highlight the importance of genetic perturbations in the cell cycle and can be used to predict the pathological effects of certain genetic perturbations, providing a framework for understanding the molecular basis of genetic diseases and diseases that affect the cell cycle. By identifying the bottlenecks and rate-limiting steps in the cell cycle, researchers may be able to develop targeted therapies that can disrupt the cell cycle and selectively kill cancer cells[70]. Therefore, future research could build on these findings to develop more accurate models of the cell cycle and other biological processes, and to explore the potential impact of genetic and environmental perturbations on these systems.

We also studied the underlying mechanism via non-equilibrium dynamics and thermodynamics of the cell cycle initiation and termination by counting MFPTs and LAPs between different cell cycle phases. In general, MFPT is defined as the average transition time of a system switching from one state to another, which can be measured experimentally accurately[71]. The positive correlation between barrier height and MFPT more intuitively shows the reasons for the difficulties of cell initiation and termination from the 3D landscape, and it is worth noting that our previous works can experimentally quantify the landscape and barrier heigh[71-73]. The analysis of EPR and MFPT reveal the existence of basal energy supply in the cell cycle, as well as the initial energy dissipation for cell cycle initiation[41], which approximates the product of EPR and MFPT from G0 phase to G1-S phase. We also found that the cell cycle is a highly energy- dissipative process, and the maintenance of the cell cycle requires thermodynamic energy-dissipation and dynamic flux to enhance the coherence of the oscillation phase against stochastic fluctuations. From a thermodynamic point of view, living systems are open systems in a non-equilibrium state, requiring material input and energy consumption to maintain their steady state and order[74,75]. The cell cycle is naturally a living system, G0 phase is a non-dividing state of the cell cycle, in which the cell is metabolically active but not actively dividing. To switch from G0 phase to the G1/S phase, the cell requires a certain amount of energy to initiate the complex set of biochemical reactions that drive cell division. Once the cell has entered the G1/S phase, it must maintain a complex system of oscillations and feedback loops in order to progress through the rest of the cell cycle. This requires additional energy input to maintain the oscillations of key cell cycle regulators such as cyclins and cyclin-dependent kinases (CDKs). Currently, certain experiments can quantify the energy flow by measuring the rate of oxygen consumption or heat production or the fluxes of metabolites including the production and comsumption of ATP[76].

In systems biology, the creation of core network regulation to analyze biological processes is an important and popular topic. In general, there are two ways to build a biological regulatory network: one is to create a small network based on the existing knowledge and database, and use the simulator to improve the network, but it is inefficient and cannot be used to build a new network model[77,78]. The second is to use bioinformatics technology, especially high-throughput sequencing data, to infer genes and gene correlations, but to ignore the actual biological modulation relationship[79]. To balance the advantages and disadvantages of these two approaches, we infer the transcriptional regulatory networks based on differentially expressed

TFs of each cycle phase from single-cell transcriptomics data. The regulatory direction between TFs (activation/inhibition) can be infered by jacobian analyses of vector fields and the hill function of the core network can be simulated by fitting jacobian in the gene expression space and response heatmap. We reveal that cell cycle regulation involve multiple positive and negative feedback loops between these regulators, which help to adjust the frequency of the oscillation without significantly affecting the amplitude, enabling widely adjustable frequencies and near-constant amplitudes and with maintaining balance between the activation and inhibition of the gene expressions, permitting for proper progression through different cell cycle phases. In this regard, further research is required to unravel such complex regulatory networks comprehensively. Our results identify that some dynamic biological processes can be accurately mapped by constructing regulatory network models based on the high-throughput sequencing data.

In summary, we have built a general framework for the quantification of landscape-flux and analysis of the underlying mechanism of the cell cycle via non-equilibrium dynamics and thermodynamics based on single-cell transcriptomic data that can be applied to numerous biological systems. More broadly, when coupled with remarkable advances in single-cell approaches, including RNA velocity[80], construction of vector fields[32], as well as spatial multi-omics[81], we will enable a deeper, quantitative understanding of the spatio-temporally underlying mechanism of the cell cycle and other biological systems via non-equilibrium dynamics and thermodynamics to help prevent various diseases, including cancer and neurodegeneration, and open up new possibilities for precision medicine.

Acknowledgements

Ligang Zhu thanks the supports from the National Natural Science Foundation of China Grant No. 12234019, 21721003. We would also like to thank Kun Zhang, Li Xu, Wenbo Li, Han Yan, Cong Yu, Linqi Wang and Songlin Yang for valuable discussions during the preparation of this manuscript.

Author contributions

L.Z. designed the research, developed the methods, carried out the stimulations, performed the analysis, and wrote the manuscript. J.W. initiated, designed and supervised the research, developed the methods, analyzed the data, and wrote the manuscript.

Declaration of interests

The authors declare no competing interests.

Methods

scRNA-seq data and scEU-seq data processing and analysis

All scRNA-seq and scEU-seq data in this study were downloaded publicly (see details in the 'Data availability' section).

(1) For the RNA sequencing data for ~1k U2OS-FUCCI data, we followed the *snakemake* pipeline at https://github.com/CellProfiling/FucciSingleCellSeqPipeline to perform scRNA-Seq data preparation and general analysis including filter, dimensionality reduction, clustering analysis and RNA velocity analysis[18].

(2) For the scEU-seq data for ~3k RPE1-FUCCI cells, we used *dyn.sample_data.scEU_seq_rpe1()* to acquire the processed data by using Dynamo[23], which includes cell cycle clustering and RNA velocity.

(3) We reconstructed the vector field based RNA velocity with *dyn.vf.VectorField()* by using Dynamo[23]. Then, we calculated the divergence, curl, acceleration and curvature to perform the differential geometry analysis. We also calculated the jacobian to perform genetic perturbation and inference gene regulatory interaction.

(4) We learned an analytical function of vector field from sparse single cell samples on the entire space robustly by *vector_field_function.* Then, we calculated stochastic dynamics by solving the Langevin equation based analytical function and quantified the non-equilibrium thermodynamics and dynamics of the cell cycle.

RNA velocity

RNA velocity calculations using Velocyto[21] (version 0.17.17), Scvelo[22] (version 0.1.24) and Dynamo[23] (1.1.0). In brief, after the gene selection and normalization, the first- and second-order moments were calculated with *scv.pp.moments()* function. The full splicing kinetics were recovered with *scv.tl.recover_dynamics()* function and the velocities were obtained with *scv.tl.velocity()* function in dynamical mode. The velocities were projected onto diffusion maps and visualized as streamlines with *scv.pl.velocity_embedding_stream()* function. The spliced vs. unspliced phase portraits of individual genes were visualized with *scv.pl.velocity()*. The pseudotimes of cells were obtain with *scv.tl.recover_latent_time()* function.

Reconstructing vector field from RNA velocity

Dynamo[23] was used to reconstruct function of the continuous and analytical velocity vector field from sparse, noisy single-cell velocity measurements with sparse approximation of vector field consensus (sparseVFC)[82]. SparseVFC uses a vector-valued kernel method built on reproducing kernel Hilbert space (RKHS) to learn the vector field, which is expressed analytically as a weighted linear combination of a set of vector-valued kernel basis functions. The learning process relies on sparse approximation to estimate the coefficients (weights) of a selected number of basis functions, each associated with a control point, which is often much smaller than the number of data points. With sparse approximation, the reconstruction scales linearly with the number of data points in both computational time and memory requirements.

To account for the noise and outliers of velocity measurements, sparseVFC relies on an EM algorithm to iteratively optimize the set of inliers, as well as the optimized coefficient set for each basis function, further improving the robustness of vector field reconstruction. With the continuous vector field that is learned in either high-dimensional principal component analysis (PCA) space, which can be projected back to the full transcriptomic space, or lower dimensional space (such as 2D uniform manifold approximation and projection [UMAP] space), or directly in the full gene expression space.

Quantification of potential landscape

The stochastic dynamics of the cell cycle system is described by Langevin equation:

$$\frac{d\mathrm{UMAP}_i(t)}{dt} = \mathbf{F}(\mathrm{UMAP}_i) + \zeta$$

where UMAP_i represents each dimension of UMAP space, which can be projected back to the full transcriptomic space. $\mathbf{F}$ is the reconstructed vector field function. The noise term ζ adopts the independent additive white Gaussian noise, $\langle\zeta(t)\rangle$ = 0 and $\langle\zeta(t)\zeta(t')\rangle = 2D\delta(t - t')$. D is the intensity of Gaussian white noise. To obtain the potential landscape of cell cycle dynamics, starting from enough random initial conditions, the system will eventually evolve to different cell cycle phase steady state, which may be monostable, multistable or limit cycle steady state. All steady-state trajectories in the UMAP space are divided into several small regions, and the steady-state trajectory density in each small region is counted as steady-state probabilityP_{ss}. It includes the effects of diffusion through the resulting steady-state probability distribution. U is the generalized potential related to the steady-state probability by $U = -\ln P_{ss}$.

Driving force decomposition to the gradient of the potential landscape and curl flux

The probabilistic evolution of diffusion equation, $\partial P/\partial t + \nabla \cdot \mathbf{J}(\mathbf{x}, t) = 0$, represents a conservation law of probability (local change is due to net flux in or out). And the probability flux vector **J** of the system in concentration space **x** is defined as $\mathbf{J}(\mathbf{x}, t) = \mathbf{F}P - \mathbf{D} \cdot \frac{\partial}{\partial \mathbf{x}} P$. In general, the dynamic driving force **F** can be decomposed into a gradient of a potential and a curl flow flux[7]:

$$\mathbf{F} = \mathbf{D}/P_{ss} \cdot \frac{\partial}{\partial \mathbf{x}} P_{ss} + \mathbf{J_{ss}}(\mathbf{x})/P_{ss} = -\mathbf{D}\frac{\partial}{\partial \mathbf{x}}\mathrm{U} + \mathbf{J_{ss}}(\mathbf{x})/P_{ss}.$$

P_{ss} represents steady state probability distribution and potential landscape U is defined as $\mathrm{U} = -\ln P_{ss}$. At steady state, $\frac{\partial P_{ss}}{\partial t} = 0$, thus $\nabla \cdot \mathbf{J_{ss}} = 0$. With detailed balance, there is no net flux and $\mathbf{J_{ss}} = 0$, the gradient of potential controls the underlying dynamics as the driving force. For nonequilibrium systems, although $\nabla \cdot \mathbf{J_{ss}} = 0$, the net flux $\mathbf{J_{ss}} \neq 0$. Therefore, both the gradient of potential landscape and nonzero steady state probability net flux of probability together form the driving force and determine the dynamics and global properties. The

dynamics of a nonequilibrium network can be described as a spiral, along the gradient direction, not as the case of the equilibrium state only following the gradient. It is similar to the electrons moving in both electric and magnetic fields[7].

EPR (Energy Production Rate)

In a nonequilibrium open system, there are constant exchanges in energy and information. This results in the dissipation of energy. The dissipation gives a global physical characterization of the nonequilibrium system. In the steady state, the dissipation of energy is closely associated with the entropy production rate. The entropy formula for the system is well known[83],

$$S = -k_B \int P(\mathbf{x}, t) \ln P(\mathbf{x}, t) dx \, .$$

By differentiating the above function, the increase of the entropy at constant temperature T is shown as follows:

$$T\dot{S} = k_B * T \int (\ln P + 1) \nabla \cdot \mathbf{J} \, dx = -\int (k_B T \nabla \ln P - \mathbf{F}) \cdot \mathbf{J} \, dx - \int \mathbf{F} \cdot \mathbf{J} \, dx$$

where $-\int (k_B T \nabla \ln P - \mathbf{F}) \cdot \mathbf{J} \, dx = EPR$ is the entropy production rate, and $\int \mathbf{F} \cdot \mathbf{J} \, dx = HDR$ is the mean rate of the heat dissipation. In steady state, $\dot{S} = 0$, and the entropy production rate EPR is equal to the heat dissipation rate HDR. In this article, we calculated the heat dissipation rate at steady state and also verified that it is the same numerically as entropy production rate at steady state.

Flux Integration

To quantify the curl probabilistic flux, we define $Flux_{Loop}$ as the flux integration along the loop trajectory ($Flux_{Loop} = \oint_0^L \mathbf{J} \cdot d\mathbf{l} / \oint_0^L d\mathbf{l}$) of the cell cycle oscillation divided by the loop length in gene expression space. From **J** = **F***P* − **D** · ∇*P*, we have $\oint_0^L \mathbf{J} \cdot d\mathbf{l} = \oint_0^L \mathbf{F} P \cdot d\mathbf{l} - D \cdot \oint_0^L \nabla P \cdot d\mathbf{l} = \oint_0^L \mathbf{F} P \cdot d\mathbf{l}$. We calculated the flux integration results in 2-dimensional UMAP space projection.

LAPs (Least Action Paths)

Considering a stochastic dynamical system described by Langevin equation:

$$\frac{d\mathrm{UMAP}_i(t)}{dt} = \mathbf{F}(\mathrm{UMAP}_i) + \zeta \, .$$

Following the approaches based on the Freidlin-Wentzell theory[84], the most probable transition path from attractor *j* at time 0 to attractor *k* at time *T*, can be acquired by minimizing the action functional over all possible paths:

$$S_T(\mathbf{x}_{jk}) = \frac{1}{2} \int_0^T \frac{\left| \dot{\mathbf{x}}_{jk} - \mathbf{F}(\mathbf{x}_{jk}) \right|^2}{D} dt$$

Where *D* is the diffusion coefficient accounting for the stochasticity of gene expression, and for simplicity here we assume it to be a constant. This path **x** is called the least action path (LAP). Specifically, the optimal path between any two cell states is searched by variating the continuous path connecting the source state to the target while minimizing its action and updating the associated transition time. The resultant LAP has the highest transition probability and is associated with a particular transition time. This type of approach has been used in many chemical and biological systems.

MFPT (Mean First Passage Time)

For rare transitions with $S_T \gg 0$, the transition rate (number of transitions per unit time) is proportional to the exponential of the actions. The Freidlin–Wentzell theorem dictates that the LAP with the minimal traversal time (which will be referred to as the optimal path below) contributes the most to this transition rate[84]:

$$R(A \to B) \approx C \exp(-S_T)$$

where *A* and *B* are two cell cycle phases, S_T is the action of the optimal path, and C a proportional factor. Furthermore, the transition time, or more specifically the mean first passage time (MFPT), is related to the transition rate:

$$\mathrm{MFPT} = \frac{1}{R(A \to B)}$$

Therefore, the action of the optimal path predicts both the likelihood and transition time for such rare transitions.

Phase diffusion and coherence time

We simulated 500 trajectories starting with the same initial condition. For the *j*-th trajectory, we obtained its *i*-th peak time t_{ij} from the trajectory $x_j(t)$ after smoothing (smooth function in MATLAB was used). The peak positions for two trajectories are shown in Figure S6. For each of the trajectories, we computed the mean of their *i*-th peak time $m_i = \sum_j t_{ij}/N$, and variance $\sigma_i^2 = \sum_j (t_{ij} - m_i)^2/(N-1)$, where N is the total number of trajectories. The average period T is given by $T = m_i/i$. Asymptotically, σ_i^2 depends linearly on m_i, and the slope of this linear dependence is the peak-time diffusion constant *D*, which has the dimension of time. The phase diffusion constant $D_\emptyset$ is linearly proportional to *D*[44]: $D_\emptyset = D\frac{(2\pi)^2}{T}$.

The autocorrelation function[44]:

$$C(t) = \frac{\langle (x(t+s) - \langle x \rangle)(x(s) - \langle x \rangle) \rangle}{\langle x^2 \rangle - \langle x \rangle^2} = \exp\left(-\frac{t}{\tau_c}\right) \times \cos\left(\frac{2\pi t}{T}\right)$$

where *s* is a time variable, the period is given by T and τ_c is the coherence time for the oscillation.

Power Spectral Density (PSD) from autocorrelation

The distribution of average power of a signal $x(t)$ in the frequency domain ω is called the power spectral density (PSD) or power density (PD) or power density spectrum. In order to drive the PSD function, consider a power signal as a limiting case of an energy signal, i.e., the signal $Z(t)$ is zero outside the interval |τ/2|, the signal $Z(t)$ is given by,

$$Z(t) = \begin{cases} x(t), & |t| < \tau/2 \\ 0, & \text{otherwise'} \end{cases}$$

Where, $x(t)$ is a power signal of same magnitude extending to infinity.
As the signal $Z(t)$ is finite duration signal of duration *τ* and thus, it is an energy signal having energy E, that is given by,

$$E = \int_{-\infty}^{\infty} |Z(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |Z(\omega)|^2 d\omega,$$

where $Z(t) \overset{FT}{\leftrightarrow} Z(\omega)$. Also,

$$\int_{-\infty}^{\infty} |Z(t)|^2 dt = \int_{-\tau/2}^{\tau/2} |x(t)|^2 dt.$$

Therefore, we have,

$$\frac{1}{\tau} \int_{-\tau/2}^{\tau/2} |x(t)|^2 dt = \frac{1}{2\pi} \cdot \frac{1}{\tau} \int_{-\infty}^{\infty} |Z(\omega)|^2 d\omega,$$

Hence, when $\tau \to \infty$, then the LHS of the above equation gives the average power (P) of the signal $x(t)$, i.e.,

$$P = \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{\tau \to \infty} \left(\frac{|Z(\omega)|^2}{\tau} \right) d\omega,$$

if $\tau \to \infty$, then $\left(\frac{|Z(\omega)|^2}{\tau}\right)$ approaches a finite value. Assume this finite value is represented by $S(\omega)$, i.e.,

$$S(\omega) = \lim_{T \to \infty} \frac{|Z(\omega)|^2}{\tau},$$

this expression is called the power spectral density (PSD) of the signal $Z(t)$. Therefore, for the function $x(t)$, the PSD function is given by,

$$S(\omega) = \lim_{T \to \infty} \frac{|x(\omega)|^2}{\tau}.$$

The power spectral density function $S(\omega)$ and the autocorrelation function $R(\tau)$ of a power signal form a Fourier transform pair, i.e.,

$$R(\tau) \overset{FT}{\leftrightarrow} S(\omega).$$

The Cross-Correlation Function and Time Irreversibility

Simulated time series data from the cell cycle dynamics exhibits noise-induced switching between the cell cycle limit cycle and the different cell cycle phase steady-state, mimicking observations from real cell cycle systems[8,9]. Practical early warning signals based on the landscape-flux potential theory demand that we quantify the nonequilibrium nature of the

system based on such time series data and hence we define the cross-correlation function as follows:

$$C_{XY}(\tau') = \langle X(0)Y(\tau')\rangle,$$

where *X* and *Y* denote the time trajectories of variables *X* and *Y* with time interval τ'[85]. $C_{XY}(\tau')$ represents the cross-correlation function forward in time and $\tilde{C}_{XY}(\tau')$ represents the cross-correlation function backward in time. The average difference in cross-correlations between the forward in time and the backward in time, defined as:

$$\Delta C = \sqrt{\frac{1}{t_f}\int_0^{t_f}(C_{XY}(\tau') - \tilde{C}_{XY}(\tau'))^2 d\tau'},$$

can be used as a quantification of the time irreversibility of the system and the degree to which it is out of equilibrium. A way for quantifying the degree of the detailed balance breaking and time irreversibility is from the differences of the two-point cross correlation function forward and backward in times[85,86].

Phase coherence

The robustness of the oscillation with respect to the diffusion coefficient D can be quantified further by the phase coherence ξ, which measures the degree of periodicity of the time evolution of a given variable[46]. The phase coherence ξ quantitatively measures the degree of persistence of the oscillatory phase and is defined as follows. First, the vector $N(t) = n_1(t)e_1 + n_2(t)e_2$ is shown in Figure S20. The unit vectors are e_1 = (1,0) and e_2 = (0, 1), and $n_1(t)$ and $n_2(t)$ are the concentration of the two kinds of protein molecules at time t. Then, $\phi(t)$ is the phase angle between $N(t)$ and $N(t+\tau)$, where τ should be smaller than the deterministic period and larger than the fast fluctuations. $\phi(t)$ > 0 to represent that the oscillation goes on the positive orientation (counterclockwise). The formula of ξ is shown as

$$\xi = \frac{2\sum_i \theta(\phi(t))\phi(t)}{\sum_i |\phi(t)|} - 1$$

where $\theta(\phi)$ = 1 when $\phi(t) >$ 0, and $\theta(\phi)$ = 0 when $\phi(t) \leq$ 0, and sums are taken over every time steps for the simulant trajectory. $\xi \approx$ 0 means the system moves stochastically and has no coherence. The oscillation is most coherent as ξ is close to 1.

Gene regulatory network inference

To construct Transcriptional Factors interaction network of the cell cycle, we first calculated and obtained the all differentially expressed TFs in different cell cycle phase (**Table S1 and S2**). After that the interaction between the differentially expressed TFs were calculated by Jacobian based on the reconstructed vector field function. Positive values from Jacobian represent activation and negative values represent inhibition[23]. The visualization of all networks was performed using Cytoscape (version 3.6.1)[87]. The nodes and edges each represent the TF and interactions in a network.

Estimating kinetic parameters by fitting the Jacobian vs. expression curve

Because the reconstructed vector field is expressed as a set of implicit basis functions, not explicitly as Hill functions, in the current framework we are not able to directly obtain kinetic parameters such as the Hill coefficient. Nevertheless, the reconstructed vector field encodes such information, and additional computations are applied to extract that information[23]. We demonstrate this possibility on simplistic network motifs such as CCNE2-CENP2 and CCNE2-KIF23, by fitting the derivatives of inhibitory or activation Hill equations to the corresponding Jacobian elements. Further efforts will be needed to make such efforts generally applicable to systems with more sophisticated mechanisms.

Formally, we assume that the activation effect of gene x on the target gene y takes the form of an activating Hill function:

$$x \rightarrow y\colon \frac{Ax^n}{(K + x^n)},$$

and that the inhibition effect assumes the form of an inhibitory Hill function:

$$x \dashv y\colon \frac{AK}{(K + x^n)}.$$

For the activation-inhibition and self-activation of two gene x and y, the ODEs can write as:

$$\frac{dx}{dt} = f_x = \frac{A_1 x^{n1}}{(K_1 + x^{n1})} + \frac{A_2 K_2}{(K_2 + y^{n2})} - g_1 \cdot x,$$
$$\frac{dy}{dt} = f_y = \frac{A_3 y^{n3}}{(K_3 + y^{n3})} + \frac{A_4 x^{n4}}{(K_4 + x^{n4})} - g_2 \cdot y.$$

The Jacobian of this system is a 2-by-2 matrix:

$$J = \begin{bmatrix} \frac{\partial f_x}{\partial x} & \frac{\partial f_x}{\partial y} \\ \frac{\partial f_y}{\partial x} & \frac{\partial f_y}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{A_1 K_1 n_1 x^{n_1 - 1}}{(K_1 + x^{n_1})^2} - g_1 & -\frac{A_2 K_2 n_2 y^{n_2 - 1}}{(K_2 + y^{n_2})^2} \\ -\frac{A_3 K_3 n_3 x^{n_3 - 1}}{(K_3 + x^{n_3})^2} & \frac{A_4 K_4 n_4 x^{n_4 - 1}}{(K_4 + y^{n_4})^2} - g_2 \end{bmatrix}.$$

In Figure 6E and S18E, the means and standard deviations of the Jacobian vs. expression profiles were calculated and fitted with the above derivatives using the Dynamo: *dyn.pl.response*, which can fit and obtain the parameters of hill function (Table S3).

Data availability

The scRNA-seq raw data of U2OS-FUCCI cell cycle from the paper by Mahdessian et al, which are available at GEO with accession GSE146773[18]. The scEU-seq data of RPE1-FUCCI cell cycle can be extracted using dynamo's CLI: *dyn.sample_data.scEU_seq_rpe1()* or from the paper by Battich et al, which are available at GEO with accession number GSE128365[19]. The processed data are available from Github (https://github.com/Zhu-1998/cellcycle). All other data supporting the findings of this study are available from the corresponding author on reasonable request.

Code availability

The source code used in this study is publicly available from Github (https://github.com/Zhu-1998/cellcycle).

References

1. Cho, R.J., Huang, M.X., Campbell, M.J., Dong, H.L., Steinmetz, L., Sapinoso, L., Hampton, G., Elledge, S.J., Davis, R.W., and Lockhart, D.J. (2001). Transcriptional regulation and function during the human cell cycle. Nat Genet *27*, 48-54. Doi 10.1038/83751.
2. Liu, L.J., Michowski, W., Kolodziejczyk, A., and Sicinski, P. (2019). The cell cycle in stem cell proliferation, pluripotency and differentiation. Nat Cell Biol *21*, 1060-1067. 10.1038/s41556-019-0384-4.
3. Teschendorff, A.E., and Feinberg, A.P. (2021). Statistical mechanics meets single-cell biology. Nat Rev Genet *22*, 459-476. 10.1038/s41576-021-00341-z.
4. MacArthur, B.D., and Lemischka, I.R. (2013). Statistical Mechanics of Pluripotency. Cell *154*, 484-489. 10.1016/j.cell.2013.07.024.
5. Tang, F.C., Lao, K.Q., and Surani, M.A. (2011). Development and applications of single-cell transcriptome analysis. Nat Methods *8*, S6-S11. 10.1038/Nmeth.1557.
6. Fang, X.N., Kruse, K., Lu, T., and Wang, J. (2019). Nonequilibrium physics in biology. Rev Mod Phys *91*, 045004. 10.1103/RevModPhys.91.045004.
7. Wang, J., Xu, L., and Wang, E.K. (2008). Potential landscape and flux framework of nonequilibrium networks: Robustness, dissipation, and coherence of biochemical oscillations. P Natl Acad Sci USA *105*, 12271-12276. 10.1073/pnas.0800579105.
8. Wang, J., Li, C.H., and Wang, E.K. (2010). Potential and flux landscapes quantify the stability and robustness of budding yeast cell cycle network. P Natl Acad Sci USA *107*, 8195-8200. 10.1073/pnas.0910331107.
9. Li, C.H., and Wang, J. (2014). Landscape and flux reveal a new global view and physical quantification of mammalian cell cycle. P Natl Acad Sci USA *111*, 14130-14135. 10.1073/pnas.1408628111.
10. Zhu, L., Li, X., Xu, F., Yin, Z., Jin, J., Liu, Z., Qi, H., and Shuai, J. (2022). Network modeling-based identification of the switching targets between pyroptosis and secondary pyroptosis. Chaos Solitons Fractals *155*, 111724. 10.1016/j.chaos.2021.111724.
11. Burkhardt, D.B., San Juan, B.P., Lock, J.G., Krishnaswamy, S., and Chaffer, C.L. (2022). Mapping Phenotypic Plasticity upon the Cancer Cell State Landscape Using Manifold Learning. Cancer Discov *12*, 1847-1859. 10.1158/2159-8290.Cd-21-0282.
12. Zhou, P.J., Wang, S.X., Li, T.J., and Nie, Q. (2021). Dissecting transition cells from single-cell transcriptome data through multiscale stochastic dynamics. Nat Commun *12*, 5609. 10.1038/s41467-021-25548-w.
13. Izzo, F., Lee, S.C., Poran, A., Chaligne, R., Gaiti, F., Gross, B., Murali, R.R., Deochand, S.D., Ang, C., Jones, P.W., et al. (2020). DNA methylation disruption reshapes the hematopoietic differentiation landscape. Nat Genet *52*, 378-387. 10.1038/s41588-020-0595-4.
14. Qin, X., Rodriguez, F.C., Sufi, J., Vlckova, P., Claus, J., and Tape, C.J. (2023). A Single-cell Perturbation Landscape of Colonic Stem Cell Polarisation. bioRxiv, 2023.2002.2015.528008. 10.1101/2023.02.15.528008.
15. Jin, S.Q., MacLean, A.L., Peng, T., and Nie, Q. (2018). scEpath: energy landscape-based inference of transition probabilities and cellular trajectories from single-cell transcriptomic data. Bioinformatics *34*, 2077-2086. 10.1093/bioinformatics/bty058.
16. Shi, J.F., Teschendorff, A.E., Chen, W.Y., Chen, L.N., and Li, T.J. (2020). Quantifying

Waddington's epigenetic landscape: a comparison of single-cell potency measures. Brief Bioinform *21*, 248-261. 10.1093/bib/bby093.

17. Wang, J., Zhang, K., Xu, L., and Wang, E. (2011). Quantifying the Waddington landscape and biological paths for development and differentiation. P Natl Acad Sci USA *108*, 8257-8262. 10.1073/pnas.1017017108.
18. Mahdessian, D., Cesnik, A.J., Gnann, C., Danielsson, F., Stenstrom, L., Arif, M., Zhang, C., Le, T., Johansson, F., Shutten, R., et al. (2021). Spatiotemporal dissection of the cell cycle with single-cell proteogenomics. Nature *590*, 649-654. 10.1038/s41586-021-03232-9.
19. Battich, N., Beumer, J., de Barbanson, B., Krenning, L., Baron, C.S., Tanenbaum, M.E., Clevers, H., and van Oudenaarden, A. (2020). Sequencing metabolically labeled transcripts in single cells reveals mRNA turnover strategies. Science *367*, 1151-1156. 10.1126/science.aax3072.
20. Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol *37*, 38-44. 10.1038/nbt.4314.
21. La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M.E., Lonnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. Nature *560*, 494-498. 10.1038/s41586-018-0414-6.
22. Bergen, V., Lange, M., Peidli, S., Wolf, F.A., and Theis, F.J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol *38*, 1408–1414. 10.1038/s41587-020-0591-3.
23. Qiu, X.J., Zhang, Y., Martin-Rufino, J.D., Weng, C., Hosseinzadeh, S., Yang, D.A., Pogson, A.N., Hein, M.Y., Min, K.H., Wang, L., et al. (2022). Mapping transcriptomic vector fields of single cells. Cell *185*, 690-711. 10.1016/j.cell.2021.12.045.
24. Zhao, Y., Zhang, W., and Li, T. (2023). EPR-Net: Constructing non-equilibrium potential landscape via a variational force projection formulation. arXiv:2301.01946. 10.48550/arXiv.2301.01946.
25. Hirsch, M.W., Smale, S., Devaney, R.L., and Hirsch, M.W. (2004). Differential equations, dynamical systems, and an introduction to chaos, 2nd Edition (Academic Press).
26. Moiseeva, V., Cisneros, A., Sica, V., Deryagin, O., Lai, Y.W., Jung, S.C., Andres, E., An, J., Segales, J., Ortet, L., et al. (2023). Senescence atlas reveals an aged-like inflamed niche that blunts muscle regeneration. Nature *613*, 169-178. 10.1038/s41586-022-05535-x.
27. Balazsi, G., van Oudenaarden, A., and Collins, J.J. (2011). Cellular Decision Making and Biological Noise: From Microbes to Mammals. Cell *144*, 910-925. 10.1016/j.cell.2011.01.030.
28. Eling, N., Morgan, M.D., and Marioni, J.C. (2019). Challenges in measuring and understanding biological noise. Nat Rev Genet *20*, 536-548. 10.1038/s41576-019-0142-2.
29. Levchenko, A., and Nemenman, I. (2014). Cellular noise and information transmission. Curr Opin Biotech *28*, 156-164. 10.1016/j.copbio.2014.05.002.
30. Pujadas, E., and Feinberg, A.P. (2012). Regulated Noise in the Epigenetic Landscape of Development and Disease. Cell *148*, 1123-1131. 10.1016/j.cell.2012.02.045.
31. Rukhlenko, O.S., Halasz, M., Rauch, N., Zhernovkov, V., Prince, T., Wynne, K., Maher, S., Kashdan, E., MacLeod, K., Carragher, N.O., et al. (2022). Control of cell state transitions. Nature *609*, 975-985. 10.1038/s41586-022-05194-y.
32. Kamimoto, K., Stringa, B., Hoffmann, C.M., Jindal, K., Solnica-Krezel, L., and Morris, S.A. (2023). Dissecting cell identity via network inference and in silico gene perturbation. Nature *614*,

742-751. 10.1038/s41586-022-05688-9.
33. Fleck, J.S., Jansen, S.M.J., Wollny, D., Zenk, F., Seimiya, M., Jain, A., Okamoto, R., Santel, M., He, Z., Camp, J.G., and Treutlein, B. (2022). Inferring and perturbing cell fate regulomes in human brain organoids. Nature. 10.1038/s41586-022-05279-8.
34. Hu, H., Feng, Z., Lin, H., Zhao, J.J., Zhang, Y.R., Xu, F., Chen, L.L., Chen, F., Ma, Y.L., Su, J.Z., et al. (2023). Modeling and analyzing single-cell multimodal data with deep parametric inference. Brief Bioinform *24*, bbad005. 10.1093/bib/bbad005.
35. Piano, V., Alex, A., Stege, P., Maffini, S., Stoppiello, G.A., Veld, P.J.H.I., Vetter, I.R., and Musacchio, A. (2021). CDC20 assists its catalytic incorporation in the mitotic checkpoint complex. Science *371*, 67-71. 10.1126/science.abc1152.
36. Huang, R.X., and Zhou, P.K. (2020). DNA damage response signaling pathways and targets for radiotherapy sensitization in cancer. Signal Transduct Tar *5*, 60. 10.1038/s41392-020-0150-x.
37. Huang, L., Wang, H.Y., Li, J.D., Wang, J.H., Zhou, Y., Luo, R.Z., Yun, J.P., Zhang, Y., Jia, W.H., and Zheng, M. (2013). KPNA2 promotes cell proliferation and tumorigenicity in epithelial ovarian carcinoma through upregulation of c-Myc and downregulation of FOXO3a. Cell Death Dis *4*, e745. 10.1038/cddis.2013.256.
38. Sonntag, R., Giebeler, N., Nevzorova, Y.A., Bangen, J.M., Fahrenkamp, D., Lambertz, D., Haas, U., Hu, W., Gassier, N., Cubero, F.J., et al. (2018). Cyclin E1 and cyclin-dependent kinase 2 are critical for initiation, but not for progression of hepatocellular carcinoma. P Natl Acad Sci USA *115*, 9282-9287. 10.1073/pnas.1807155115.
39. Graf, T., and Enver, T. (2009). Forcing cells to change lineages. Nature *462*, 587-594. 10.1038/nature08533.
40. Feng, H.D., Zhang, K., and Wang, J. (2014). Non-equilibrium transition state rate theory. Chem Sci *5*, 3761-3769. 10.1039/c4sc00831f.
41. Papagiannakis, A., Niebel, B., Wit, E.C., and Heinemann, M. (2017). Autonomous Metabolic Oscillations Robustly Gate the Early and Late Cell Cycle. Mol Cell *65*, 285-295. 10.1016/j.molcel.2016.11.018.
42. Zhu, J.J., and Thompson, C.B. (2019). Metabolic regulation of cell growth and proliferation. Nat Rev Mol Cell Bio *20*, 436-450. 10.1038/s41580-019-0123-5.
43. Ferrell, J.E., Tsai, T.Y.C., and Yang, Q.O. (2011). Modeling the Cell Cycle: Why Do Certain Circuits Oscillate? Cell *144*, 874-885. 10.1016/j.cell.2011.03.006.
44. Cao, Y.S., Wang, H.L., Ouyang, Q., and Tu, Y.H. (2015). The free-energy cost of accurate biochemical oscillations. Nat Phys *11*, 772-778. 10.1038/Nphys3412.
45. Fei, C.Y., Cao, Y.S., Ouyang, Q., and Tu, Y.H. (2018). Design principles for enhancing phase sensitivity and suppressing phase fluctuations simultaneously in biochemical oscillatory systems. Nat Commun *9*, 1434. 10.1038/s41467-018-03826-4.
46. Yoda, M., Ushikubo, T., Inoue, W., and Sasai, M. (2007). Roles of noise in single and coupled multiple genetic oscillators. J Chem Phys *126*, 115101. 10.1063/1.2539037.
47. Potoyan, D.A., and Wolynes, P.G. (2014). On the dephasing of genetic oscillators. P Natl Acad Sci USA *111*, 2391-2396. 10.1073/pnas.1323433111.
48. Aubin-Frankowski, P.C., and Vert, J.P. (2020). Gene regulation inference from single-cell RNA-seq data with linear differential equations and velocity inference. Bioinformatics *36*, 4774-4780. 10.1093/bioinformatics/btaa576.

49. Aibar, S., Gonzalez-Blas, C.B., Moerman, T., Van, A.H.T., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory network inference and clustering. Nat Methods *14*, 1083-1086. 10.1038/Nmeth.4463.
50. Bocci, F., Zhou, P.J., and Nie, Q. (2022). spliceJAC: transition genes and state-specific gene regulation from single-cell transcriptome data. Mol Syst Biol *18*, e11176. 10.15252/msb.202211176.
51. Chan, T.E., Stumpf, M.P.H., and Babtie, A.C. (2017). Gene Regulatory Network Inference from Single-Cell Data Using Multivariate Information Measures. Cell Syst *5*, 251-267. 10.1016/j.cels.2017.08.014.
52. Chen, S.N., and Mar, J.C. (2018). Evaluating methods of inferring gene regulatory networks highlights their lack of performance for single cell gene expression data. Bmc Bioinformatics *19*, 232. 10.1186/s12859-018-2217-z.
53. Dong, X., Tang, K., Xu, Y.F., Wei, H.L., Han, T., and Wang, C.F. (2022). Single-cell gene regulation network inference by large-scale data integration. Nucleic Acids Res *50*, e126. 10.1093/nar/gkac819.
54. Gupta, A., Martin-Rufino, J.D., Jones, T.R., Subramanian, V., Qiu, X.J., Grody, E.I., Bloemendal, A., Weng, C., Niu, S.Y., Min, K.H., et al. (2022). Inferring gene regulation from stochastic transcriptional variation across single cells at steady state. P Natl Acad Sci USA *119*, e2207392119. 10.1073/pnas.2207392119.
55. Yuan, Z., Riera, A., Bai, L., Sun, J., Nandi, S., Spanos, C., Chen, Z.A., Barbon, M., Rappsilber, J., Stillman, B., et al. (2017). Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6 and Cdt1. Nat Struct Mol Biol *24*, 316-324. 10.1038/nsmb.3372.
56. Wei, L., and Zhao, X.L. (2016). A new MCM modification cycle regulates DNA replication initiation. Nature Structural & Molecular Biology *23*, 209-216. 10.1038/nsmb.3173.
57. Bertran-Alamillo, J., Cattan, V., Schoumacher, M., Codony-Servat, J., Gimenez-Capitan, A., Cantero, F., Burbridge, M., Rodriguez, S., Teixido, C., Roman, R., et al. (2019). AURKB as a target in non-small cell lung cancer with acquired resistance to anti-EGFR therapy. Nat Commun *10*, 1812. 10.1038/s41467-019-09734-5.
58. Basu, S., Greenwood, J., Jones, A.W., and Nurse, P. (2022). Core control principles of the eukaryotic cell cycle. Nature *607*, 381-386. 10.1038/s41586-022-04798-8.
59. Boutros, R., Lobjois, V., and Ducommun, B. (2007). CDC25 phosphatases in cancer cells: key players? Good targets? Nat Rev Cancer *7*, 495-507. 10.1038/nrc2169.
60. Hou, S., Li, N., Zhang, Q., Li, H., Wei, X.Y., Hao, T., Li, Y., Azam, S., Liu, C.G., Cheng, W., et al. (2016). XAB2 functions in mitotic cell cycle progression via transcriptional regulation of CENPE. Cell Death Dis *7*, e2409. 10.1038/cddis.2016.313.
61. Pedersen, R.S., Karemore, G., Gudjonsson, T., Rask, M.B., Neumann, B., Heriche, J.K., Pepperkok, R., Ellenberg, J., Gerlich, D.W., Lukas, J., and Lukas, C. (2016). Profiling DNA damage response following mitotic perturbations. Nat Commun *7*, 13887. 10.1038/ncomms13887.
62. Paeschke, K., Bochman, M.L., Garcia, P.D., Cejka, P., Friedman, K.L., Kowalczykowski, S.C., and Zakian, V.A. (2013). Pif1 family helicases suppress genome instability at G-quadruplex motifs. Nature *497*, 458-462. 10.1038/nature12149.
63. Wadsworth, P. (2015). TPX2. Curr Biol *25*, R1156-R1158. 10.1016/j.cub.2015.10.003.
64. Maurice, D.H., Ke, H.M., Ahmad, F., Wang, Y.S., Chung, J., and Manganiello, V.C. (2014).

Advances in targeting cyclic nucleotide phosphodiesterases. Nat Rev Drug Discov *13*, 290-314. 10.1038/nrd4228.

65. Geng, A., Qiu, R., Murai, K., Liu, J., Wu, X., Zhang, H., Farhoodi, H., Duong, N., Jiang, M., Yee, J.K., et al. (2018). KIF20A/MKLP2 regulates the division modes of neural progenitor cells during cortical development. Nat Commun *9*, 2707. 10.1038/s41467-018-05152-1.
66. Tsai, T.Y.C., Choi, Y.S., Ma, W.Z., Pomerening, J.R., Tang, C., and Ferrell, J.E. (2008). Robust, tunable biological oscillations from interlinked positive and negative feedback loops. Science *321*, 126-129. 10.1126/science.1156951.
67. Xie, M.Q., and Fussenegger, M. (2018). Designing cell function: assembly of synthetic gene circuits for cell biology applications. Nat Rev Mol Cell Bio *19*, 507-525. 10.1038/s41580-018-0024-z.
68. Xu, L., Patterson, D., Staver, A.C., Levin, S.A., and Wang, J. (2021). Unifying deterministic and stochastic ecological dynamics via a landscape-flux approach. P Natl Acad Sci USA *118*, e2103779118. 10.1073/pnas.2103779118.
69. Xu, L., Patterson, D., Levin, S.A., and Wang, J. (2023). Non-equilibrium early-warning signals for critical transitions in ecological systems. Proc Natl Acad Sci U S A *120*, e2218663120. 10.1073/pnas.2218663120.
70. Vasan, N., Baselga, J., and Hyman, D.M. (2019). A view on drug resistance in cancer. Nature *575*, 299-309. 10.1038/s41586-019-1730-1.
71. Fang, X.N., Liu, Q., Bohrer, C., Hensel, Z., Han, W., Wang, J., and Xiao, J. (2018). Cell fate potentials and switching kinetics uncovered in a classic bistable genetic switch. Nat Commun *9*, 2787. 10.1038/s41467-018-05071-1.
72. Jiang, Z.L., Tian, L., Fang, X.N., Zhang, K., Liu, Q., Dong, Q.Z., Wang, E.K., and Wang, J. (2019). The emergence of the two cell fates and their associated switching for a negative auto-regulating gene. Bmc Biol *17*, 49. 10.1186/s12915-019-0666-0.
73. Sheng, L., Qiong, L., Erkang, W., and Jin, W. (2023). Global quantitative understanding of nonequilibrium cell fate decision making in response to pheromone. bioRxiv, 2022.2007.2004.498707. 10.1101/2022.07.04.498707.
74. Fang, X.N., and Wang, J. (2020). Nonequilibrium Thermodynamics in Cell Biology: Extending Equilibrium Formalism to Cover Living Systems. Annual Review of Biophysics *49*, 227-246. 10.1146/annurev-biophys-121219-081656.
75. Vonbertalanffy, L. (1950). The Theory of Open Systems in Physics and Biology. Science *111*, 23-29. DOI 10.1126/science.111.2872.23.
76. Yang, X.B., Heinemann, M., Howard, J., Huber, G., Iyer-Biswas, S., Le Treut, G., Lynch, M., Montooth, K.L., Needleman, D.J., Pigolotti, S., et al. (2021). Physical bioenergetics: Energy fluxes, budgets, and constraints in cells. P Natl Acad Sci USA *118*, e2026786118. 10.1073/pnas.2026786118.
77. Fei, X., Xiang, L., Rui, W., Hong, Q., Jun, J., Zhilong, L., Yuning, W., Hai, L., Chuansheng, S., and Jianwei, S. (2023). Topological design principle for the robustness of necroptosis biphasic, emergent, and coexistent (BEC) dynamics. bioRxiv, 2023.2001.2023.525173. 10.1101/2023.01.23.525173.
78. Le Novere, N. (2015). Quantitative and logic modelling of molecular and gene networks. Nat Rev Genet *16*, 146-158. 10.1038/nrg3885.
79. Pratapa, A., Jalihal, A.P., Law, J.N., Bharadwaj, A., and Murali, T.M. (2020). Benchmarking

algorithms for gene regulatory network inference from single-cell transcriptomic data. Nat Methods *17*, 147-154. 10.1038/s41592-019-0690-6.

80. Li, S., Zhang, P., Chen, W., Ye, L., Brannan, K.W., Le, N.T., Abe, J.I., Cooke, J.P., and Wang, G. (2023). A relay velocity model infers cell-dependent RNA velocity. Nat Biotechnol. 10.1038/s41587-023-01728-5.
81. Vandereyken, K., Sifrim, A., Thienpont, B., and Voet, T. (2023). Methods and applications for single-cell and spatial multi-omics. Nat Rev Genet, 1-22. 10.1038/s41576-023-00580-2.
82. Ma, J.Y., Zhao, J., Tian, J.W., Bai, X., and Tu, Z.W. (2013). Regularized vector field learning with sparse approximation for mismatch removal. Pattern Recogn *46*, 3519-3532. 10.1016/j.patcog.2013.05.017.
83. Qian, H. (2002). Mesoscopic nonequilibrium thermodynamics of single macromolecules and dynamic entropy-energy compensation. Phys Rev E *65*, 016102. 10.1103/PhysRevE.65.016102.
84. Freĭdlin, M.I., Wentzell, A.D., and Szücs, J. (2012). Random perturbations of dynamical systems, Third Edition. Edition (Springer).
85. Qian, H., and Elson, E.L. (2004). Fluorescence correlation spectroscopy with high-order and dual-color correlation to probe nonequilibrium steady states. P Natl Acad Sci USA *101*, 2828-2833. 10.1073/pnas.0305962101.
86. Li, C.H., Wang, E.K., and Wang, J. (2011). Landscape, Flux, Correlation, Resonance, Coherence, Stability, and Key Network Wirings of Stochastic Circadian Oscillation. Biophys J *101*, 1335-1344. 10.1016/j.bpj.2011.08.012.
87. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504. 10.1101/gr.1239303.

Supplementary information

Supplementary Table S1-3 and Supplementary Figure S1–20.

Figures and legends

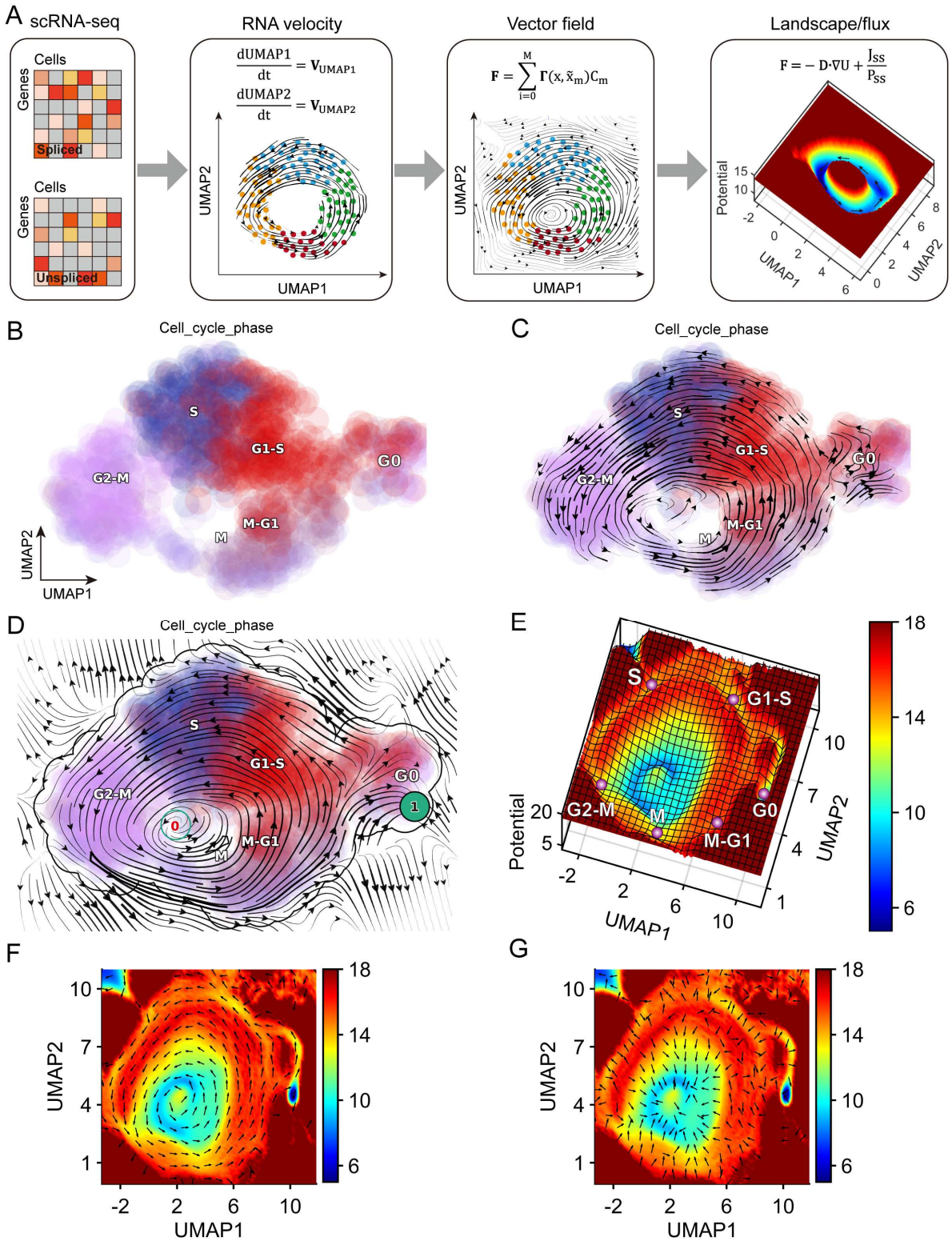

Figure 1. Workflow for analyzing cell cycle landscape-flux from single-cell transcriptomics data

(A) Workflow of constructing cell cycle landscape-flux by single-cell transcriptomics data.

(B) Cell cycle phase clusters by UMAP of U2OS cells.

(C) RNA velocity of cell cycle dynamics of U2OS cells.
(D) Reconstructed vector field of cell cycle dynamics of U2OS cells.
(E) Potential landscape of cell cycle dynamics in UMAP of U2OS cells.
(F) Curl flux (black arrow) of cell cycle dynamics landscape in UMAP of U2OS cells.
(G) The gradient force (black arrow) of cell cycle dynamics landscape in UMAP of U2OS cells.

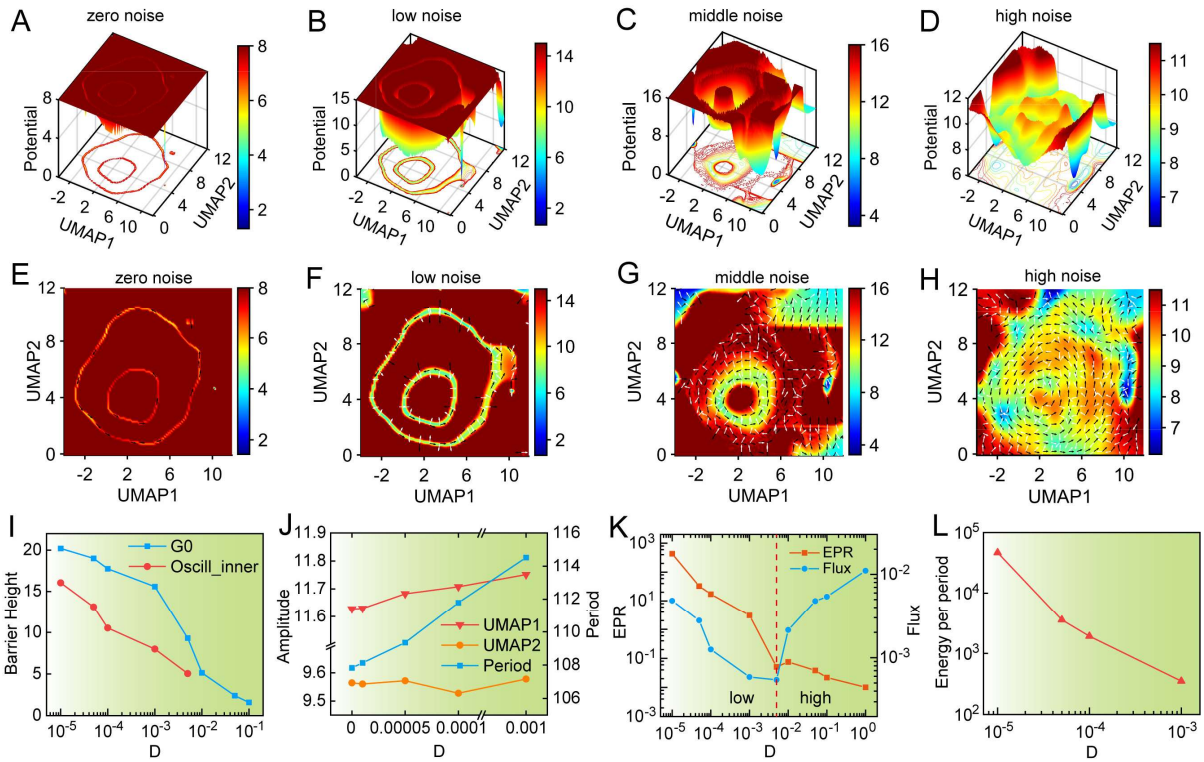

Figure 2. Cell cycle nonequilibrium landscape-flux affected by noise intensity

(A-D) Cell cycle global dynamics landscape of U2OS cells changes in UMAP when noise intensity is 0 (A), 0.00005 (B), 0.005 (C), 0.05 (D).

(E-H) Curl flux (black arrow) of cell cycle dynamics landscape in UMAP when noise intensity is 0 (E), 0.00005 (F), 0.005 (G), 0.05 (H).

(I) The change of Barrier Height of G0 phase (blue) and oscillation center (red) in landscape when the noise intensity is changed.

(J) The change of amplitude of UMAP1 and UMAP1 and period of cell cycle oscillatory dynamics when the noise intensity is changed.

(K) The change of EPR and Flux of cell cycle when the noise intensity is changed.

(L) The change of Energy per period of cell cycle when the noise intensity is changed.

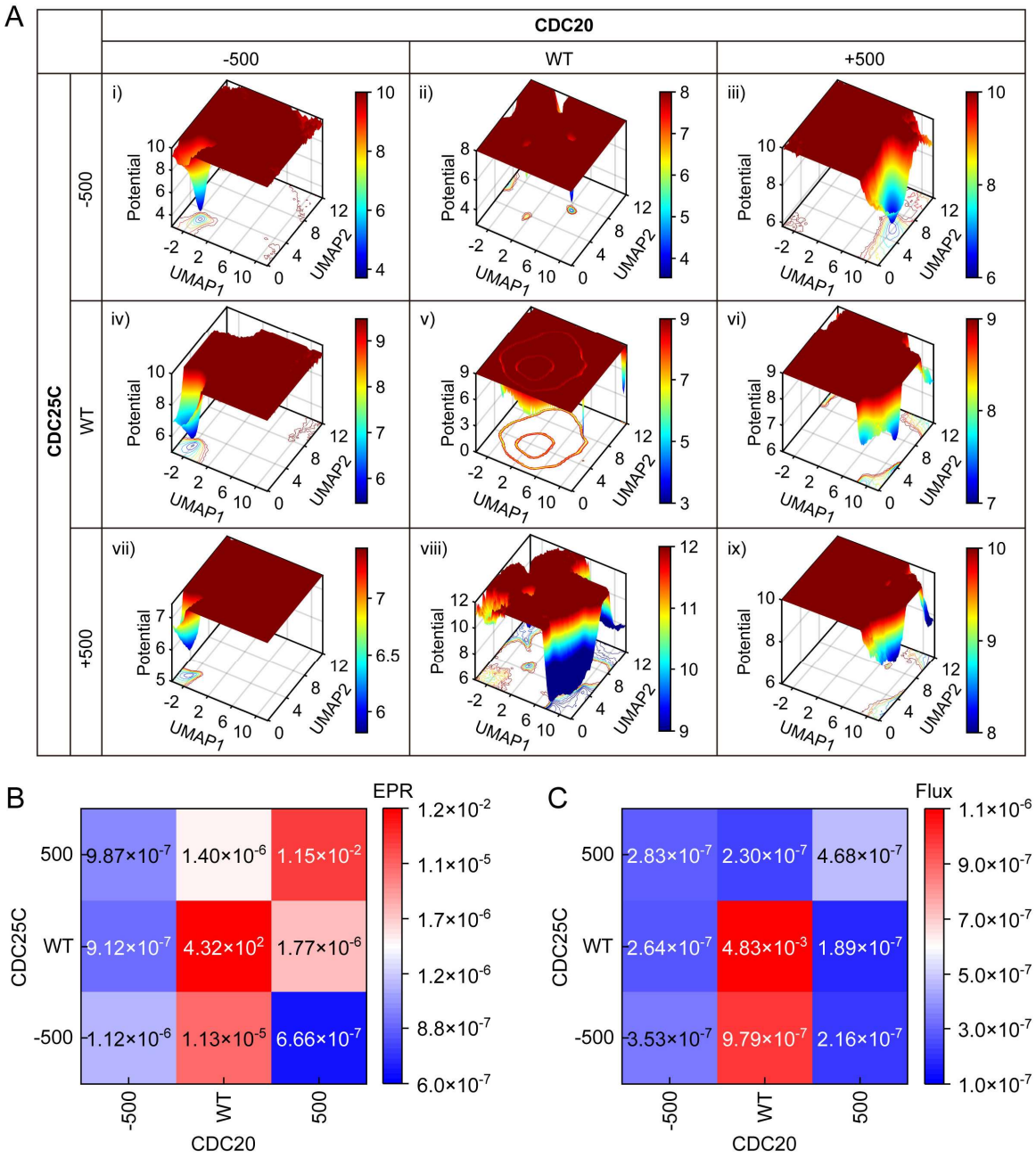

Figure 3. The genetic perturbation alters cell cycle landscape/flux

(A) Cell cycle global dynamics landscape of U2OS cells in UMAP with different genetic perturbation. Validation of in silico trajectory predictions. (i) Suppression of both CDC20 and CDC25C. (ii) Suppression of CDC25C only. (iii) Activation of CDC20 and suppression of CDC25C. (iv) Suppression of CDC20 only. (v) WT. (vi) Activation of CDC20 only. (vii) Suppression of CDC20 and activation of CDC25C. (viii) Activation of CDC25C only. (ix) Activation of both CDC20 and CDC25C.

(B) EPR of cell cycle nonequilibrium thermodynamics with different genetic perturbation.

(C) The average Flux of cell cycle nonequilibrium dynamics with different genetic perturbation.

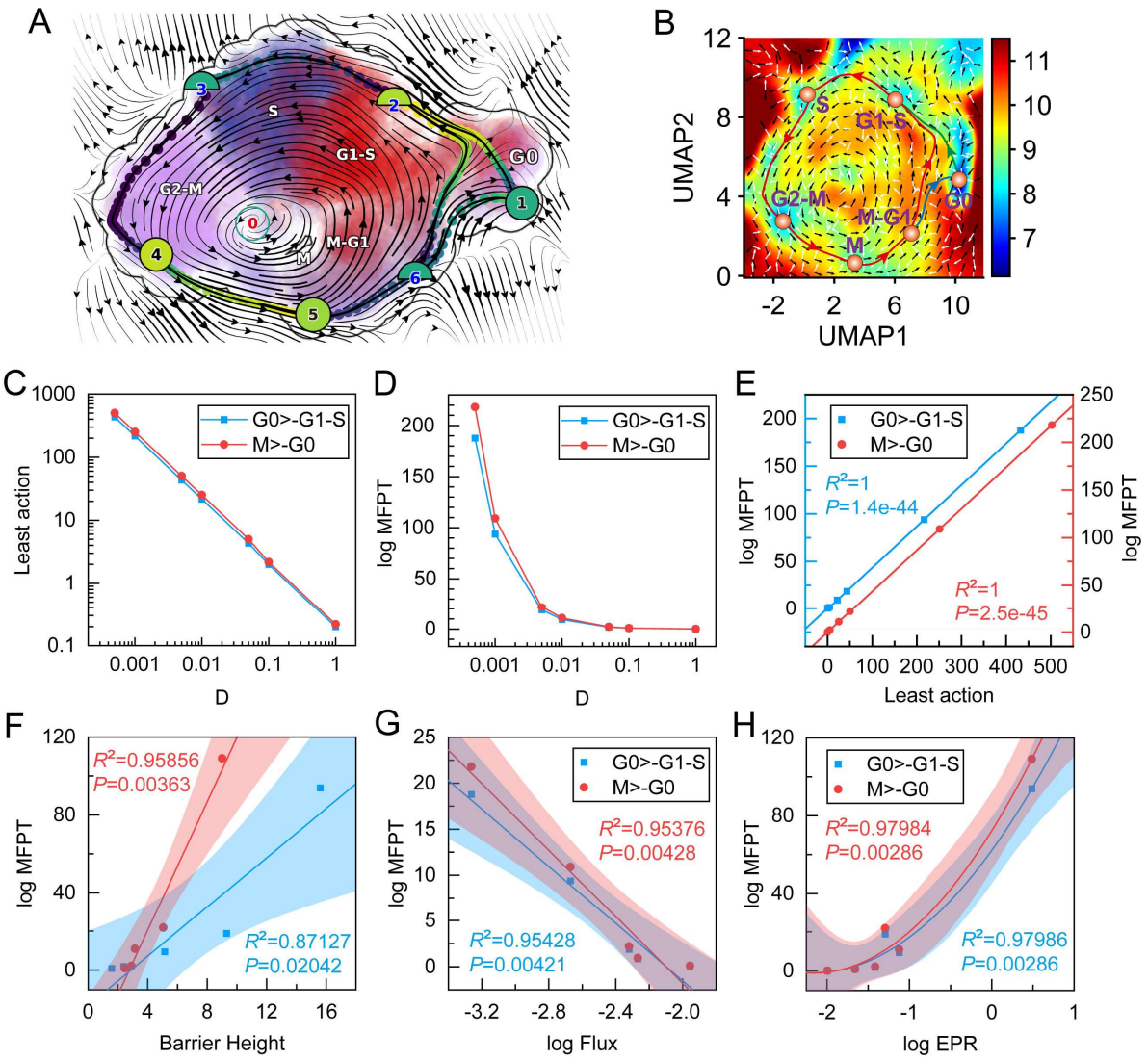

Figure 4. LAPs and MFPT in cell cycle initiation and termination

(A) LAPs between different cell cycle phases in the velocity vector field of U2OS cells.

(B) LAPs between different cell cycle phases in 2D landscape of U2OS cells. The black arrows present the curl flux and the white arrows present the gradient force in the landscape.

(C) The least action of G0 phase to G1-S phase and M phase to G0 phase versus diffusion coefficient D.

(D) The logarithm of MFPT of G0 phase to G1-S phase and M phase to G0 phase versus diffusion coefficient D.

(E) The correlation between the least action and the logarithm of MFPT when diffusion coefficients (fluctuations) are changed.

(F) The correlation between barrier height and the logarithm of MFPT when diffusion coefficients are changed. The line is a fitting correlation line, and the shaded part is the 95% fitting confidence interval.

(G) The correlation between flux and the logarithm of MFPT when diffusion coefficients are changed.

(H) The correlation between the logarithm of EPR and the logarithm of MFPT when diffusion coefficients are changed.

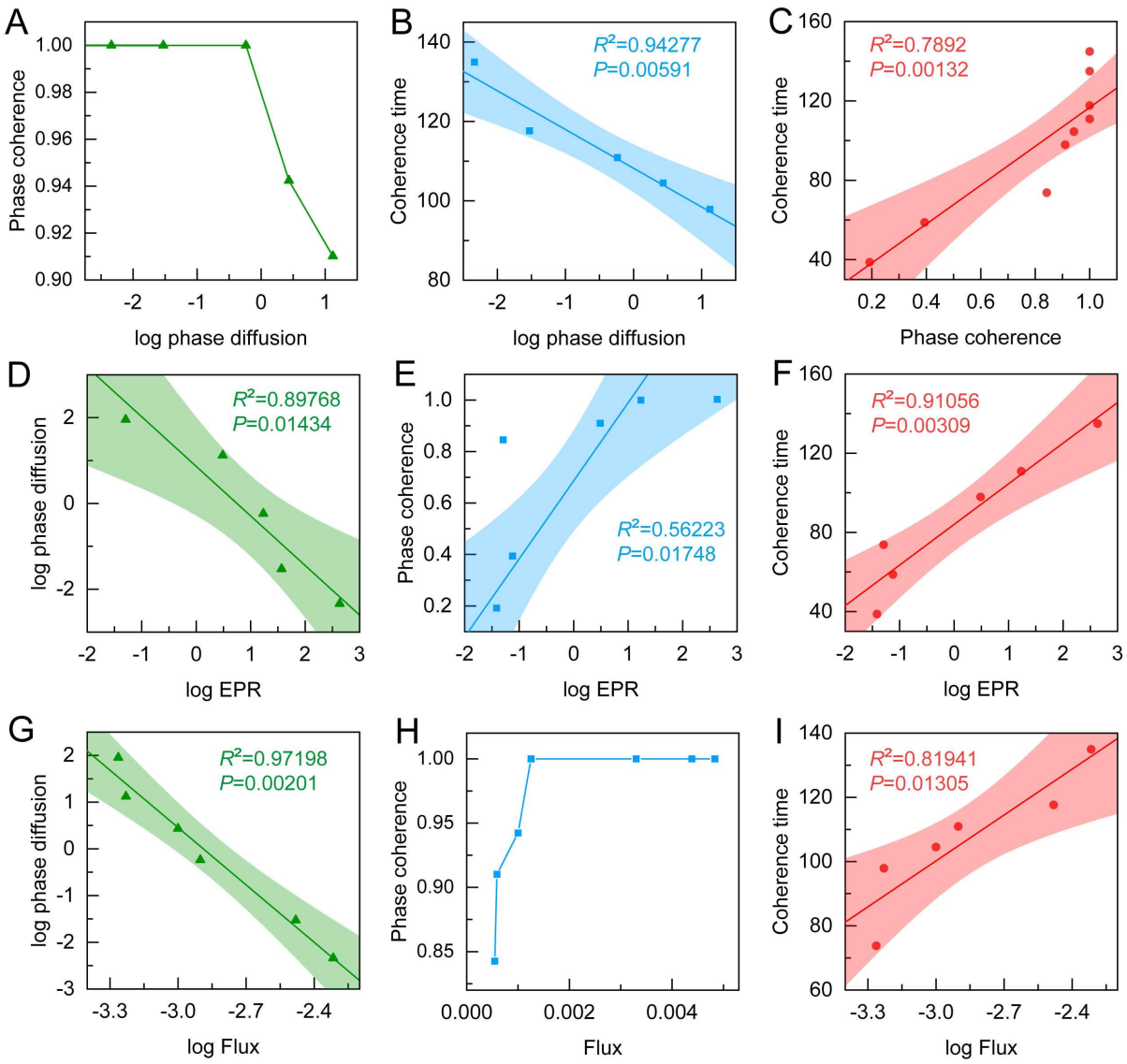

Figure 5. Relationship between phase diffusion, phase coherence, coherence time and EPR and flux

(A) Phase coherence versus the logarithm of phase diffusion.

(B) The correlation between coherence time and the logarithm of phase diffusion when diffusion coefficients (fluctuations) are changed. The line is a fitting correlation line, and the shaded part is the 95% fitting confidence interval.

(C) The correlation between coherence time and phase coherence when diffusion coefficients are changed.

(D) The correlation between the logarithm of phase diffusion and the logarithm of EPR when diffusion coefficients are changed.

(E) The correlation between phase coherence and the logarithm of EPR when diffusion coefficients are changed.

(F) The correlation between phase coherence and the logarithm of EPR when diffusion coefficients are changed.

(G) The correlation between the logarithm of phase diffusion and the logarithm of Flux when diffusion coefficients are changed.

(H) The change of phase coherence when Flux is changed.
(I) The correlation between coherence time and the logarithm of Flux when diffusion coefficients are changed.

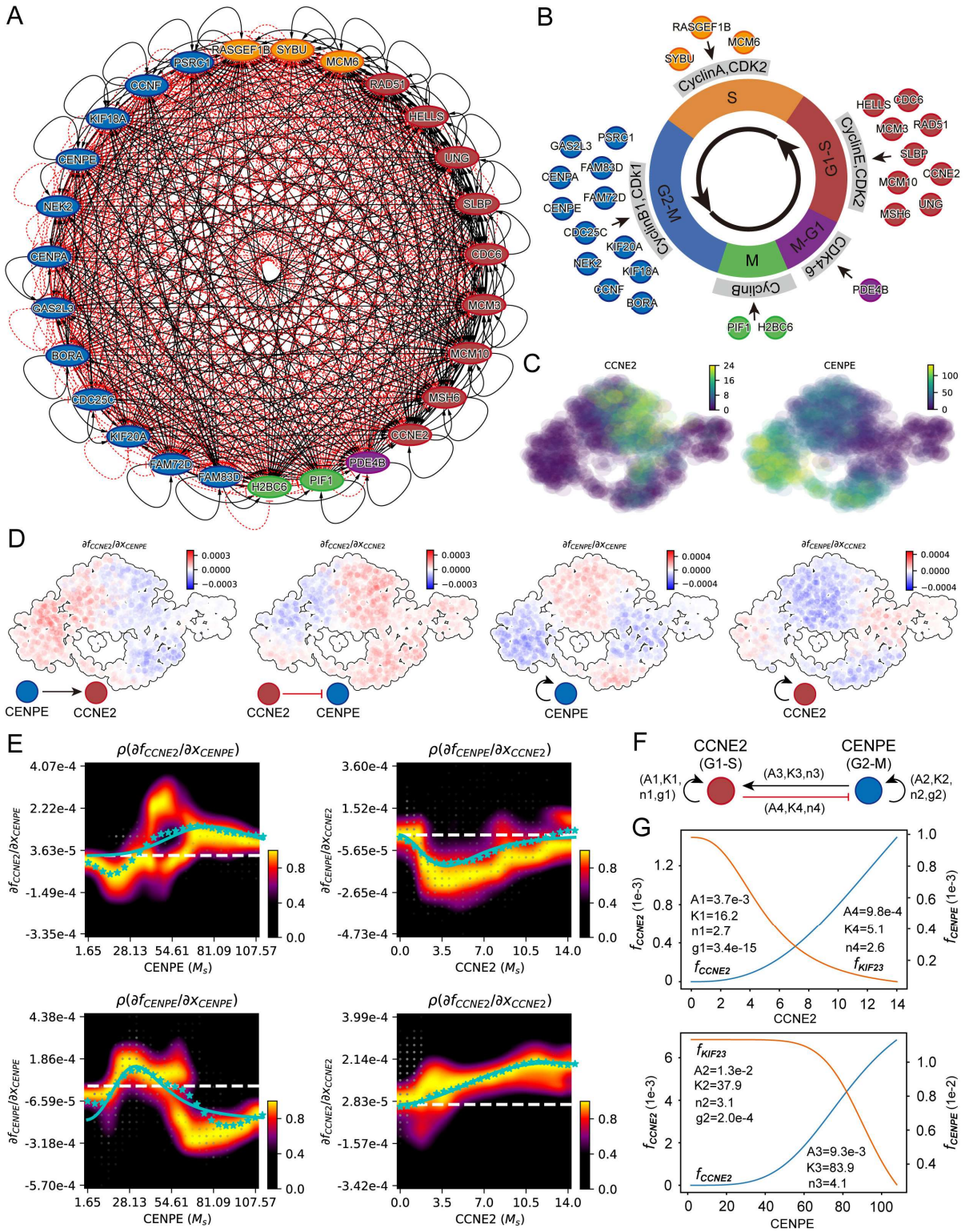

Figure 6. Inference of transcription factor regulatory networks for the cell cycle by single-cell transcriptomics data

(A) The interaction network of TFs of each cell cycle phase of U2OS cells. All differentially expressed TFs in each cell cycle phase are used to construct network; nodes (TFs) with more than one edge are shown. Node colors represent different cycle phases (red: G1-S, yellow: S, blue: G2-M, green: M, purple: M-G1), the black arrows represent the activation and the red arrows represent the inhibition.

(B) The diagram for the cell cycle model with key TFs of each cell cycle phase.
(C) CCNE2 has high expression in G1-S phase and CENPE has high expression in G2-M phase.
(D) Molecular mechanisms underlying the maintenance of the cell cycle. (i) CENPE activates CCNE2. (ii) Repression of CENPE by CCNE2. (iii) Self-activation of CENPE. (iv) Self-activation of CCNE2.
(E) Fitting the function of Jacobian versus gene expression with derivatives of a simplistic inhibitory or activation Hill equation. White dashed line corresponds to the zero Jacobian value. The blue stars at each x axis grid point correspond to the weighted mean of the Jacobian values for that point. The blue solid lines are the resultant fittings for the Jacobian.
(F) Schematic summarizing the interactions involving CCNE2 and CENPE.
(G) The velocity kinetic curves over gene expression changes of the corresponding fitted Hill equations of (E).

Supplementary Information

Quantifying landscape-flux via single-cell transcriptomics uncovers the underlying mechanism of cell cycle

Ligang Zhu, Jin Wang

Supplementary Tables

Table S1: Differential genes for U2OS cell cycle regulation.

test_group	gene	versus_group	exp_frac	log2_fc	pval	qval
G1-S	CCNE2	S	0.798898	1.023973	2.64E-16	1.33E-15
	MSH6	G2-M	0.997245	1.013831	1.13E-63	9.91E-62
	MCM10	G2-M	0.942149	1.043559	6.55E-37	5.05E-36
	MCM3	G2-M	0.997245	1.0391	1.85E-71	4.84E-69
	CDC6	G2-M	0.964187	1.027817	4.11E-61	1.54E-59
	SLBP	G2-M	0.983471	1.212273	5.23E-60	1.25E-58
	UNG	G2-M	0.972452	1.619185	2.23E-62	9.73E-61
	HELLS	G2-M	0.936639	1.178466	1.32E-24	8.06E-24
	RAD51	G2-M	0.895317	1.311513	9.65E-34	7.23E-33
S	MCM6	G2-M	0.974576	1.21473	5.60E-30	5.31E-29
	SYBU	G2-M	0.241525	1.409025	8.43E-03	1.11E-02
	RASGEF1B	G2-M	0.165254	1.180672	4.32E-02	4.97E-02
G2-M	PSRC1	G1-S	0.95	1.155866	1.27E-51	2.04E-50
	CCNF	G1-S	0.954167	1.222709	2.62E-47	3.31E-46
	KIF18A	G1-S	0.920833	1.06701	4.85E-35	3.80E-34
	FAM83D	G1-S	0.941667	1.360211	6.35E-49	8.77E-48
	CENPE	G1-S	0.920833	1.115475	4.32E-33	2.91E-32
	NEK2	G1-S	0.929167	1.095565	1.63E-35	1.31E-34
	CENPA	G1-S	0.9	1.013723	5.63E-33	3.71E-32
	GAS2L3	G1-S	0.904167	1.190276	3.53E-24	1.50E-23
	BORA	G1-S	0.904167	1.480186	4.58E-34	3.24E-33
	CDC25C	G1-S	0.883333	1.064069	7.03E-22	2.68E-21
	KIF20A	S	0.9125	1.027906	8.60E-33	5.42E-32
	FAM72D	S	0.516667	1.097402	1.03E-06	2.57E-06
M	PIF1	G1-S	0.792593	1.813076	5.90E-10	6.23E-10
	H2BC6	G2-M	0.37037	1.017269	2.59E-02	2.68E-02
M-G1	PDE4B	G2-M	0.27439	1.303967	7.50E-03	7.77E-03

Table S2: Differential genes for RPE1 cell cycle regulation.

test_group	gene	versus_group	exp_frac	log2_fc	pval	qval
G1-S	CCNE2	M-G1	0.486813	1.217845	6.65e-25	1.77e-24
S	H4C3	M-G1	0.932743	1.014606	1.87e-67	4.49e-66
	AURKB	M-G1	0.838938	1.000883	8.96e-39	7.16e-38
	CDK1	M-G1	0.725664	1.172389	8.75e-36	5.25e-35
G2-M	TOP2A	G1-S	0.939759	1.022381	4.68e-52	1.76e-50
	KIF23	G1-S	0.822289	1.216331	1.51e-30	1.22e-29
	HJURP	G1-S	0.734940	1.112817	8.81e-21	3.43e-20
	KPNA2	G1-S	0.849398	1.140699	2.22e-31	2.09e-30
M	VEGFA	G1-S	0.363395	1.066943	1.79e-06	2.81e-06
	TPX2	G1-S	0.904509	1.033970	1.80e-48	3.23e-47
	CKAP2	G1-S	0.694960	1.134864	1.20e-23	3.61e-23
	GAS2L3	G1-S	0.533156	1.090759	1.18e-12	2.23e-12
	CDC20	G1-S	0.867374	1.324284	1.96e-52	4.71e-51
M-G1	MATR3	G1-S	0.211823	1.241768	1.28e-07	2.99e-07
	KIF20A	G1-S	0.571429	1.108737	1.43e-23	4.99e-23
	CCNB1	G1-S	0.420361	1.109880	4.92e-14	1.29e-13

Table S3: The hill function and parameters of gene interaction related to Figure 6G and S18G.

ODEs	Interaction term	Parameters
The interaction between CCNE2 and CENPE		
$\frac{d[CCNE2]}{dt} = f_{CCNE2} = V_1 + V_2 - V_3$	Self-activation	A_1 =3.7e-3
	$V_1 = \frac{A_1[CCNE2]^{n1}}{(K_1 + [CCNE2]^{n1})}$	K_1 =16.2 n_1 =2.7
	CENPE activates CCNE2	A_3 =9.3e-3
	$V_2 = \frac{A_3[CENPE]^{n3}}{(K_3 + [CENPE]^{n3})}$	K_3 =83.9 n_3 =4.1
	Degradation	g_1 =3.4e-15
	$V_3 = g_1 \cdot [CCNE2]$	
$\frac{d[CENPE]}{dt} = f_{CENPE} = V_4 + V_5 - V_6$	Self-activation	A_2 =1.3e-2
	$V_4 = \frac{A_2[CENPE]^{n2}}{(K_2 + [CENPE]^{n2})}$	K_2 =37.9 n_2 =3.1
	CCNE2 inhibits CENPE	A_4 =9.8e-4
	$V_5 = \frac{A_4 K_4}{(K_4 + [CCNE2]^{n4})}$	K_4 =5.1 n_4 =2.6
	Degradation	g_2 =2.0e-4
	$V_6 = g_2 \cdot [CENPE]$	
The interaction between CCNE2 and KIF23		
$\frac{d[CCNE2]}{dt} = f_{CCN} \quad = V_1 + V_2 - V_3$	Self-activation	A_1 = 4.7e-4
	$V_1 = \frac{A_1[CCNE2]^{n1}}{(K_1 + [CCNE2]^{n1})}$	K_1 =2.3 n_1 =1.1
	KIF23 inhibits CCNE2	A_4 =3.2e-4
	$V_2 = \frac{A_4 K_4}{(K_4 + [KIF23]^{n4})}$	K_4 =0.9 n_4 =3.1
	Degradation	g_1 =1.3e-4
	$V_3 = g_1 \cdot [CCNE2]$	
$\frac{d[KIF23]}{dt} = f_{KIF23} = V_4 + V_5 - V_6$	Self-activation	A_2 =1.5e-4
	$V_4 = \frac{A_2[KIF23]^{n2}}{(K_2 + [KIF23]^{n2})}$	K_2 =1.7 n_2 =6.6
	CCNE2 activates KIF23	A_3 =4.1e-4
	$V_5 = \frac{A_3[CCNE2]^{n3}}{(K_3 + [CCNE2]^{n3})}$	K_3 =1.2 n_3 =2.4
	Degradation	g_2 =6.7e-27
	$V_6 = g_2 \cdot [KIF23]$	

Supplementary Figures

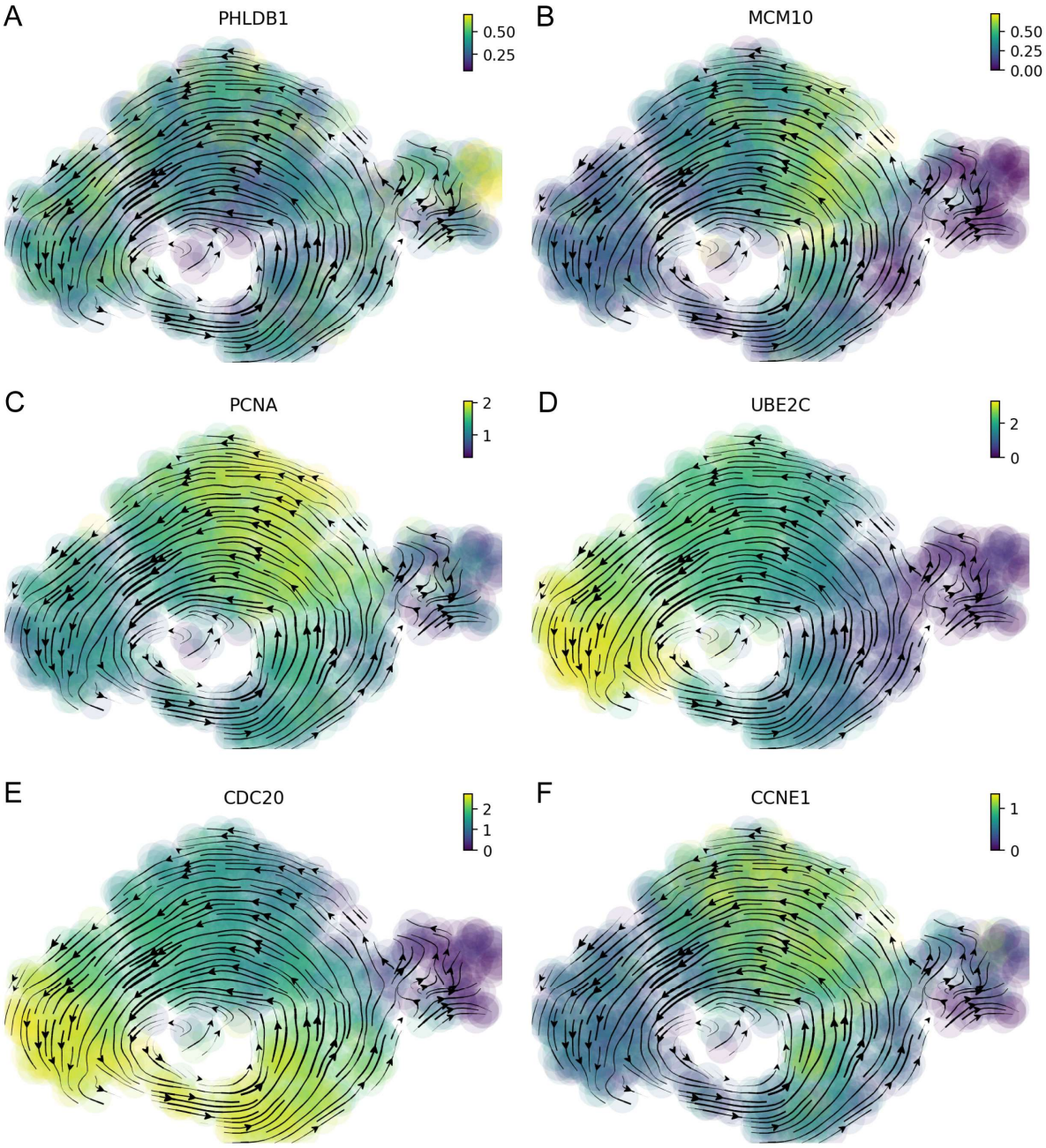

Figure S1. The key TF expression of different cell cycle phases, related to Figure 1

(A) PHLDB1 has high expression in the G0 phase.

(B) MCM10 has high expression in the G1-S phase.

(C) PCNA has high expression in the S phase.

(D) UBE2C has high expression in the G2-M phase.

(E) CDC20 has high expression in the M phase.

(F) CCNE1 has high expression in the M-G1 phase.

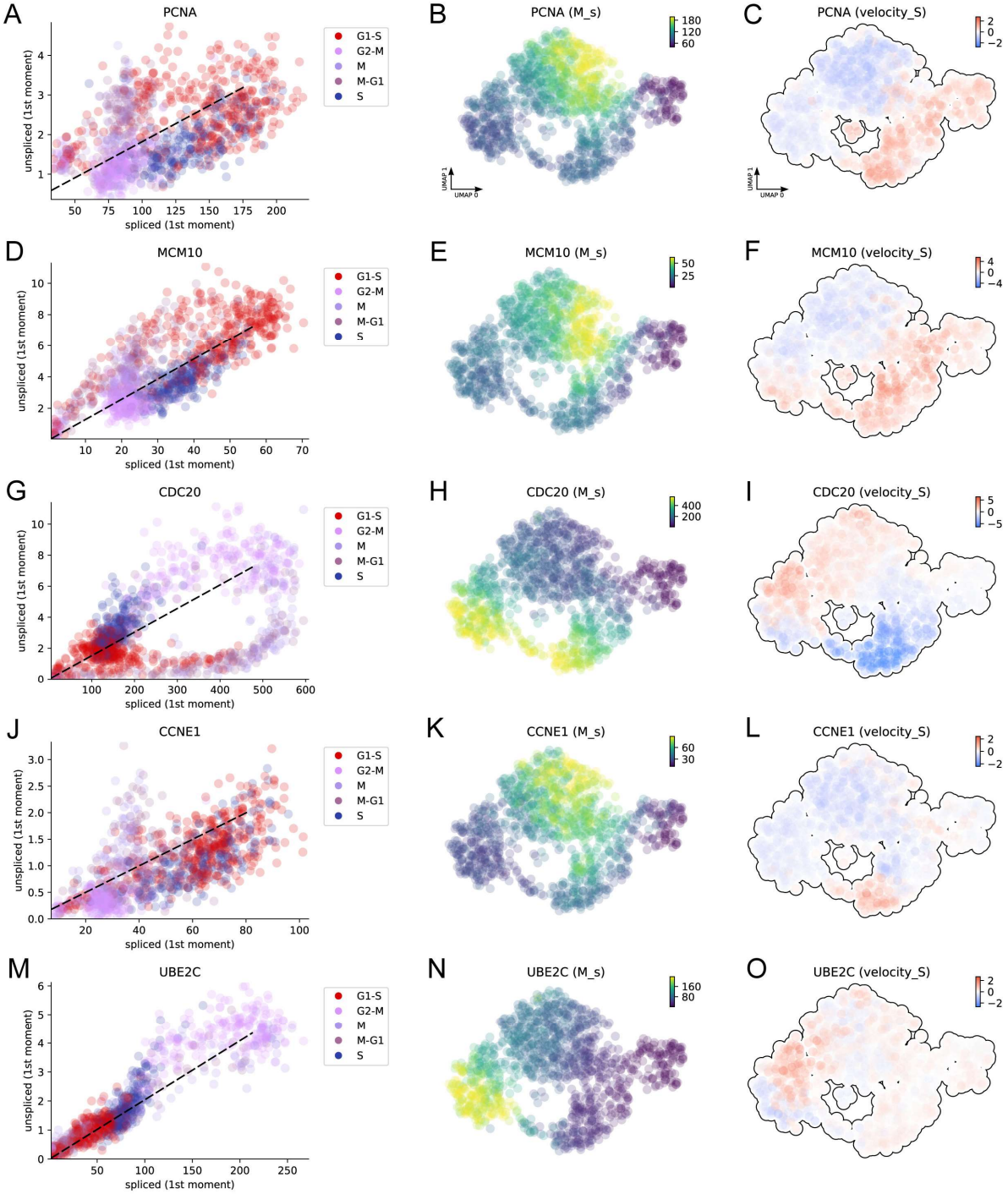

Figure S2. Phase portraits of different cell cycle dynamics, related to Figure 1

(A) Phase portraits between unspliced and spliced PCNA.

(B) The expression of spliced PCNA in different phases.

(C) Spliced RNA velocity of PCNA.

(D) Same as in (A) but for MCM10.

(E) Same as in (B) but for MCM10

(F) Same as in (C) but for MCM10

(G) Same as in (A) but for CDC20.

(H) Same as in (B) but for CDC20.

(I) Same as in (C) but for CDC20.

(J) Same as in (A) but for CCNE1.
(K) Same as in (B) but for CCNE1.
(L) Same as in (C) but for CCNE1.
(M) Same as in (A) but for UBE2C.
(N) Same as in (B) but for UBE2C.
(O) Same as in (C) but for UBE2C.

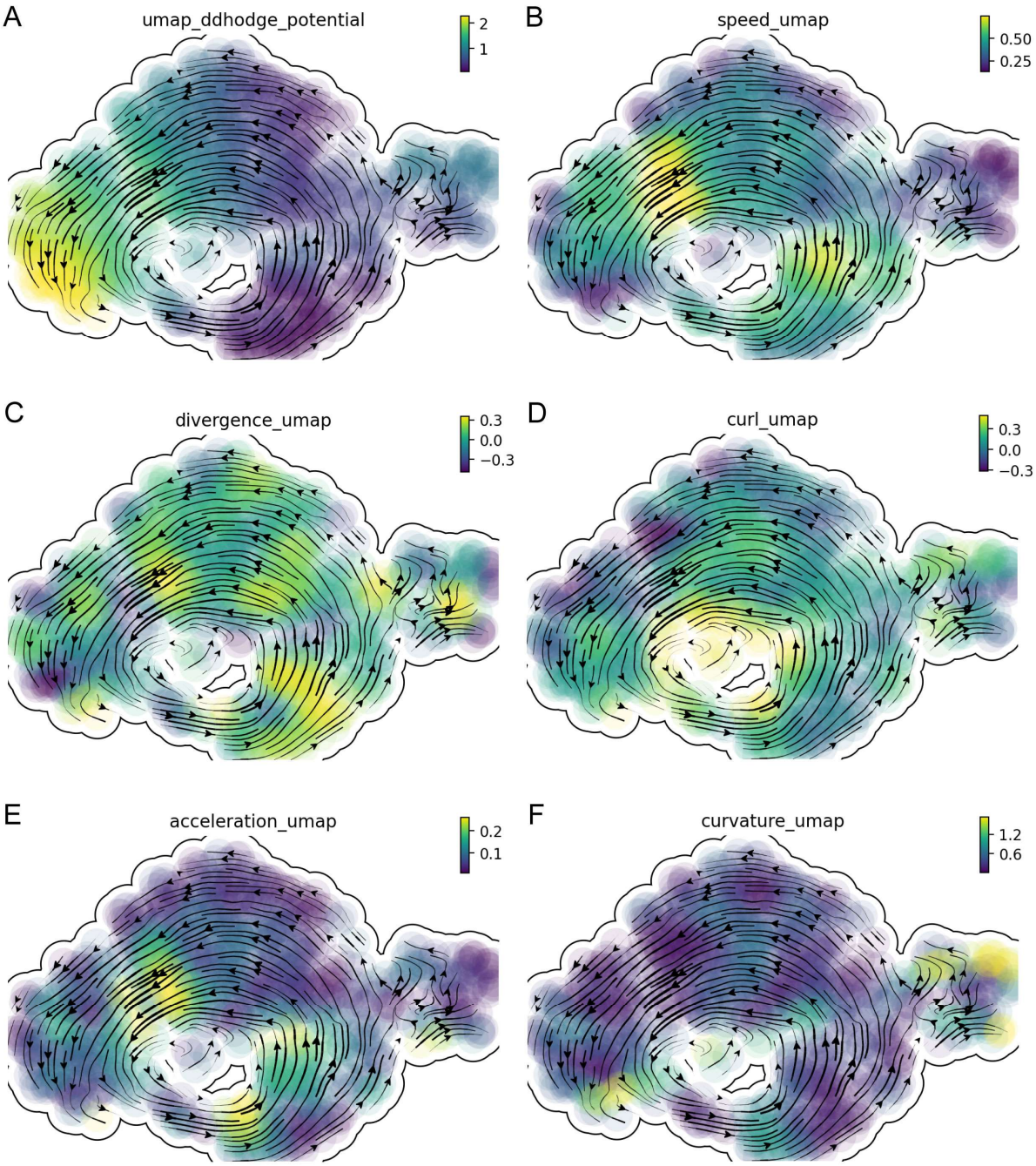

Figure S3. The differential geometric analysis of cell cycle dynamics of U2OS cells, related to Figure 1

(A) Ddhodge potential of the reconstructed vector field among all cell cycle phases.

(B) Same as in (A) but for the speed.

(C) Same as in (A) but for the divergence.

(D) Same as in (A) but for the curl.

(E) Same as in (A) but for the acceleration.

(F) Same as in (A) but for the curvature.

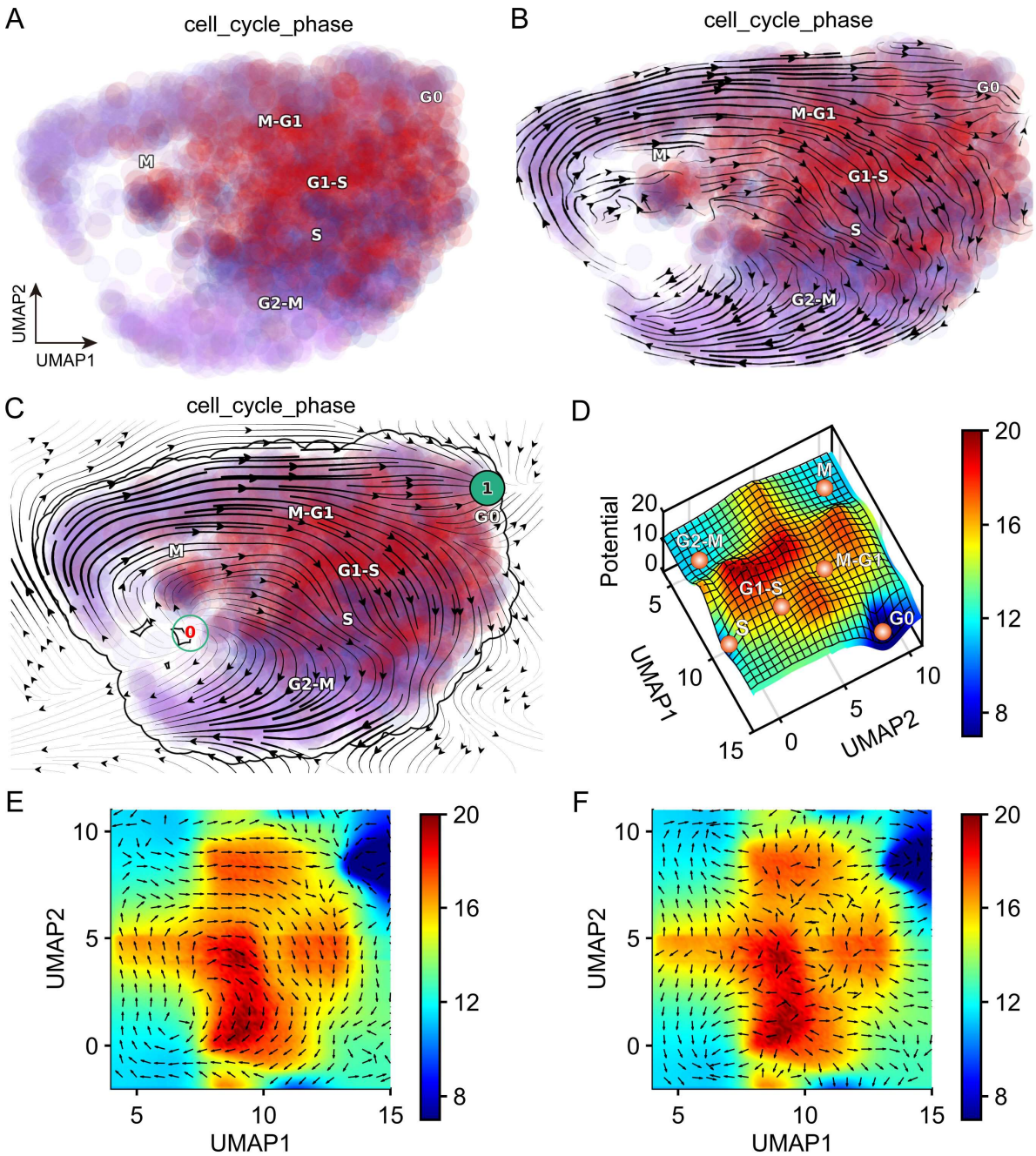

Figure S4. Constructing cell cycle landscape-flux by scEU-seq data of RPE1 cells, related to Figure 1

(A) Cell cycle phase clusters by UMAP.

(B) RNA velocity of cell cycle dynamics.

(C) Reconstructed vector field of cell cycle dynamics.

(D) Potential landscape of cell cycle dynamics in UMAP.

(E) Curl flux (black arrow) of cell cycle dynamics landscape in UMAP.

(F) The gradient force (black arrow) of cell cycle dynamics landscape in UMAP.

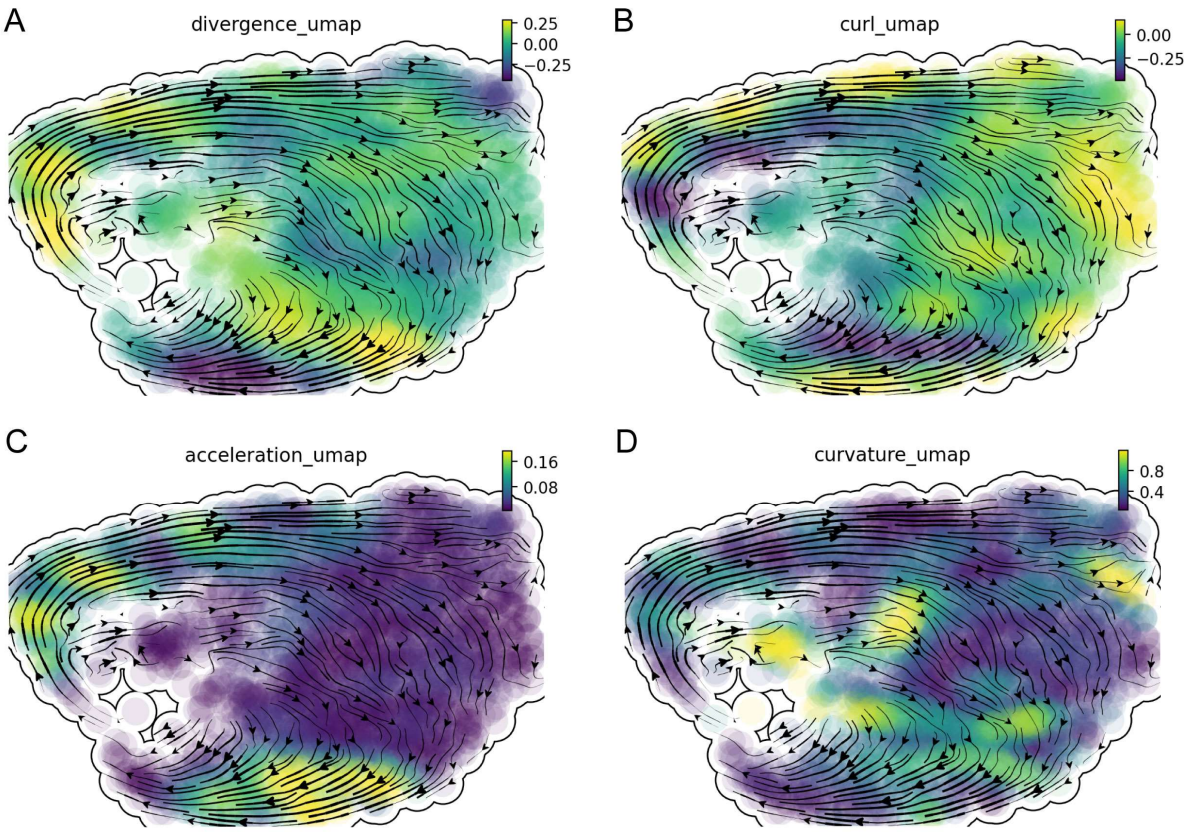

Figure S5. The differential geometric analysis of cell cycle dynamics of RPE1 cell, related to Figure 1

(A) Divergence of the reconstructed vector field among all cell cycle phases.

(B) Same as in (A) but for the curl.

(C) Same as in (A) but for the acceleration.

(D) Same as in (A) but for the curvature.

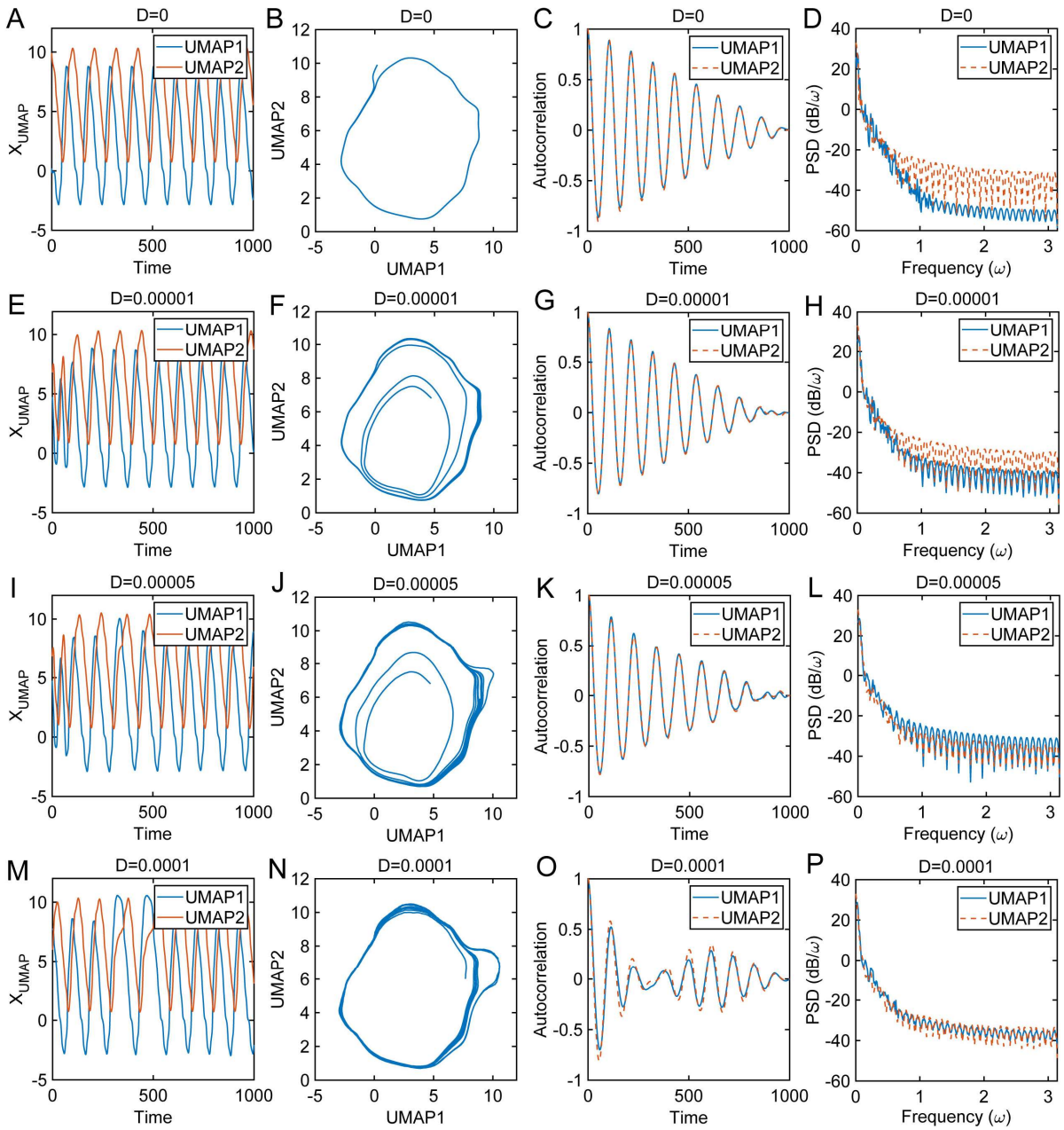

Figure S6. The trajectory, autocorrelation and power spectral density of cell cycle oscillation dynamics with different strength of noise, related to Figure 2

(A) The trajectories along time of the cell cycle for different diffusion coefficient D=0.

(B) The trajectories in phase space corresponding to (A) for different diffusion coefficient D=0.

(C) The autocorrelation corresponding to (A) for different diffusion coefficient D=0.

(D) The power spectral density corresponding to (A) for different diffusion coefficient D=0.

(E) Same as in (A) but for different diffusion coefficient D=0.00001.

(F) Same as in (B) but for different diffusion coefficient D=0.00001.

(G) Same as in (C) but for different diffusion coefficient D=0.00001.

(H) Same as in (D) but for different diffusion coefficient D=0.00001.

(I) Same as in (A) but for different diffusion coefficient D=0.00005.

(J) Same as in (B) but for different diffusion coefficient D=0.00005.

(K) Same as in (C) but for different diffusion coefficient D=0.00005.

(L) Same as in (D) but for different diffusion coefficient D=0.00005.
(M) Same as in (A) but for different diffusion coefficient D=0.0001.
(N) Same as in (B) but for different diffusion coefficient D=0.0001.
(O) Same as in (C) but for different diffusion coefficient D=0.0001.
(P) Same as in (D) but for different diffusion coefficient D=0.0001.

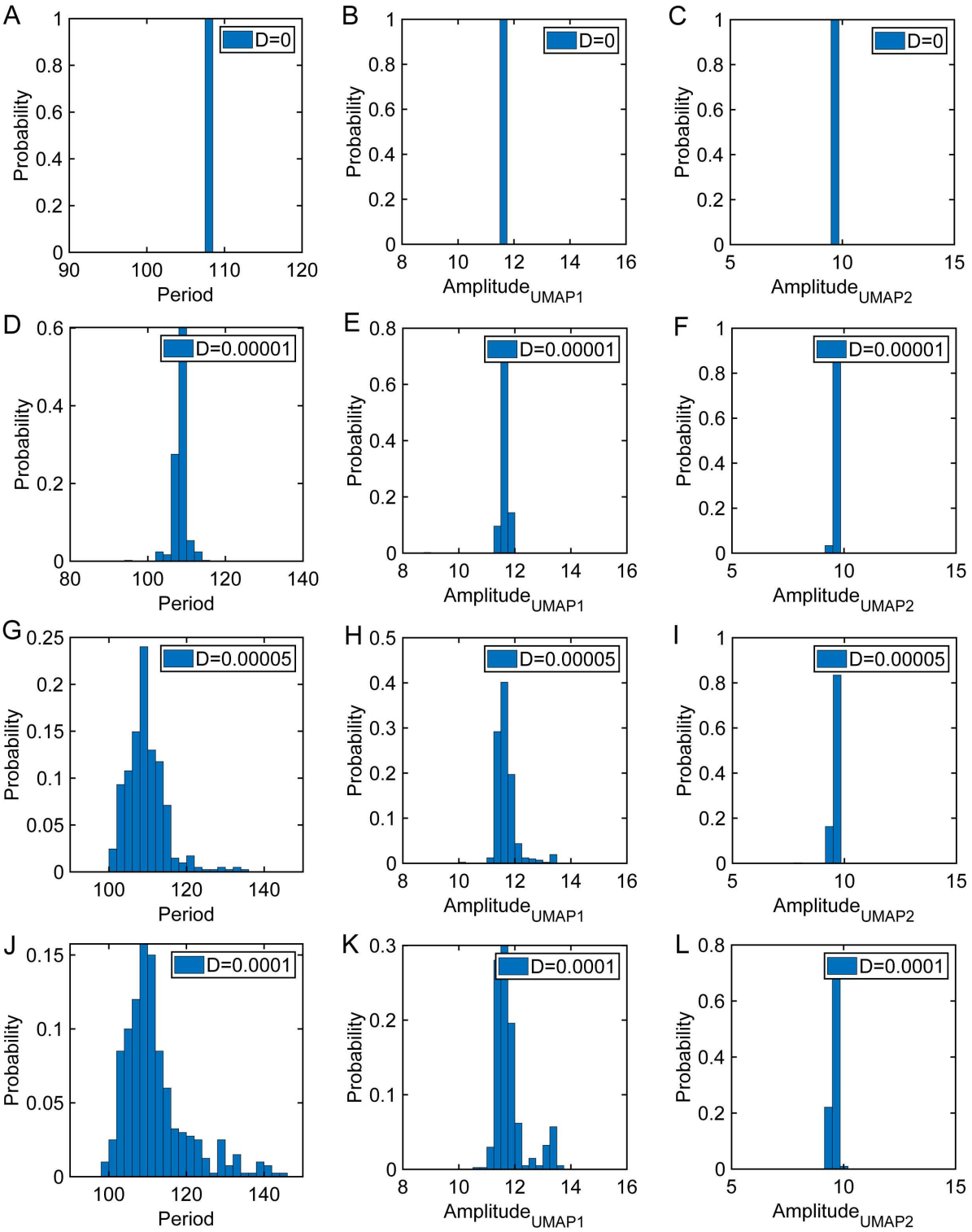

Figure S7. The period and amplitude distribution of cell cycle oscillation dynamics with different strength of noise, related to Figure 2

(A) The distribution of cell cycle period for different diffusion coefficient D=0.

(B) The distribution of amplitude in UMAP1 for different diffusion coefficient D=0.

(C) The distribution of amplitude in UMAP2 for different diffusion coefficient D=0.

(D) Same as in (A) but for different diffusion coefficient D=0.00001.

(E) Same as in (B) but for different diffusion coefficient D=0.00001.

(F) Same as in (C) but for different diffusion coefficient D=0.00001.
(G) Same as in (A) but for different diffusion coefficient D=0.00005.
(H) Same as in (B) but for different diffusion coefficient D=0.00005.
(I) Same as in (C) but for different diffusion coefficient D=0.00005.
(J) Same as in (A) but for different diffusion coefficient D=0.0001.
(K) Same as in (B) but for different diffusion coefficient D=0.0001.
(L) Same as in (C) but for different diffusion coefficient D=0.0001.

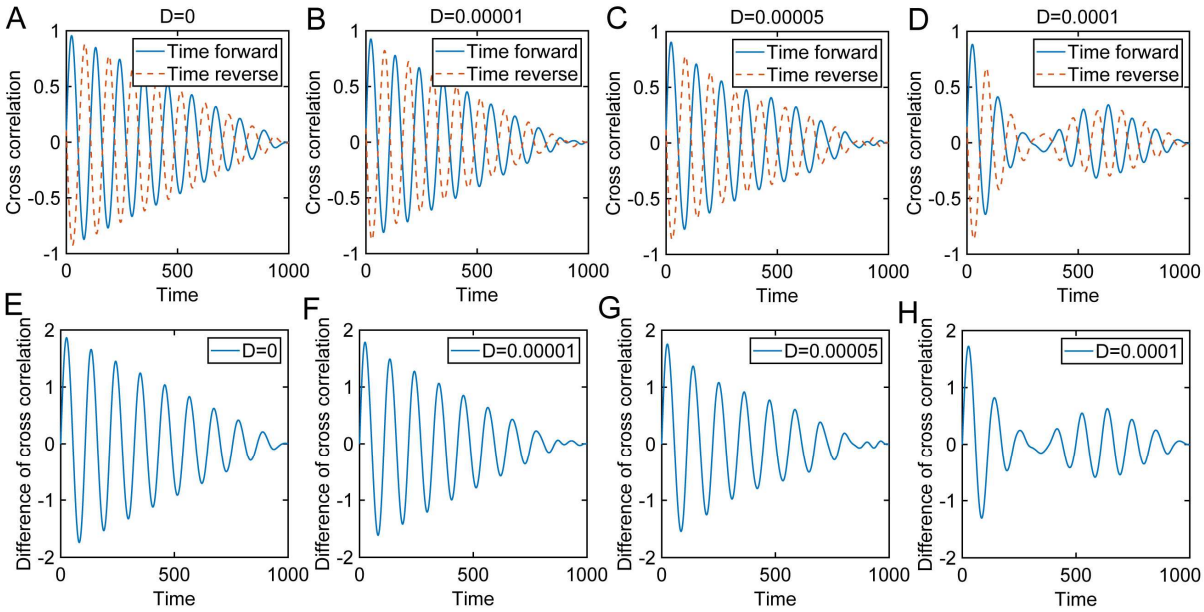

Figure S8. The cross-correlation function in time forward direction and time reverse direction and their difference at different noise of cell cycle dynamics, related to Figure 2

(A-D) The cross-correlation function in time forward direction and time reverse direction corresponding to Figure S6 for different diffusion coefficient D, D = 0 (A), D = 0.00001 (B), D = 0.00005 (C) and D = 0.0001 (D).

(E-H) The difference of cross correlation at different noise D, D = 0 (E), D = 0.00001 (F), D = 0.00005 (G) and D = 0.0001 (H).

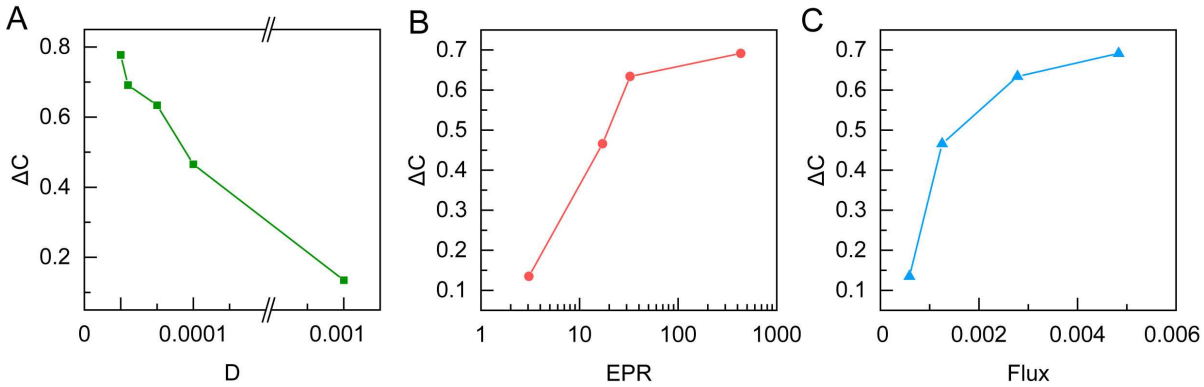

Figure S9. The average difference in cross-correlations between the forward in time and the backward in time with different strength of noise, related to Figure 2

(A) The average difference in cross-correlations between the forward in time and the backward in time (ΔC) with different diffusion coefficient.

(B) The change of ΔC when EPR is changed.

(C) The change of ΔC when Flux is changed.

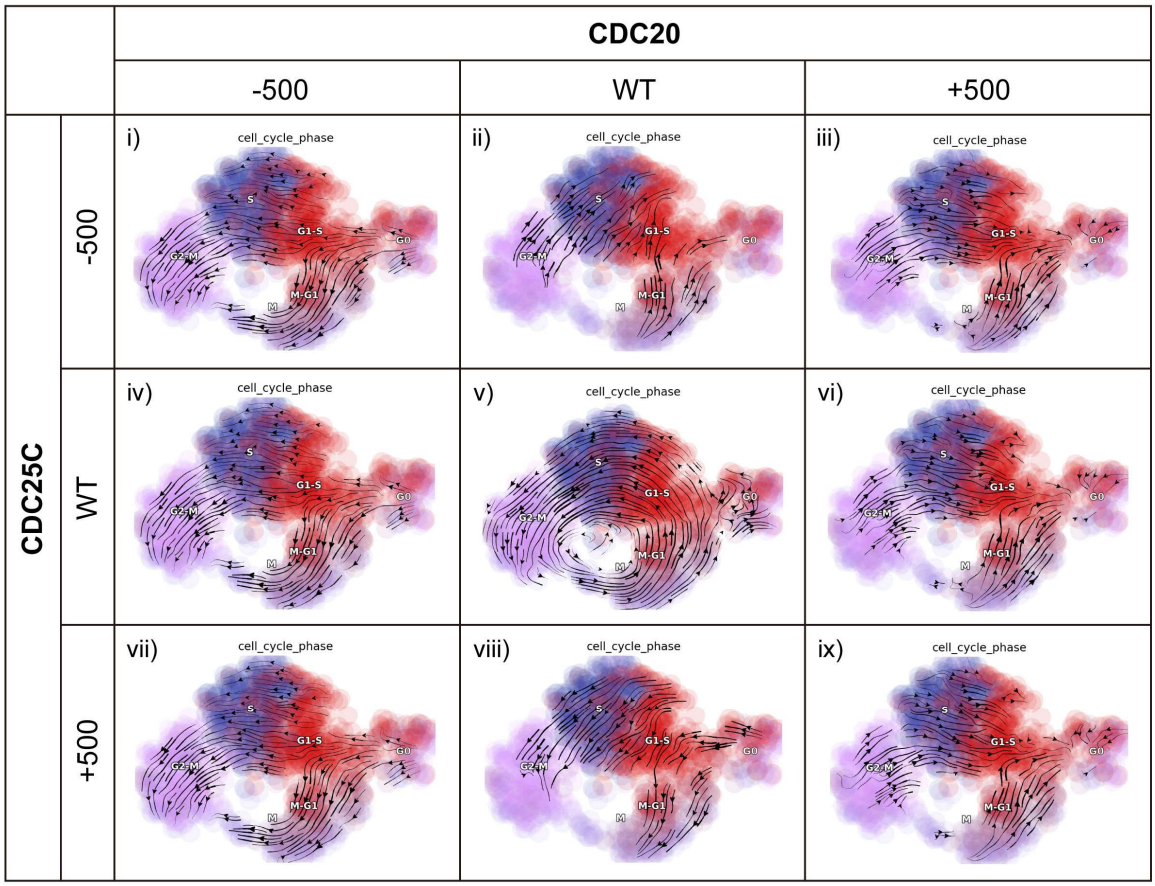

Figure S10. RNA velocity of cell cycle under genetic perturbation, related to Figure 3
In silico perturbation trajectory predictions. (i) Suppression of both CDC20 and CDC25C. (ii) Suppression of CDC25C only. (iii) Activation of CDC20 and suppression of CDC25C. (iv) Suppression of CDC20 only. (v) WT. (vi) Activation of CDC20 only. (vii) Suppression of CDC20 and activation of CDC25C. (viii) Activation of CDC25C only. (ix) Activation of both CDC20 and CDC25C.

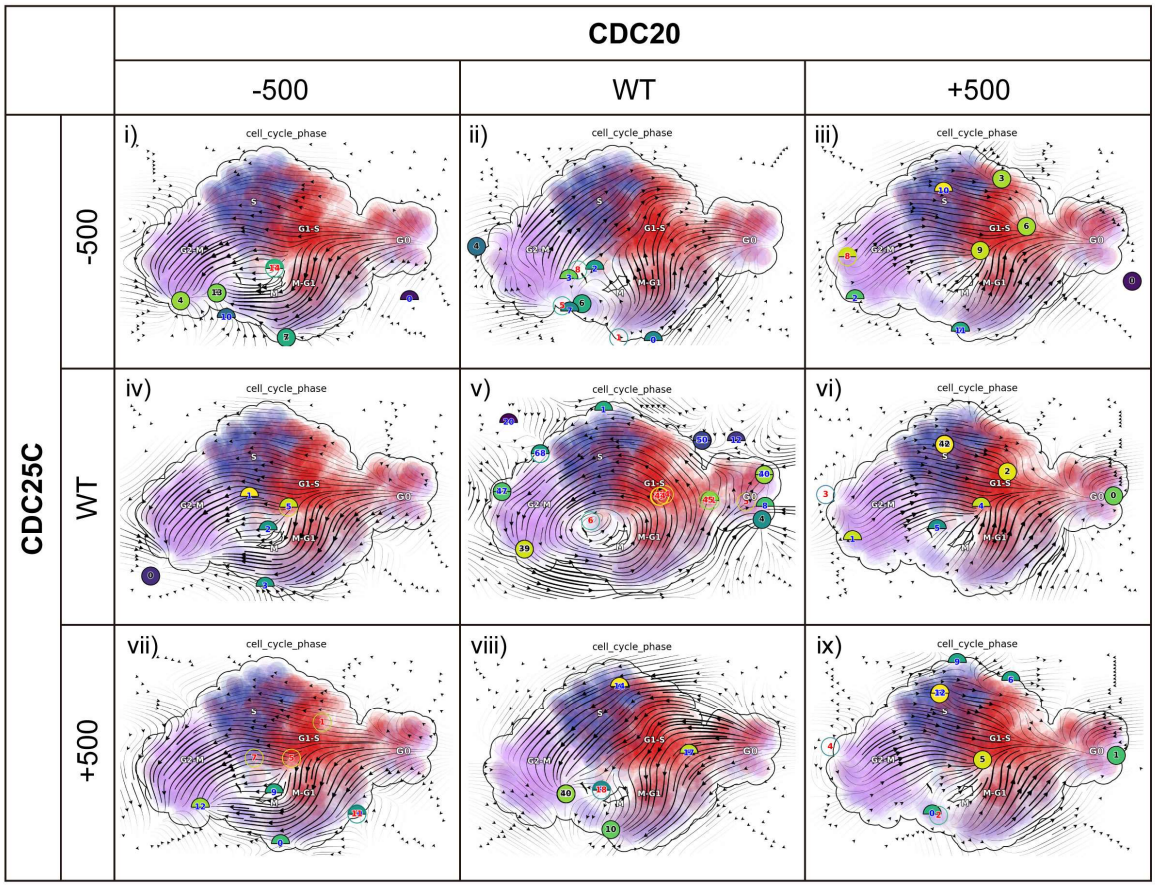

Figure S11. Reconstructed vector field of cell cycle under genetic perturbation, related to Figure 3

In silico perturbation vector field predictions. (i) Suppression of both CDC20 and CDC25C. (ii) Suppression of CDC25C only. (iii) Activation of CDC20 and suppression of CDC25C. (iv) Suppression of CDC20 only. (v) WT. (vi) Activation of CDC20 only. (vii) Suppression of CDC20 and activation of CDC25C. (viii) Activation of CDC25C only. (ix) Activation of both CDC20 and CDC25C.

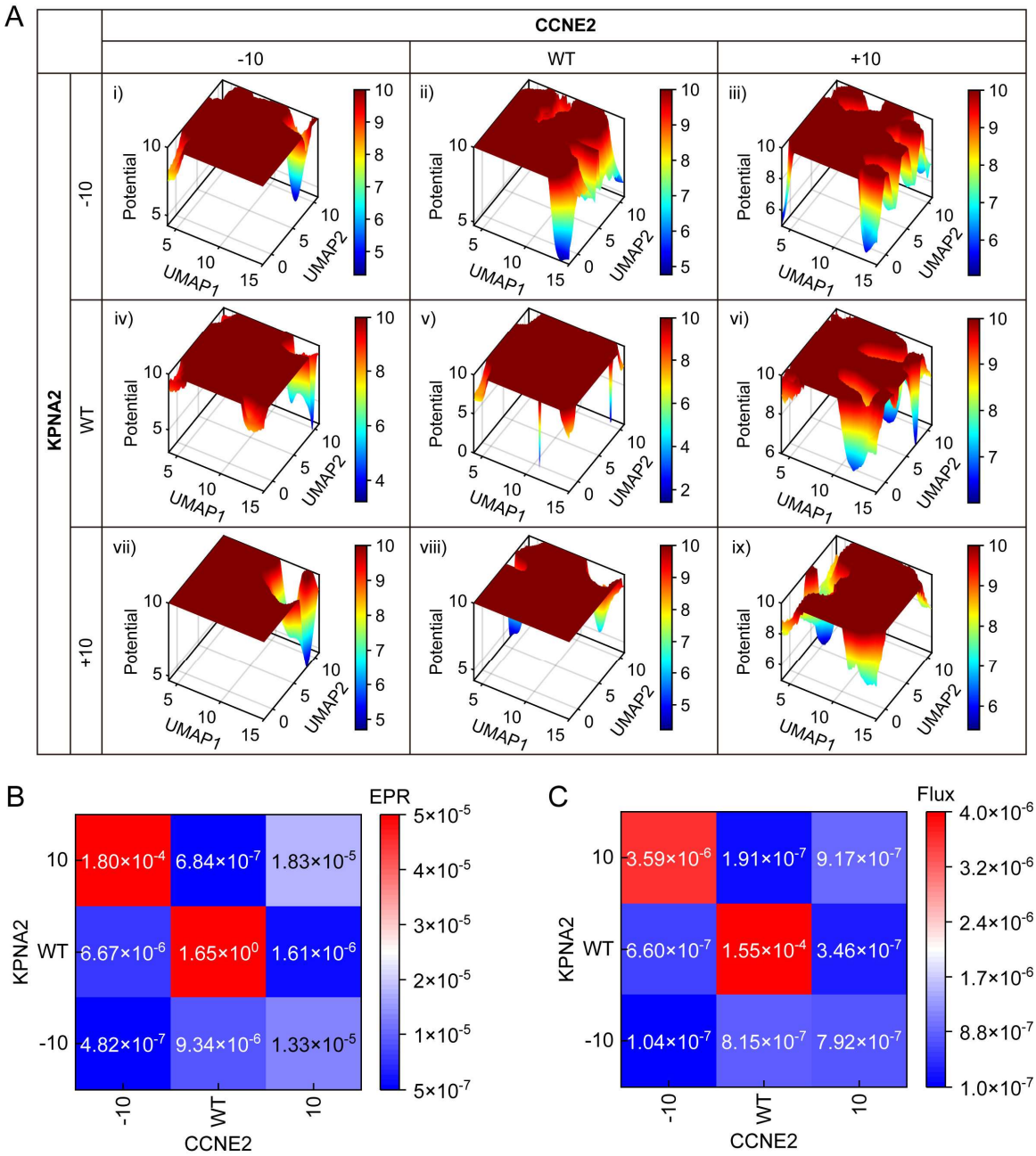

Figure S12. The genetic perturbation alters of RPE1 cell cycle landscape-flux, related to Figure 3

(A) Cell cycle global dynamics landscape of REP1 cells in UMAP with different genetic perturbation. (i) Suppression of both CCNE2 and KPNA2. (ii) Suppression of KPNA2 only. (iii) Activation of CCNE2 and suppression of KPNA2. (iv) Suppression of CCNE2 only. (v) WT. (vi) Activation of CCNE2 only. (vii) Suppression of CCNE2 and activation of KPNA2. (viii) Activation of KPNA2 only. (ix) Activation of both CCNE2 and KPNA2.

(B) EPR of cell cycle nonequilibrium thermodynamics with different genetic perturbation.

(C) The average Flux of cell cycle nonequilibrium dynamics with different genetic perturbation.

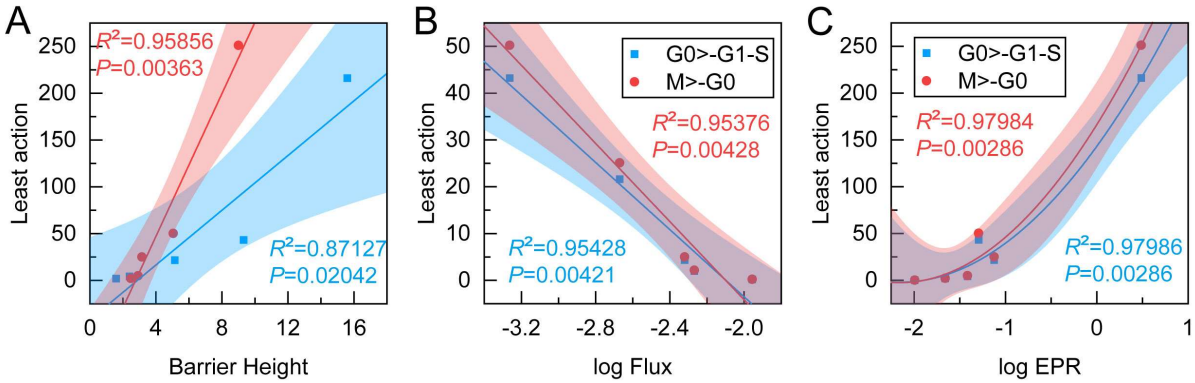

Figure S13. The correlation between least action and barrier height, flux and EPR, separately, related to Figure 4

(A) The correlation between barrier height and least action when diffusion coefficients are changed. The line is a fitting correlation line, and the shaded part is the 95% fitting confidence interval.

(B) The correlation between the logarithm of flux and least action when diffusion coefficients are changed.

(C) The correlation between the logarithm of EPR and least action when diffusion coefficients are changed.

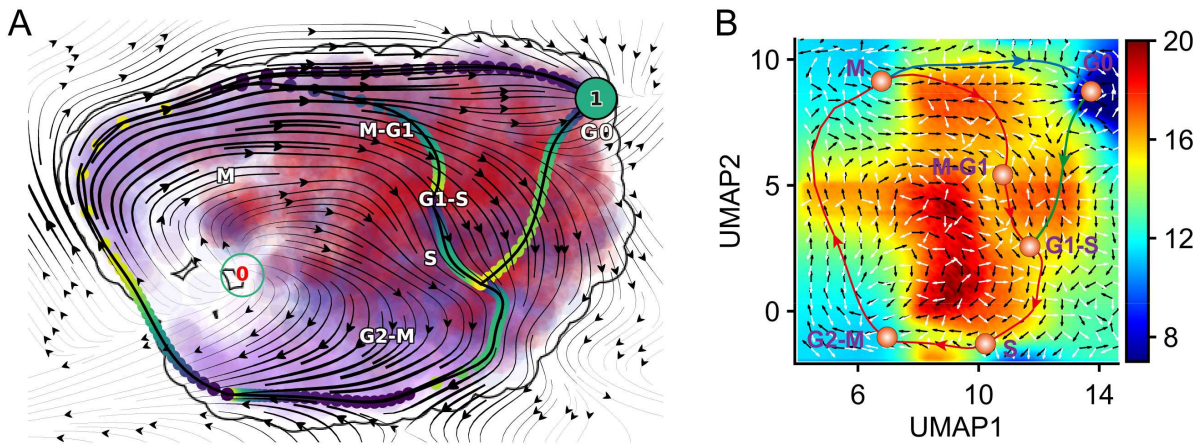

Figure S14. LAPs and MFPT in RPE1 cell cycle initiation and termination, related to Figure 4

(A) LAPs between different cell cycle phases in the velocity vector field of RPE1 cells.

(B) LAPs between different cell cycle phases in 2D landscape of RPE1 cells. The black arrows present the curl flux and the white arrows present the gradient force in the landscape.

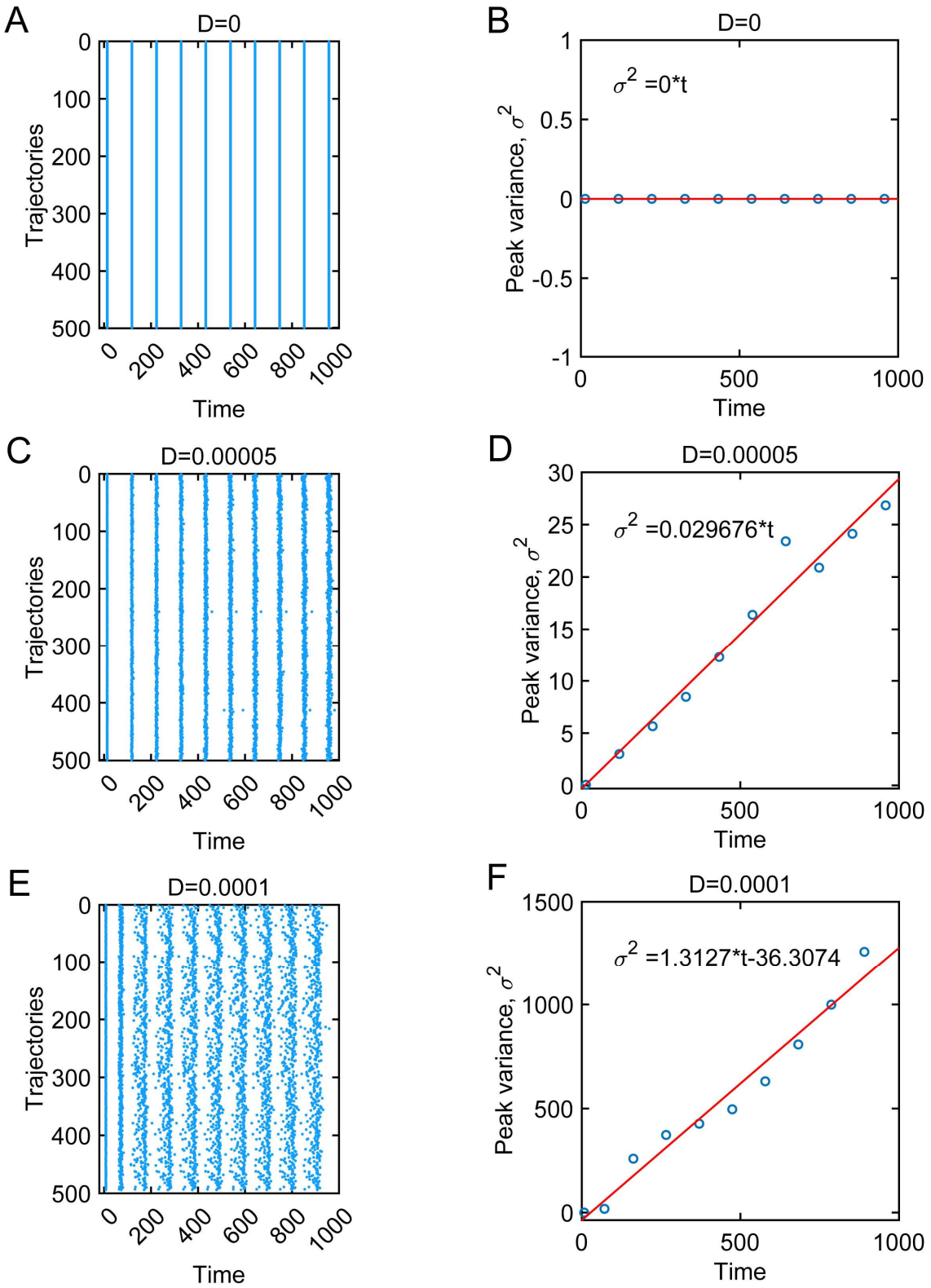

Figure S15. Phase diffusion of cell cycle oscillation with different noise, related to Figure 5

(A) Raster plot of the peak times for 500 different trajectories starting with the same initial condition at the diffusion coefficient D=0.

(B) Peak time variance σ^2 goes linearly with the average peak time for the diffusion coefficient D=0, with the linear coefficient defined as the peak-time diffusion constant.

(C) Same as in (A) but for the diffusion coefficient D=0.00005.
(D) Same as in (B) but for the diffusion coefficient D=0.00005.
(E) Same as in (A) but for the diffusion coefficient D=0.0001.
(F) Same as in (B) but for the diffusion coefficient D=0.0001.

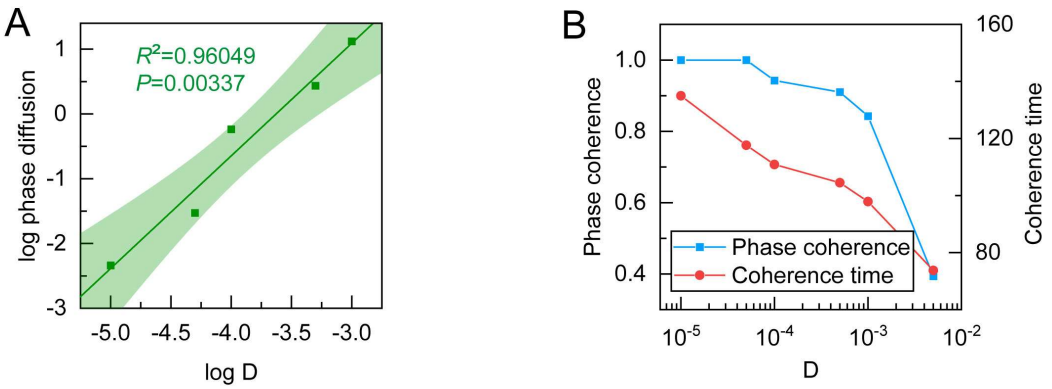

Figure S16. Relation between the phase properties of cell cycle oscillation and noisy fluctuations, related to Figure 5

(A) The logarithm of phase diffusion and the logarithm of diffusion coefficient D. The line is a fitting correlation line, and the shaded part is the 95% fitting confidence interval.

(B) The change of coherence time and phase coherence when diffusion coefficients (fluctuations) are changed.

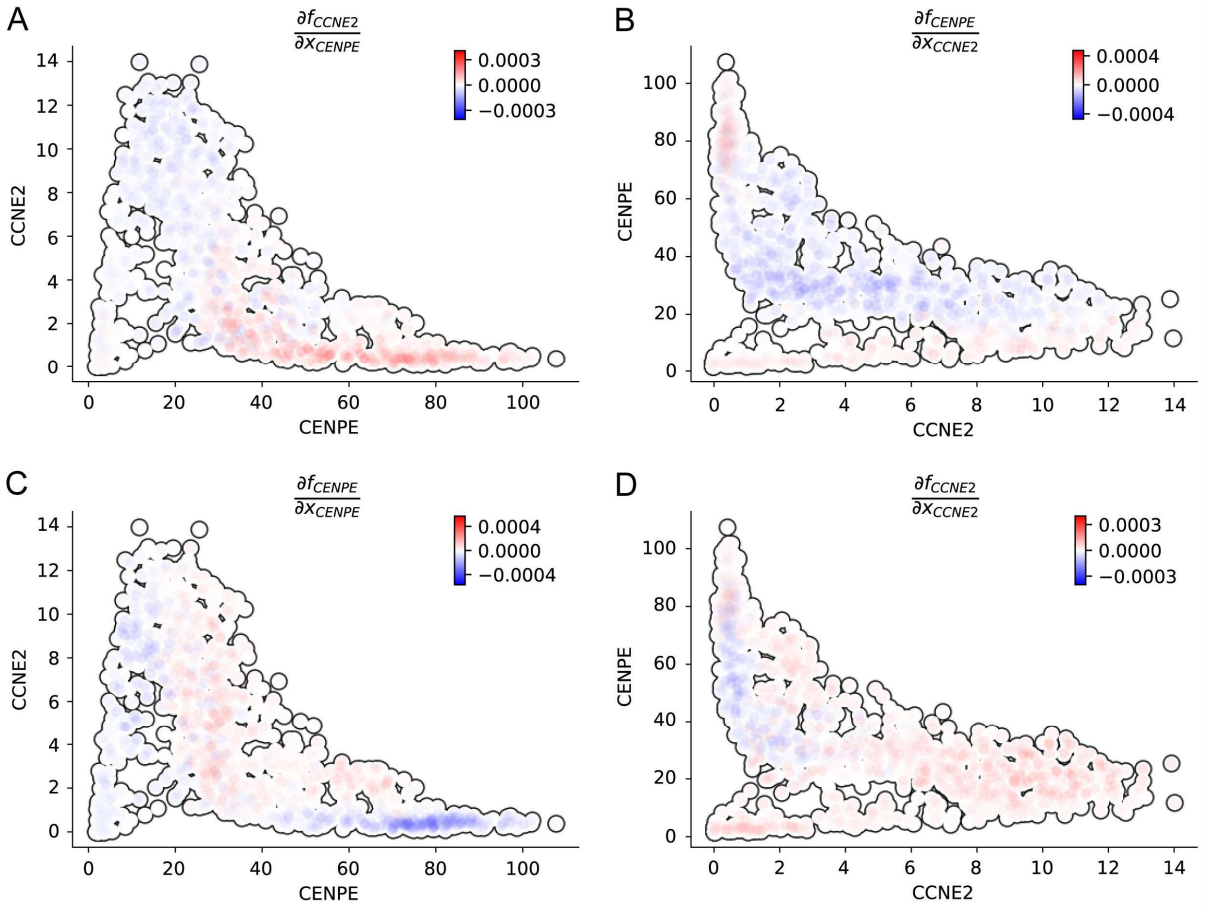

Figure S17. Jacobian analyses of the U2OS cell cycle regulatory interactions in the CCNE2 and CENPE expression space, related to Figure 6

(A) Activation from CENPE to CCNE2 in the CENPE and CCNE2 expression space.

(B) Repression from CCNE2 to CENPE in the CCNE2 and CENPE expression space.

(C) Self-activation of CENPE in the CENPE and CCNE2 expression space.

(D) Self-activation of CCNE2 in the CCNE2 and CENPE expression space.

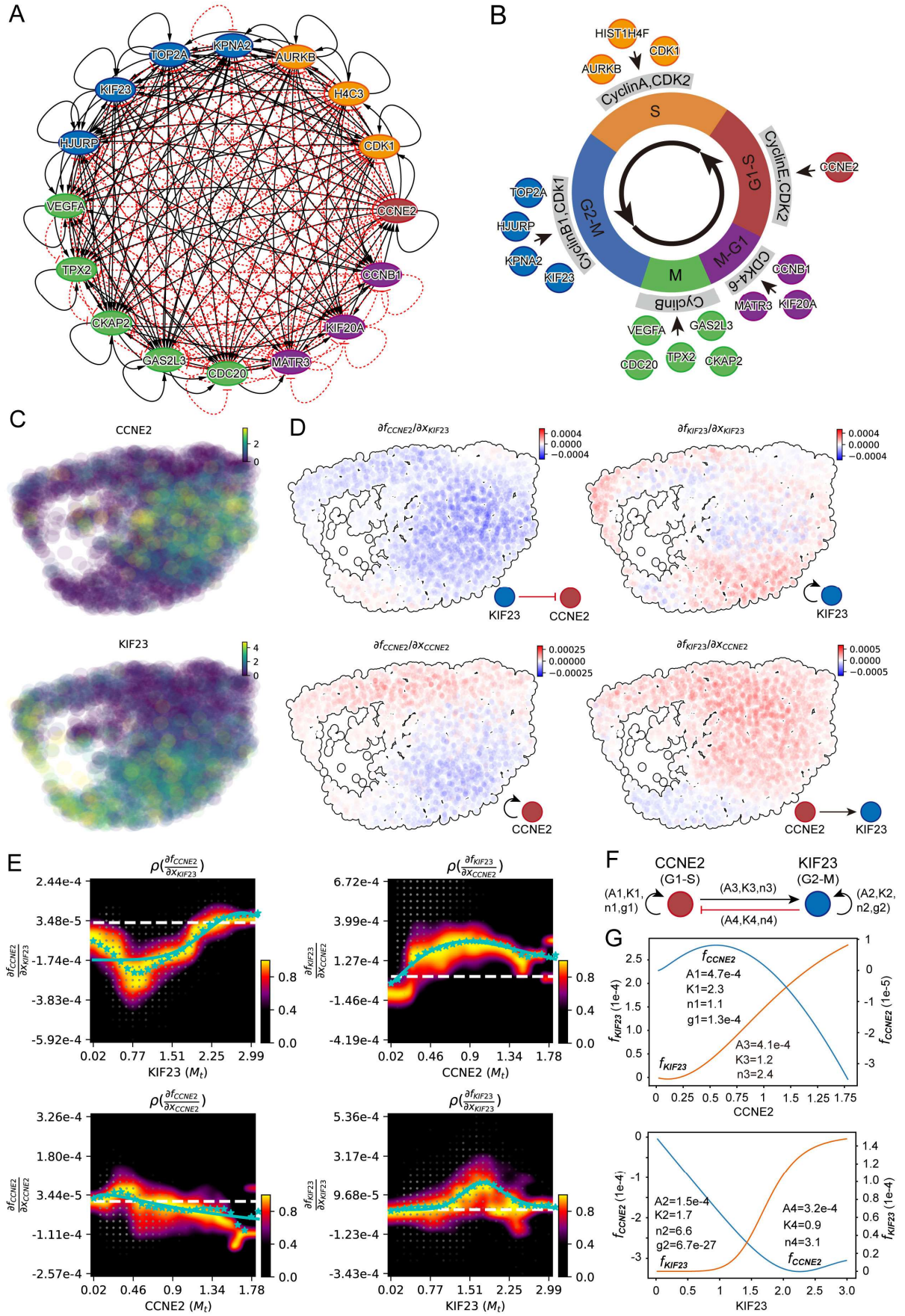

Figure S18. Inference of transcription factor regulatory networks for the RPE1 cell cycle

by single-cell transcriptomics data, related to Figure 6

(A) The interaction network of TFs of each cell cycle phase of RPE1 cells. All differentially expressed TFs in each cell cycle phase are used to construct network; nodes (TFs) with more than one edge are shown. Node colors represent different cycle phases (red: G1-S, yellow: S, blue: G2-M, green: M, purple: M-G1), the black arrows represent the activation and the red arrows represent the inhibition.

(B) The diagram for the cell cycle model with key TFs of each cell cycle phase.

(C) CCNE2 has high expression in G1-S phase and KIF23 has high expression in G2-M phase.

(D) Molecular mechanisms underlying the maintenance of the cell cycle. (i) Repression of CCNE2 by KIF23. (ii) Self-activation of KIF23. (iii) Self-activation of CCNE2. (iv) CCNE2 activates KIF23.

(E) Fitting the function of Jacobian versus gene expression with derivatives of a simplistic inhibitory or activation Hill equation. White dashed line corresponds to the zero Jacobian value. The blue stars at each x axis grid point correspond to the weighted mean of the Jacobian values for that point. The blue solid lines are the resultant fittings for the Jacobian.

(F) Schematic summarizing the interactions involving CCNE2 and KIF23.

(G) The velocity kinetic curves over gene expression changes of the corresponding fitted Hill equations of (E).

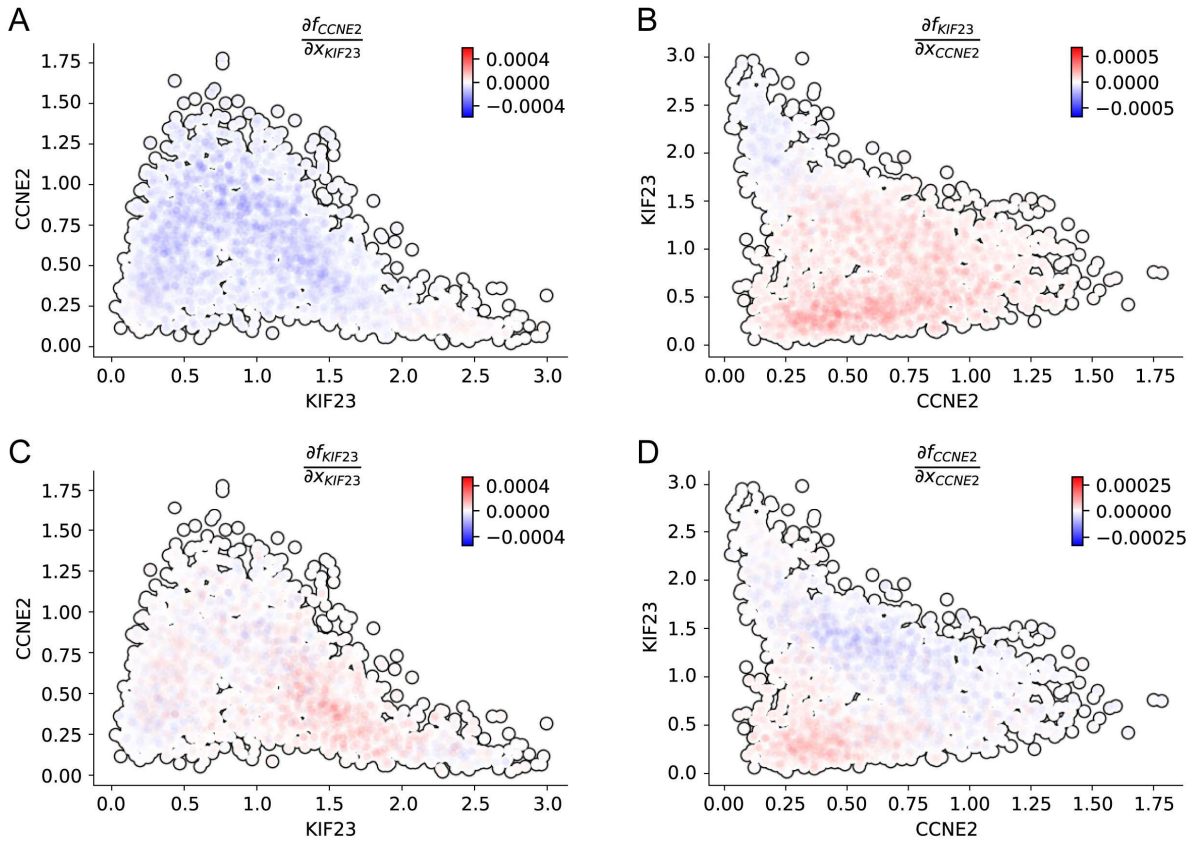

Figure S19: Jacobian analyses of the RPE1 cell cycle regulatory interactions in the CCNE2 and KIF23 expression space, related to Figure 6

(A) Repression from KIF23 to CCNE2 in the KIF23 and CCNE2 expression space.

(B) Activation from CCNE2 to KIF23 in the CCNE2 and KIF23 expression space.

(C) Self-activation of KIF23 in the KIF23 and CCNE2 expression space.

(D) Self-activation of CCNE2 in the CCNE2 and KIF23 expression space.

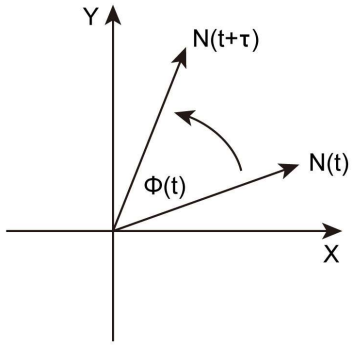

Figure S20: Sketch map for the definition of phase coherence related to Methods