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

*** For correspondence:** jin.wang.1@stonybrook.edu

Abstract

Recent developments in single-cell sequencing technology enable the acquisition of the whole transcriptome data. However, understanding the underlying mechanism and identifying the driving force of the transcriptional regulation of the cell function directly from these data remains challenging. To address this urgent need, we reconstruct a continuous vector field of cell cycle based on the discrete single-cell RNA velocity to quantify the single-cell global non-equilibrium dynamic landscape-flux. We reveal that large fluctuations disrupt the global landscape and genetic perturbations alter landscape-flux, thus identifying key genes in maintaining cell cycle dynamics and predicting associated effects on function. Additionally, we quantify the fundamental energy cost of the cell cycle initiation and reveal that sustaining the cell cycle requires curl flux and dissipation to maintain the oscillatory phase coherence. We enable the inference of the cell cycle gene regulatory networks directly from the single-cell transcriptomic data, including the feedback mechanisms and interaction intensity. This provides a golden opportunity to experimentally verify the landscape-flux theory and also obtain its associated quantifications. Our study also offers a unique framework for combining the landscape-flux theory and single-cell high-through sequencing experiments together for understanding the underlying mechanisms of the cell cycle and can be extended to other non-equilibrium 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 of a system, linking the dynamical and thermodynamic aspects of the system. Since flux originates from energy input and generates thermodynamic dissipation, it can act as a dynamic driving force (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-throughput sequencing and theoretical landscape-flux approaches to understand the origin and driving forces of the cell cycle. The purpose of our research is to utilize high-throughput data to experimentally verify the landscape and flux theory and provides its associated quantifications. 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 (13-16). 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 experimental data collected by single-cell high throughput sequencing.

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 human bone osteosarcoma epithelial cells (U2OS cells)(18) and single-cell 5-ethynyl-uridine-labeled RNA sequencing (scEU-seq) data from human retinal pigment epithelial-1 cells (RPE1 cells)(19). We first 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. 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 (24). 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 S2B). 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 in the vector field (Figure S2C). 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 S2D). 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 S2E). 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 S2F). 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 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 (25). 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 (26). 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. This potential landscape also reflects the stability of different states in the cell cycle and highlights key transition points. 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 (27).

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 S4), and quantified the non-equilibrium landscape-flux (Figure S3). These findings verify the presence of the curl flux as another component of the driving force for the nonequilibrium dynamics in

202 additional to the gradient of the cell cycle landscape by experimental single-cell high
203 throughput sequencing data, providing deeper insights into the mechanisms behind the
204 stability and behavior of intrinsic nonequilibrium cell cycle progression that cannot be
205 discerned by the landscape quantification alone. In fact, the cell cycle is one of the best
206 examples of demonstrating the nonequilibrium dynamics being dictated by both the landscape
207 and the flux. What's more, the creation of the cell cycle non-equilibrium landscape-flux through
208 single-cell transcriptomics data presents a new and powerful tool for the study of cell cycle
209 dynamics, allowing for insights into underlying mechanisms that drive cell cycle progression.
210

211 **Noise alters the global cell cycle dynamics and thermodynamics**

212 Variability in biomolecular noise or changes in molecular states can affect intracellular gene
213 regulations and networks, which are crucial for the uncertain, multi-destiny behavior of the
214 system (28-30). To investigate the effects of noise intensity on the underlying mechanism of
215 the cell cycle via nonequilibrium dynamics and thermodynamics, we performed simulations at
216 varying noise intensities and analyzed the resulting global changes in the cell cycle dynamics
217 and thermodynamics. Firstly, we examined changes to the global cell cycle dynamic
218 landscape and show that as the noise intensity increases, the global dynamics landscape of
219 the cell cycle becomes more complex and less predictable. Specifically, with zero noise, the
220 cell cycle progresses very smoothly and predictably (Figure 2A), while the ring valley of cell
221 cycle becomes wider as noise intensity increases (Figure 2B and 2C), whereas the ring valley
222 is broken when the noise is high, showing the discrete attraction basins, at which time the cell
223 cycle is broken and the cells arrest in the different phases (Figure 2D). Moreover, we found
224 that the barrier height decreased in the G0 phase and oscillation center in the landscape as
225 noise intensity increased (Figure 2I). Barrier height of G0 phase is defined as the barrier from
226 the G0 basin to the saddle point (between G0 and G1-S), which quantifies how easily cells
227 can switch from quiescent G0 phase to the cyclic state. The barrier height of the oscillation
228 center is defined as the potential difference between the minimum potential of the landscape
229 and the local maximum inside the limit cycle, which characterizes the global stability of the
230 oscillation system (7). We also investigated and found that the noise increases the cell cycle
231 period, indicative of the rate at which the cell cycle oscillates and how fast it grows (Figures
232 2J and S6). Oscillation amplitude of UMAP1 increases with the noise intensity, resulting in a
233 wider range of the oscillation limit cycle (Figures 2J, S5 and S6). Increased variability or noise
234 effectively makes potential landscape basins shallower, thereby broadening the distribution of
235 biological variables (31). By quantifying the trajectory autocorrelation function and the
236 corresponding power spectrum density of the oscillating trajectory for different diffusion
237 coefficient D, we can see that as the noise increases (the stability of the oscillation system
238 decreases), the peak height of the power spectrum density decreases and the distribution of
239 the power spectrum becomes more spread out (Figure S5). This indicates that the peak height
240 and the width of the power spectrum density can be a measure of the stability of the oscillation
241 system. To be clear, when D is in the range of 0.01 to 0.1, the landscape is more realistic, with

242 obvious cell cycle checkpoints, and it is consistent with the true level of experimental
243 measurements (32, 33).
244 Furthermore, we quantified the curl flux as a major component of the driving force of the cell
245 cycle and the entropy production rate (EPR) as the measure of the thermodynamic cost
246 association to the nonequilibrium curl flux of the cell cycle dynamic landscape to investigate
247 how noise intensity affected the underlying mechanism of the cell cycle via nonequilibrium
248 dynamics and thermodynamics. As shown in Figure 2E-H, increasing noise intensity resulted
249 in a more widely distributed flux, with cells flowing between different phases less predictably.
250 We observed that in the small noise range, the flux decreases as the noise increases, while
251 the cell cycle is still oscillating (Figure 2K). Larger flux derived from the energy input or nutrient
252 supply implies a larger cell cycle driving force and subsequently results in a shorter time to
253 complete the cell cycle oscillation (period is smaller). Indeed, the speed of the cell cycle is a
254 hallmark of cancer. However, when the noise is very large, although the flux becomes larger
255 with the increase of noise, the cell cycle has been destroyed. At this point, the flux drive on
256 the cell cycle can no longer resist the destruction of noise. In addition to global dynamics
257 landscape alteration, noise intensity also affects the thermodynamics and energetics of the
258 cell cycle. We found that the EPR of the cell cycle decreases with increasing noise intensity
259 (Figure 2K). Furthermore, the energy per period of the cell cycle decreases with increasing
260 noise intensity (Figure 2L), indicating a change in the energy requirements of the system.
261 These results imply that the cell cycle becomes less efficient at higher noise levels, suggesting
262 that the system requires energy to maintain the coherent cell cycle oscillations.
263 Taken together, these results demonstrate that noise intensity alters the underlying
264 mechanism of the cell cycle via global nonequilibrium dynamics and thermodynamics,
265 suggesting that the feedback of intracellular gene interactions plays an important role in
266 maintaining the robustness and stability of the cell cycle in the face of environmental
267 fluctuations. These findings have crucial implications for our understanding of the mechanisms
268 that control cell cycle progression and may have important applications in the development of
269 novel therapeutic strategies for the treatment of diseases associated with cell cycle
270 dysregulation. These changes suggest that noise intensity may play a critical role in regulating
271 the onset and termination of cell cycle phases.
272

273 ***In silico* perturbation analysis cell cycle dynamics and thermodynamics under genetic**
274 **perturbation**

275 Genetic perturbation refers to the alteration of the function of a biological system by external
276 or internal factors, such as environmental stimuli, drug inhibition, and gene knockdown (34).
277 It is used to study how genes interact with each other and affect cellular processes and
278 phenotypes (35-37). To investigate the effects of the genetic perturbations on the underlying
279 mechanism of the cell cycle via nonequilibrium dynamics and thermodynamics, we performed
280 *in silico* perturbations of different genes in a controlled and systematic manner. Key genes
281 involved in cell cycle regulation, such as cyclin-dependent kinases (CDKs), cyclins, and tumor
282 suppressor genes, were perturbed in this study. Similar to Figure 1, we first estimated the RNA
283 velocity of U2OS cells under different genetic backgrounds by *in silico* perturbation (Figure
284 S7A), and then reconstructed the different vector fields based on the perturbed RNA velocity

285 (Figure S7B). Finally, we quantified the global cell cycle dynamic landscapes under different
286 genetic perturbations by adding noise in the dynamic simulation (Figure 3A). We found that
287 perturbing distinct genes resulted in specific changes in the cell cycle landscape and
288 alterations in EPR and flux (Figure 3B and 3C).
289 Specifically, the negative perturbation of *CDC20* led to an increase in the barrier height of
290 the G2-M phase with higher chance of arrest in this phase, while positive perturbation of
291 *CDC20* led to a deeper basin of the G1 phase with higher chance of arrest in this phase (Figure
292 3A). As the experiment showed, *CDC20* plays a crucial role in cell division, being the target
293 protein in the spindle checkpoint during mitosis (38). We also found that the negative
294 perturbation of *CDC25C* caused disruption of cell cycle oscillation and led to a deeper well of
295 the G2-M phase leading to possible arrest, albeit less obviously than the effect caused by
296 perturbing *CDC20* (Figure 3A). Biologically, the downregulation of *CDC25C* induces cell cycle
297 arrest in G2/M phase in response to DNA damage via p53-mediated signal transduction, and
298 its abnormal expression is associated with cancer initiation, development, metastasis,
299 occurrence and poor prognosis (39). Similarly, genetic perturbations in RPE1 cells are
300 manifested that the negative perturbation of *KPNA2* led to a higher barrier height of the G2-M
301 phase, whereas the positive perturbation of *CDC20* led to a higher barrier height of the M
302 phase (Figure S8A). This finding aligns with experimental evidence: *KPNA2* knockdown
303 downregulated *CCNB2* and *CDK1*, inhibited cell proliferation, and induced cell cycle arrest in
304 the G2-M phase (40). *CCNE2* is a regulatory subunit of cyclin-dependent kinase 2 (Cdk2) and
305 is thought to control the transition of quiescent cells into the cell cycle (41). Our genetic
306 perturbations also confirm the cell cycle initiation function of *CCNE2*, as reflected in the
307 deepening of the G0 well under negative perturbation of *CCNE2*, and disappearance of the
308 G0 basin with the appearance of other periodic phases following positive perturbation of
309 *CCNE2* (Figure S8A).
310 Figure 3B shows the EPR of cell cycle nonequilibrium thermodynamics with different genetic
311 perturbations. We observed that different genetic perturbations can cause significant changes
312 in EPR, with some resulting in increased energy consumption and others in decreased
313 consumption. However, the EPR is lower under genetic perturbations than in WT, strong
314 perturbations disrupt cell oscillatory dynamics (Figures 3B and S8B). Thus, maintaining
315 continuous cell oscillation requires greater energetic costs than arresting in different periodic
316 phases. We also quantified the average flux of cell cycle nonequilibrium dynamics with various
317 genetic perturbations and found that WT cells exhibit the largest curl flux, where the flux as a
318 driving force of the cell cycle is greater than the negative gradient force, and the flux drives
319 the cell cycle to continue moving forward to maintain the coherent oscillation (Figures 3C and
320 S8C). In contrast, specific genetic perturbations result in a flux less than the negative gradient
321 force, insufficient to drive the cell cycle forward, and causing cells to arrest at certain cell cycle
322 phases.
323 Overall, the results demonstrate the efficacy of using the reconstructed vector field based
324 on RNA velocity derived from single-cell omics data of the cell cycle to predict how genetic
325 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 (42). The quantified Waddington landscape provides a way of describing and revealing biological pathways of developmental processes (43). 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 (Figures 4A and S9D). 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 (Figures 4B and S9E). 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) (Figures 4F and S9A). 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

369 logarithm of EPR (Figure 4H). The longer the cell takes to escape, the more energy it
370 consumes. To overcome this barrier, cells must expend a certain amount of energy, known as
371 cell cycle initiation energy (44). In summary, these results provide a quantitative assessment
372 of the initial energy required for the progression of the cell cycle (e.g., nutrient supply capacity
373 and metabolic activity (45)) and provide graphical, physical, and quantitative explanations for
374 the checkpoint mechanisms of the cell cycle.
375

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

378 Biological systems need to function accurately in the presence of strong noise, and various
379 regulatory mechanisms have evolved to suppress the fluctuation and environmental effects in
380 processing vital life processes such as cell cycle and development (46). In a noisy
381 environment, cell cycle oscillations can be highly inaccurate due to phase fluctuation, the
382 phase of oscillation follows diffusive dynamics (47, 48). A previous study found that phase
383 coherence is a property of waves that describes how well-aligned the phases of the waves
384 are. When waves are perfectly coherent, their peaks and troughs align precisely, creating a
385 stable and predictable pattern (7, 49). We quantified the relationship between noise intensity
386 and cell cycle oscillation properties (phase diffusion, phase coherence, and coherent time)
387 and found that noise fluctuations increase phase diffusion, decrease phase coherence, and
388 make the coherence time (cell cycle maintenance time) to become lower (Figure S10). In the
389 system of cell cycle regulation, phase diffusion refers to the random fluctuation of the phase
390 of oscillate waves. After certain times, the oscillations can lose the memory or track of the
391 previous dynamics. This makes it harder to maintain the periodicity. The phase coherence is
392 important for maintaining accurate and reliable oscillations in noisy environments. Coherent
393 time refers to the duration over which a system exhibits coherence, or the ability to maintain a
394 stable and predictable pattern (50). In other words, it is the length of time during which
395 oscillating waves in the cell cycle system remain perfectly aligned. There are significant
396 correlations between these factors: increasing phase diffusion leads to a corresponding
397 decrease in phase coherence, and phase diffusion can reduce the coherent time of the cell
398 cycle system, especially in large fluctuation systems (Figure 5A and 5B). Decreased phase
399 coherence leads to decreased coherent time because less well-aligned waves become less
400 stable and predictable over time (Figure 5C).
401 To identify the non-equilibrium mechanism of the cell maintenance cycle progression, we
402 further investigated the effects of dynamic flux and thermodynamic dissipation on the
403 characteristic features of cell cycle oscillations. Our results indicate that the logarithm of phase
404 diffusion is negatively correlated with the logarithm of EPR, while phase coherence and
405 coherent time are positively correlated with the logarithm of EPR (Figure 5D-5F). This
406 suggests that thermodynamic dissipation plays an essential role in regulating the overall
407 dynamics of the cell cycle oscillation. As cells enter the cycle, metabolic processes generate
408 ATP and other energy molecules to provide thermodynamic energy dissipation to maintain the
409 cell cycle oscillations. This added energy can help counteracting the effects of phase diffusion,
410 maintaining coherence, and resulting in more reliable and effective signaling between different
411 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, which align with the principles of thermodynamic uncertainty relationship (TUR). TUR establishes a trade-off between the dissipation of energy and the precision of observed quantities. In the context of cell cycle oscillating systems, maintaining coherence in the oscillation phases may require the dissipation of energy to counteract the natural tendency of the system to relax towards equilibrium. The coherent oscillation necessitates energy input to sustain the oscillatory behavior and counteract the dissipative processes that tend to disrupt coherence. the TUR can guide the understanding of how energy allocation influences the coherence and precision of oscillation phases. 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 (51-57). The use of gene 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 genes of U2OS cell and RPE1 cell, separately (Tables 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 genes regulates and maintains cells into different cycle stages via corresponding niche components (Figures 6A and S12A). The nodes in the network represent genes colored according to their associated cell cycle phase. We also identified key genes for each cell cycle phase and constructed a wiring diagram for the cell cycle network model with these key genes (Figures 6B and S12B). The GRN identifies key genes 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 (41, 58). While the S phase is regulated by *MCM6*, *AURKB*, and *CDK1*, which are related to DNA replication

455 and phosphorylate several key proteins to control the transition from G1 to S phase of the cell
456 cycle (59-61). Other important genes identified include *CDC25C*, *CENPE*, and *KIF23*, which
457 regulate the G2/M checkpoint by affecting mitotic cell cycle (62-64). Additionally, the GRN
458 reveals that the M phase is regulated by *CDC20* (38), *PIF1* (65), and *TPX2* (66), *PDE4B* (67),
459 and *KIF20A* (68), highlighting the crucial role of these factors in controlling the M-G1 phase.
460 Although there are many methods for inferring gene regulatory networks, most of them
461 cannot obtain the strength of regulation, and only a few tools can do it, including Dynamo (23),
462 spliceJAC (53), etc. In our study, we used Dynamo, based on the differential geometry
463 analysis of the reconstructed cell cycle vector field from the single-cell transcriptomics data,
464 we reveal the quantitative information about gene regulation and further uncover the molecular
465 mechanism underlying the maintenance of the cell cycle and the interactions between the key
466 genes. Taking U2OS cells as an example, we demonstrate via our inferred regulatory network
467 that *CCNE2* and *CENPE* regulate the G1-S and G2-M phase, respectively, showing peak
468 expression in their respective regulatory cycle phases (Figure 6C). Jacobian analysis of vector
469 fields can help us study cellular state-dependent gene interactions, with positive and negative
470 values of the Jacobian matrix reflecting activation and inhibition regulation, respectively (23).
471 Across all cells, Jacobian analyses demonstrate the activation of *CCNE2* by *CENPE*,
472 repression of *CENPE* by *CCNE2*, self-activation of *CENPE*, and self-activation of *CCNE2*
473 (Figure 6D). These results collectively suggest that the interlinked positive and negative
474 feedback loops between *CCNE2*, which regulates the G1-S checkpoint, and *CENPE*, which
475 regulates the G2-M checkpoint, underlies the dynamics of cell cycle oscillations, revealing the
476 critical role of two transcription factors in maintaining cell cycle progression (Figure 6F).
477 Although only one negative feedback loop is required for biological oscillation, we found that
478 *CCNE2* and *CENPE* have additional self-activating positive feedback loops that promote
479 modularity of cell cycle frequency while essentially unchanged amplitude (Figures 2J and S6).
480 Previous computational studies have shown that it is often difficult to adjust the frequency of
481 a negative feedback oscillator without affecting the amplitude, whereas with positive plus
482 negative feedback, a widely adjustable frequency and a near-constant amplitude can be
483 achieved (69). This also reveals that keeping the cell cycle more robust and easier to evolve
484 requires a lot of positive and negative feedback coupling. To further quantitatively understand
485 the feedback regulation mechanism between regulators, we first plotted distributions of the
486 four Jacobian elements versus expression of each gene (Figure S11A-S11D). Then, we
487 estimated kinetic parameters with derivatives of a simplistic inhibitory or activation hill equation
488 by fitting the Jacobian vs. expression curve in the response heatmap (Figure 6E). Finally, we
489 theoretically and numerically simulate the velocity dynamics curve of the hill equation as a
490 function of gene expression (Figure 6G). The GRN inferred above not only indicates the
491 presence or absence of the interaction between genes but does provide quantitative
492 information on positive and negative regulation relationships.
493 Similarly, based on the REP1 cell cycle regulation network inferred from scEU-seq data, we
494 further revealed the feedback mechanism between *CCNE1* (a G1-S phase regulator) and
495 *KIF23* (a G2-M phase regulator), and quantified the intensity of feedback between each other
496 (Figures S11E-S11H and S12). By inferring the feedback mechanism between different cell
497 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 (70).

Discussion

Spectacular advances in experimentation during the last decade, for example, single-cell dynamics in microscopy and in high throughput data acquisition, have yielded an unprecedented wealth of data on cell dynamics, genetic regulation, and organismal development. These experimental data have motivated the development and refinement of concepts and tools to dissect the physical mechanisms underlying biological processes. An appropriate analytical framework is needed to use these data to study the underlying mechanisms through the non-equilibrium dynamics and thermodynamic behavior of cells. Our study utilizes high-throughput sequencing experimental data to verify the landscape and flux theory and provides its associated quantifications. 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. 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, 43, 71, 72). 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 (73). 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 (74). 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 (74-76). 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 (44), 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 (77, 78). The Thermodynamic Uncertainty Relationship is a result in non-equilibrium statistical mechanics that relates the dissipation of energy or entropy production to the precision of a thermodynamic observable. In systems far from equilibrium, there is a trade-off between the dissipation of energy (or entropy production) and the fluctuations in the observed quantities. The TUR establishes a lower bound on the fluctuations of a thermodynamic observable based on the rate of entropy production or energy dissipation in the system. The cell cycle is naturally a

584 living system, G0 phase is a non-dividing state of the cell cycle, in which the cell is
585 metabolically active but not actively dividing. To switch from G0 phase to the G1/S phase, the
586 cell requires a certain amount of energy to initiate the complex set of biochemical reactions
587 that drive cell division. Once the cell has entered the G1/S phase, it must maintain a complex
588 system of oscillations and feedback loops in order to progress through the rest of the cell
589 cycle. This requires additional energy input to maintain the oscillations of key cell cycle
590 regulators such as cyclins and cyclin-dependent kinases (CDKs). Currently, certain
591 experiments can quantify the energy flow by measuring the rate of oxygen consumption or
592 heat production or the fluxes of metabolites including the production and comsumption of ATP
593 (79).
594 In systems biology, the creation of core network regulation to analyze biological processes
595 is an important and popular topic. In general, there are two ways to build a biological regulatory
596 network: one is to create a small network based on the existing knowledge and database, and
597 use the simulator to improve the network, but it is inefficient and cannot be used to build a new
598 network model (80, 81). The second is to use bioinformatics technology, especially high-
599 throughput sequencing data, to infer genes and gene correlations, but to ignore the actual
600 biological modulation relationship (82). To balance the advantages and disadvantages of
601 these two approaches, we infer the transcriptional regulatory networks based on differentially
602 expressed genes of each cycle phase from single-cell transcriptomics data. The regulatory
603 direction between genes (activation/inhibition) can be inferred by Jacobian analyses of vector
604 fields and the hill function of the core network can be simulated by fitting Jacobian in the gene
605 expression space and response heatmap. We reveal that cell cycle regulation involve multiple
606 positive and negative feedback loops between these regulators, which help to adjust the
607 frequency of the oscillation without significantly affecting the amplitude, enabling widely
608 adjustable frequencies and near-constant amplitudes and with maintaining balance between
609 the activation and inhibition of the gene expressions, permitting for proper progression through
610 different cell cycle phases. In this regard, further research is required to unravel such complex
611 regulatory networks comprehensively. Our results identify that some dynamic biological
612 processes can be accurately mapped by constructing regulatory network models based on
613 the high-throughput sequencing data.
614 In summary, we have built a general framework for the quantification of landscape-flux and
615 analysis of the underlying mechanism of the cell cycle via non-equilibrium dynamics and
616 thermodynamics based on single-cell transcriptomic data that can be applied to numerous
617 biological systems. The mathematical model behind RNA velocity quantity describes only the
618 biological transformation from DNA to RNA, which does not fully capture the intricacies of
619 post-translational regulation from RNA to protein and protein phosphorylation. Corresponding
620 to landscape-flux, the flux driving force may become smaller when protein phosphorylation is
621 lacking. It is worthwhile noticing although some details are missing, the main topology of the
622 Mexican landscape and corresponding flux are similar to the protein level studies (9). This
623 implies that the global nature of the cell cycle network may have already been captured at the
624 transcription level. Of course, more information will need to be complemented at the protein
625 level to study the detailed physical mechanisms. For further research in the future, we need
626 to combine proteomics to evaluate protein velocity and even combine epigenomics with

627 chromatin velocity to understand the regulation mechanism of the cell cycle more
628 comprehensively. More broadly, when coupled with remarkable advances in single-cell
629 approaches, including multi-velocity (83), construction of vector fields (35), as well as spatial
630 multi-omics (84), we will enable a deeper, quantitative understanding of the spatiotemporally
631 underlying mechanism of the cell cycle and other biological systems via non-equilibrium
632 dynamics and thermodynamics to help prevent various diseases, including cancer and
633 neurodegeneration, and open up new possibilities for precision medicine.
634

635 Materials and methods

636 Single-cell sequencing data processing and analysis

637 (1) For the scRNA-seq data for ~1k U2OS-FUCCI data, we followed the snakemake pipeline
638 at https://github.com/CellProfiling/FucciSingleCellSeqPipeline to perform scRNA-Seq data
639 preparation and general analysis including filter, dimensionality reduction, clustering analysis
640 and RNA velocity analysis (18).
641 (2) For the scEU-seq data for ~3k RPE1-FUCCI cells, we used
642 *dyn.sample_data.scEU_seq_rpe1* to acquire the processed data by using Dynamo (23), which
643 includes cell cycle clustering and RNA velocity.
644 (3) We reconstructed the vector field based RNA velocity with *dyn.vf.VectorField* by using
645 Dynamo (23). Then, we calculated the divergence, curl, acceleration and curvature to perform
646 the differential geometry analysis. We also calculated the jacobian to perform genetic
647 perturbation and inference gene regulatory interaction.
648 (4) We learned an analytical function of vector field from sparse single cell samples on the
649 entire space robustly by *vector_field_function.* Then, we calculated stochastic dynamics by
650 solving the Langevin equation based analytical function and quantified the non-equilibrium
651 thermodynamics and dynamics of the cell cycle.

652 RNA velocity

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

661 Reconstructing vector field from RNA velocity

662 Dynamo (23) was used to reconstruct function of the continuous and analytical velocity vector
663 field from sparse, noisy single-cell velocity measurements with sparse approximation of vector
664 field consensus (sparseVFC)(87). SparseVFC uses a vector-valued kernel method built on
665 reproducing kernel Hilbert space (RKHS) to learn the vector field, which is expressed
666 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:

$$\frac{d\mathrm{UMAP}_i(t)}{dt} = \mathbf{F}(\mathrm{UMAP}_i) + \zeta, \tag{1}$$

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 diffusion coefficient matrix. The noise term is associated with the intensity of cellular fluctuations either from environmental external fluctuations or intrinsic fluctuations. Under large n expansions, the process follows Brownian dynamics. 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, multi-stable 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 probability P_{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):

$$\mathbf{F} = \frac{\mathbf{D}}{P_{ss}} \cdot \frac{\partial}{\partial \mathbf{x}} P_{ss} + \frac{\mathbf{J_{ss}}(\mathbf{x})}{P_{ss}} = -\mathbf{D}\frac{\partial}{\partial \mathbf{x}}\mathrm{U} + \frac{\mathbf{J_{ss}}(\mathbf{x})}{P_{ss}}. \tag{2}$$

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 (Entropy 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 (88),

$$S = -k_B \int P(\mathbf{x}, t) \ln P(\mathbf{x}, t) dx \,. \tag{3}$$

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

$$T\dot{S} = k_B * T \int (\ln P + 1) \nabla \cdot \mathbf{J}\, dx = -\int (k_B T \nabla \ln P - \mathbf{F}) \cdot \mathbf{J}\, dx - \int \mathbf{F} \cdot \mathbf{J}\, dx \,, \tag{4}$$

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 $\mathbf{J} = \mathbf{F}P - \mathbf{D} \cdot \nabla P$, 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.

In silico perturbation

For genes in cells, genetic perturbations affect their velocity and thus the changes in the vector field. For small perturbations ($\Delta\mathbf{x}$) of the gene expression in cells, specific to the influence of perturbation dx_1, dx_2, ..., dx_n of each gene on velocity or vector field, it can be calculated with the exact differential:

$$\mathrm{df_i} = \frac{\partial \mathrm{f_i}}{\partial \mathrm{x_1}} \mathrm{dx_1} + \frac{\partial \mathrm{f_i}}{\partial \mathrm{x_1}} \mathrm{dx_1} + \cdots + \frac{\partial \mathrm{f_i}}{\partial \mathrm{x_1}} \mathrm{dx_1}, \tag{5}$$

In vectorized form:

$$\Delta \mathbf{f} = \mathbf{J} \Delta \mathbf{x} \,, \tag{6}$$

J is the Jacobian matrix of vector field and $\Delta\mathbf{f}$ is the velocity and vector field response *in silico* genetic perturbation. In practice, a proportionality constant c is often added to the perturbation $\Delta\mathbf{x}$ to amplify the response $\Delta\mathbf{f}$.

We first used the *dyn.vf.perturbation* function to perform *in silico* perturbation and visualize the RNA velocity from the perturbation effect vectors. Then, reconstruct the perturbed vector field from the perturbed RNA velocity with *dyn.vf.vectorfield* function. Finally, we perform random dynamics simulation on the vector field of the genetic perturbation to obtain the perturbed landscape, and then quantify the relevant flux and EPR.

LAPs (Least Action Paths)

Considering a stochastic dynamical system described by Langevin equation:

$$\frac{d\mathrm{UMAP}_i(t)}{dt} = \mathbf{F}(\mathrm{UMAP}_i) + \zeta. \tag{7}$$

Following the approaches based on the Freidlin-Wentzell theory (89), 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:

$$S_T(\mathbf{x}_{jk}) = \frac{1}{2}\int_0^T \frac{\left|\dot{\mathbf{x}}_{jk} - \mathbf{F}(\mathbf{x}_{jk})\right|^2}{D}\,dt. \tag{8}$$

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 (89):

$$R(A \to B) \approx C\,\exp(-S_T), \tag{9}$$

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:

$$\mathrm{MFPT} = \frac{1}{R(A \to B)}. \tag{10}$$

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 S5. 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* (47):

$$D_\emptyset = D\frac{(2\pi)^2}{T}. \tag{11}$$

The autocorrelation function (47):

$$C(t) = \frac{\langle (x(t+s) - \langle x\rangle)(x(s) - \langle x\rangle)\rangle}{\langle x^2\rangle - \langle x\rangle^2} = \exp\left(-\frac{t}{\tau_c}\right) \times \cos\left(\frac{2\pi t}{T}\right), \tag{12}$$

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,

$$Z(t) = \begin{cases} x(t), & |t| < \frac{\tau}{2} \\ 0, & \text{otherwise} \end{cases}, \tag{13}$$

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,

$$E = \int_{-\infty}^{\infty} |Z(t)|^2 dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} |Z(\omega)|^2 d\omega, \tag{14}$$

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

$$\int_{-\infty}^{\infty} |Z(t)|^2 dt = \int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} |x(t)|^2 dt. \tag{15}$$

Therefore, we have,

$$\frac{1}{\tau}\int_{-\frac{\tau}{2}}^{\frac{\tau}{2}} |x(t)|^2 dt = \frac{1}{2\pi}\cdot\frac{1}{\tau}\int_{-\infty}^{\infty} |Z(\omega)|^2 d\omega, \tag{16}$$

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

$$P = \frac{1}{2\pi}\int_{-\infty}^{\infty} \lim_{\tau\to\infty}\left(\frac{|Z(\omega)|^2}{\tau}\right) d\omega\,, \tag{17}$$

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.,

$$S(\omega) = \lim_{T\to\infty} \frac{|Z(\omega)|^2}{\tau}, \tag{18}$$

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,

$$S(\omega) = \lim_{T\to\infty} \frac{|x(\omega)|^2}{\tau}. \tag{19}$$

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

$$R(\tau) \overset{FT}{\leftrightarrow} S(\omega). \tag{20}$$

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 (49). 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 S10I. 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

$$\xi = \frac{2\sum_i \theta\big(\phi(t)\big)\phi(t)}{\sum_i |\phi(t)|} - 1, \tag{21}$$

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 genes interaction network of the cell cycle, we first calculated and obtained the all differentially expressed genes in different cell cycle phase (Tables S1 and S2). After that the interaction between the differentially expressed genes 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 (86). The nodes and edges each represent the gene and interactions in a network.

Estimating kinetic parameters by fitting the Jacobian vs. expression curve

The Hill function is a mathematical model commonly used to describe various regulatory interactions in genetic and biochemical networks. It is used to represent the relationship between the concentration of a regulator (e.g., transcription factor) and the rate of transcriptional activation or repression. Its versatility lies in its ability to capture a wide range of phenomena and dynamics commonly observed in biological systems. The Hill function for quantifying the gene regulations works reasonably well when regulating binding/unbinding is relatively fast compared to protein synthesis and degradation. Derivation from the Hill function is expected when the regulatory binding/unbinding is relatively slow or comparable compared to protein synthesis and degradation (90, 91). 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:

$$x \rightarrow y: \frac{Ax^n}{(K + x^n)}, \tag{22}$$

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

$$x \dashv y: \frac{AK}{(K + x^n)}. \tag{23}$$

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

$$\begin{aligned} \frac{dx}{dt} &= f_x = \frac{A_1 x^{n1}}{(K_1 + x^{n1})} + \frac{A_2 K_2}{(K_2 + y^{n2})} - g_1 \cdot x, \\ \frac{dy}{dt} &= f_y = \frac{A_3 y^{n3}}{(K_3 + y^{n3})} + \frac{A_4 x^{n4}}{(K_4 + x^{n4})} - g_2 \cdot y. \end{aligned} \tag{24}$$

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

$$J = \begin{bmatrix} \frac{\partial f_x}{\partial x} & \frac{\partial f_x}{\partial y} \\ \frac{\partial f_y}{\partial x} & \frac{\partial f_y}{\partial y} \end{bmatrix} = \begin{bmatrix} \frac{A_1 K_1 n_1 x^{n_1 - 1}}{(K_1 + x^{n_1})^2} - g_1 & -\frac{A_2 K_2 n_2 y^{n_2 - 1}}{(K_2 + y^{n_2})^2} \\ -\frac{A_3 K_3 n_3 x^{n_3 - 1}}{(K_3 + x^{n_3})^2} & \frac{A_4 K_4 n_4 x^{n_4 - 1}}{(K_4 + y^{n_4})^2} - g_2 \end{bmatrix}. \tag{25}$$

In Figures 6E and S12E, 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).

Quantification and statistical analysis

The calculation of the amplitude and period of cell cycle oscillations was averaged multiple times using MATLAB. All correlation analyses and confidence intervals are calculated in Origin. The fitting evaluation of phase diffusion was performed by MATLAB.

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.

Supplemental information

Supplementary Table S1-S3 and Supplementary Figure S1–S12.

Data and code availability

The scRNA-seq raw data of U2OS-FUCCI cell cycle from the paper by Mahdessian et al (18), which are available at GEO with accession GSE14677318. 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 (19), which are available at GEO with accession number GSE12836519. The processed data are available from Github (https://github.com/Zhu-1998/cellcycle). The source codes used in this study is publicly available from Github (https://github.com/Zhu-1998/cellcycle). Requests for additional program codes and data generate and/or analyzed during the current study should be fulfilled on reasonable request by the Lead Contact.

References

1. R. J. Cho *et al.*, Transcriptional regulation and function during the human cell cycle. *Nat Genet* **27**, 48-54 (2001).
2. L. J. Liu, W. Michowski, A. Kolodziejczyk, P. Sicinski, The cell cycle in stem cell proliferation, pluripotency and differentiation. *Nat Cell Biol* **21**, 1060-1067 (2019).
3. A. E. Teschendorff, A. P. Feinberg, Statistical mechanics meets single-cell biology. *Nat Rev Genet* **22**, 459-476 (2021).
4. B. D. MacArthur, I. R. Lemischka, Statistical Mechanics of Pluripotency. *Cell* **154**, 484-489 (2013).
5. F. C. Tang, K. Q. Lao, M. A. Surani, Development and applications of single-cell transcriptome analysis. *Nat Methods* **8**, S6-S11 (2011).
6. X. N. Fang, K. Kruse, T. Lu, J. Wang, Nonequilibrium physics in biology. *Rev Mod Phys* **91**, 045004 (2019).

7. J. Wang, L. Xu, E. K. Wang, Potential landscape and flux framework of nonequilibrium networks: Robustness, dissipation, and coherence of biochemical oscillations. *P Natl Acad Sci USA* **105**, 12271-12276 (2008).
8. J. Wang, C. H. Li, E. K. Wang, Potential and flux landscapes quantify the stability and robustness of budding yeast cell cycle network. *P Natl Acad Sci USA* **107**, 8195-8200 (2010).
9. C. H. Li, J. Wang, Landscape and flux reveal a new global view and physical quantification of mammalian cell cycle. *P Natl Acad Sci USA* **111**, 14130-14135 (2014).
10. L. Zhu *et al.*, Network modeling-based identification of the switching targets between pyroptosis and secondary pyroptosis. *Chaos Solitons Fractals* **155**, 111724 (2022).
11. D. B. Burkhardt, B. P. San Juan, J. G. Lock, S. Krishnaswamy, C. L. Chaffer, Mapping Phenotypic Plasticity upon the Cancer Cell State Landscape Using Manifold Learning. *Cancer Discov* **12**, 1847-1859 (2022).
12. P. J. Zhou, S. X. Wang, T. J. Li, Q. Nie, Dissecting transition cells from single-cell transcriptome data through multiscale stochastic dynamics. *Nat Commun* **12**, 5609 (2021).
13. F. Izzo *et al.*, DNA methylation disruption reshapes the hematopoietic differentiation landscape. *Nat Genet* **52**, 378-387 (2020).
14. X. Qin *et al.*, A Single-cell Perturbation Landscape of Colonic Stem Cell Polarisation. *bioRxiv* 10.1101/2023.02.15.528008, 2023.2002.2015.528008 (2023).
15. S. Q. Jin, A. L. MacLean, T. Peng, Q. Nie, scEpath: energy landscape-based inference of transition probabilities and cellular trajectories from single-cell transcriptomic data. *Bioinformatics* **34**, 2077-2086 (2018).
16. J. F. Shi, A. E. Teschendorff, W. Y. Chen, L. N. Chen, T. J. Li, Quantifying Waddington's epigenetic landscape: a comparison of single-cell potency measures. *Brief Bioinform* **21**, 248-261 (2020).
17. J. Wang, K. Zhang, L. Xu, E. Wang, Quantifying the Waddington landscape and biological paths for development and differentiation. *P Natl Acad Sci USA* **108**, 8257-8262 (2011).
18. D. Mahdessian *et al.*, Spatiotemporal dissection of the cell cycle with single-cell proteogenomics. *Nature* **590**, 649-654 (2021).
19. N. Battich *et al.*, Sequencing metabolically labeled transcripts in single cells reveals mRNA turnover strategies. *Science* **367**, 1151-1156 (2020).
20. E. Becht *et al.*, Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* **37**, 38-44 (2019).
21. G. La Manno *et al.*, RNA velocity of single cells. *Nature* **560**, 494-498 (2018).
22. V. Bergen, M. Lange, S. Peidli, F. A. Wolf, F. J. Theis, Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* **38**, 1408–1414 (2020).
23. X. J. Qiu *et al.*, Mapping transcriptomic vector fields of single cells. *Cell* **185**, 690-711 (2022).
24. M. Kazumitsu, O. Yasuyuki, Modeling latent flows on single-cell data using the Hodge decomposition. *bioRxiv* 10.1101/592089, 592089 (2019).

25. Y. Zhao, W. Zhang, T. Li, EPR-Net: Constructing non-equilibrium potential landscape via a variational force projection formulation. 10.48550/arXiv.2301.01946, arXiv:2301.01946 (2023).
26. M. W. Hirsch, S. Smale, R. L. Devaney, M. W. Hirsch, *Differential equations, dynamical systems, and an introduction to chaos*, Pure and applied mathematics a series of monographs and textbooks (Academic Press, San Diego, CA, ed. 2nd, 2004), pp. xiv, 417 p.
27. V. Moiseeva *et al.*, Senescence atlas reveals an aged-like inflamed niche that blunts muscle regeneration. *Nature* **613**, 169-178 (2023).
28. G. Balazsi, A. van Oudenaarden, J. J. Collins, Cellular Decision Making and Biological Noise: From Microbes to Mammals. *Cell* **144**, 910-925 (2011).
29. N. Eling, M. D. Morgan, J. C. Marioni, Challenges in measuring and understanding biological noise. *Nat Rev Genet* **20**, 536-548 (2019).
30. A. Levchenko, I. Nemenman, Cellular noise and information transmission. *Curr Opin Biotech* **28**, 156-164 (2014).
31. E. Pujadas, A. P. Feinberg, Regulated Noise in the Epigenetic Landscape of Development and Disease. *Cell* **148**, 1123-1131 (2012).
32. M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, Stochastic gene expression in a single cell. *Science* **297**, 1183-1186 (2002).
33. Z. Hensel *et al.*, Stochastic expression dynamics of a transcription factor revealed by single-molecule noise analysis. *Nature Structural & Molecular Biology* **19**, 797-802 (2012).
34. O. S. Rukhlenko *et al.*, Control of cell state transitions. *Nature* **609**, 975-985 (2022).
35. K. Kamimoto *et al.*, Dissecting cell identity via network inference and in silico gene perturbation. *Nature* **614**, 742-751 (2023).
36. J. S. Fleck *et al.*, Inferring and perturbing cell fate regulomes in human brain organoids. *Nature* 10.1038/s41586-022-05279-8 (2022).
37. H. Hu *et al.*, Modeling and analyzing single-cell multimodal data with deep parametric inference. *Brief Bioinform* **24**, bbad005 (2023).
38. V. Piano *et al.*, CDC20 assists its catalytic incorporation in the mitotic checkpoint complex. *Science* **371**, 67-71 (2021).
39. R. X. Huang, P. K. Zhou, DNA damage response signaling pathways and targets for radiotherapy sensitization in cancer. *Signal Transduct Tar* **5**, 60 (2020).
40. L. Huang *et al.*, KPNA2 promotes cell proliferation and tumorigenicity in epithelial ovarian carcinoma through upregulation of c-Myc and downregulation of FOXO3a. *Cell Death Dis* **4**, e745 (2013).
41. R. Sonntag *et al.*, 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 (2018).
42. T. Graf, T. Enver, Forcing cells to change lineages. *Nature* **462**, 587-594 (2009).
43. H. D. Feng, K. Zhang, J. Wang, Non-equilibrium transition state rate theory. *Chem Sci* **5**, 3761-3769 (2014).

994 44. A. Papagiannakis, B. Niebel, E. C. Wit, M. Heinemann, Autonomous Metabolic
995 Oscillations Robustly Gate the Early and Late Cell Cycle. *Mol Cell* **65**, 285-295 (2017).
996 45. J. J. Zhu, C. B. Thompson, Metabolic regulation of cell growth and proliferation. *Nat*
997 *Rev Mol Cell Bio* **20**, 436-450 (2019).
998 46. J. E. Ferrell, T. Y. C. Tsai, Q. O. Yang, Modeling the Cell Cycle: Why Do Certain
999 Circuits Oscillate? *Cell* **144**, 874-885 (2011).
1000 47. Y. S. Cao, H. L. Wang, Q. Ouyang, Y. H. Tu, The free-energy cost of accurate
1001 biochemical oscillations. *Nat Phys* **11**, 772-778 (2015).
1002 48. C. Y. Fei, Y. S. Cao, Q. Ouyang, Y. H. Tu, Design principles for enhancing phase
1003 sensitivity and suppressing phase fluctuations simultaneously in biochemical
1004 oscillatory systems. *Nat Commun* **9**, 1434 (2018).
1005 49. M. Yoda, T. Ushikubo, W. Inoue, M. Sasai, Roles of noise in single and coupled
1006 multiple genetic oscillators. *J Chem Phys* **126**, 115101 (2007).
1007 50. D. A. Potoyan, P. G. Wolynes, On the dephasing of genetic oscillators. *P Natl Acad*
1008 *Sci USA* **111**, 2391-2396 (2014).
1009 51. P. C. Aubin-Frankowski, J. P. Vert, Gene regulation inference from single-cell RNA-
1010 seq data with linear differential equations and velocity inference. *Bioinformatics* **36**,
1011 4774-4780 (2020).
1012 52. S. Aibar *et al.*, SCENIC: single-cell regulatory network inference and clustering. *Nat*
1013 *Methods* **14**, 1083-1086 (2017).
1014 53. F. Bocci, P. J. Zhou, Q. Nie, spliceJAC: transition genes and state-specific gene
1015 regulation from single-cell transcriptome data. *Mol Syst Biol* **18**, e11176 (2022).
1016 54. T. E. Chan, M. P. H. Stumpf, A. C. Babtie, Gene Regulatory Network Inference from
1017 Single-Cell Data Using Multivariate Information Measures. *Cell Syst* **5**, 251-267 (2017).
1018 55. S. N. Chen, J. C. Mar, Evaluating methods of inferring gene regulatory networks
1019 highlights their lack of performance for single cell gene expression data. *Bmc*
1020 *Bioinformatics* **19**, 232 (2018).
1021 56. X. Dong *et al.*, Single-cell gene regulation network inference by large-scale data
1022 integration. *Nucleic Acids Res* **50**, e126 (2022).
1023 57. A. Gupta *et al.*, Inferring gene regulation from stochastic transcriptional variation
1024 across single cells at steady state. *P Natl Acad Sci USA* **119**, e2207392119 (2022).
1025 58. Z. Yuan *et al.*, Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6
1026 and Cdt1. *Nature Structural & Molecular Biology* **24**, 316-324 (2017).
1027 59. L. Wei, X. L. Zhao, A new MCM modification cycle regulates DNA replication initiation.
1028 *Nature Structural & Molecular Biology* **23**, 209-216 (2016).
1029 60. J. Bertran-Alamillo *et al.*, AURKB as a target in non-small cell lung cancer with acquired
1030 resistance to anti-EGFR therapy. *Nat Commun* **10**, 1812 (2019).
1031 61. S. Basu, J. Greenwood, A. W. Jones, P. Nurse, Core control principles of the
1032 eukaryotic cell cycle. *Nature* **607**, 381-386 (2022).
1033 62. R. Boutros, V. Lobjois, B. Ducommun, CDC25 phosphatases in cancer cells: key
1034 players? Good targets? *Nat Rev Cancer* **7**, 495-507 (2007).
1035 63. S. Hou *et al.*, XAB2 functions in mitotic cell cycle progression via transcriptional
1036 regulation of CENPE. *Cell Death Dis* **7**, e2409 (2016).

1037 64. R. S. Pedersen *et al.*, Profiling DNA damage response following mitotic perturbations.
1038 *Nat Commun* **7**, 13887 (2016).
1039 65. K. Paeschke *et al.*, Pif1 family helicases suppress genome instability at G-quadruplex
1040 motifs. *Nature* **497**, 458-462 (2013).
1041 66. P. Wadsworth, TPX2. *Curr Biol* **25**, R1156-R1158 (2015).
1042 67. D. H. Maurice *et al.*, Advances in targeting cyclic nucleotide phosphodiesterases. *Nat*
1043 *Rev Drug Discov* **13**, 290-314 (2014).
1044 68. A. Geng *et al.*, KIF20A/MKLP2 regulates the division modes of neural progenitor cells
1045 during cortical development. *Nat Commun* **9**, 2707 (2018).
1046 69. T. Y. C. Tsai *et al.*, Robust, tunable biological oscillations from interlinked positive and
1047 negative feedback loops. *Science* **321**, 126-129 (2008).
1048 70. M. Q. Xie, M. Fussenegger, Designing cell function: assembly of synthetic gene circuits
1049 for cell biology applications. *Nat Rev Mol Cell Bio* **19**, 507-525 (2018).
1050 71. L. Xu, D. Patterson, A. C. Staver, S. A. Levin, J. Wang, Unifying deterministic and
1051 stochastic ecological dynamics via a landscape-flux approach. *P Natl Acad Sci USA*
1052 **118**, e2103779118 (2021).
1053 72. L. Xu, D. Patterson, S. A. Levin, J. Wang, Non-equilibrium early-warning signals for
1054 critical transitions in ecological systems. *Proc Natl Acad Sci U S A* **120**, e2218663120
1055 (2023).
1056 73. N. Vasan, J. Baselga, D. M. Hyman, A view on drug resistance in cancer. *Nature* **575**,
1057 299-309 (2019).
1058 74. X. N. Fang *et al.*, Cell fate potentials and switching kinetics uncovered in a classic
1059 bistable genetic switch. *Nat Commun* **9**, 2787 (2018).
1060 75. Z. L. Jiang *et al.*, The emergence of the two cell fates and their associated switching
1061 for a negative auto-regulating gene. *Bmc Biol* **17**, 49 (2019).
1062 76. L. Sheng, L. Qiong, W. Erkang, W. Jin, Global quantitative understanding of
1063 nonequilibrium cell fate decision making in response to pheromone. *bioRxiv*
1064 10.1101/2022.07.04.498707, 2022.2007.2004.498707 (2023).
1065 77. X. N. Fang, J. Wang, Nonequilibrium Thermodynamics in Cell Biology: Extending
1066 Equilibrium Formalism to Cover Living Systems. *Annual Review of Biophysics* **49**, 227-
1067 246 (2020).
1068 78. L. Vonbertalanffy, The Theory of Open Systems in Physics and Biology. *Science* **111**,
1069 23-29 (1950).
1070 79. X. B. Yang *et al.*, Physical bioenergetics: Energy fluxes, budgets, and constraints in
1071 cells. *P Natl Acad Sci USA* **118**, e2026786118 (2021).
1072 80. X. Fei *et al.*, Topological design principle for the robustness of necroptosis biphasic,
1073 emergent, and coexistent (BEC) dynamics. *bioRxiv* 10.1101/2023.01.23.525173,
1074 2023.2001.2023.525173 (2023).
1075 81. N. Le Novere, Quantitative and logic modelling of molecular and gene networks. *Nat*
1076 *Rev Genet* **16**, 146-158 (2015).
1077 82. A. Pratapa, A. P. Jalihal, J. N. Law, A. Bharadwaj, T. M. Murali, Benchmarking
1078 algorithms for gene regulatory network inference from single-cell transcriptomic data.
1079 *Nat Methods* **17**, 147-154 (2020).

83. S. Li *et al.*, A relay velocity model infers cell-dependent RNA velocity. *Nat Biotechnol* 10.1038/s41587-023-01728-5 (2023).
84. K. Vandereyken, A. Sifrim, B. Thienpont, T. Voet, Methods and applications for single-cell and spatial multi-omics. *Nat Rev Genet* **24**, 494-515 (2023).
85. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology* **19**, 15 (2018).
86. P. Shannon *et al.*, Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).
87. J. Y. Ma, J. Zhao, J. W. Tian, X. Bai, Z. W. Tu, Regularized vector field learning with sparse approximation for mismatch removal. *Pattern Recogn* **46**, 3519-3532 (2013).
88. H. Qian, Mesoscopic nonequilibrium thermodynamics of single macromolecules and dynamic entropy-energy compensation. *Phys Rev E* **65**, 016102 (2002).
89. M. I. Freĭdlin, A. D. Wentzell, J. Szücs, *Random perturbations of dynamical systems*, Grundlehren der mathematischen wissenschaften (Springer, Heidelberg, ed. Third Edition., 2012), pp. xxviii, 458 pages.
90. J. E. M. Hornos *et al.*, Self-regulating gene: An exact solution. *Phys Rev E* **72**, 051907 (2005).
91. H. D. Feng, B. Han, J. Wang, Adiabatic and Non-Adiabatic Non-Equilibrium Stochastic Dynamics of Single Regulating Genes. *J Phys Chem B* **115**, 1254-1261 (2011).

Figures and legends

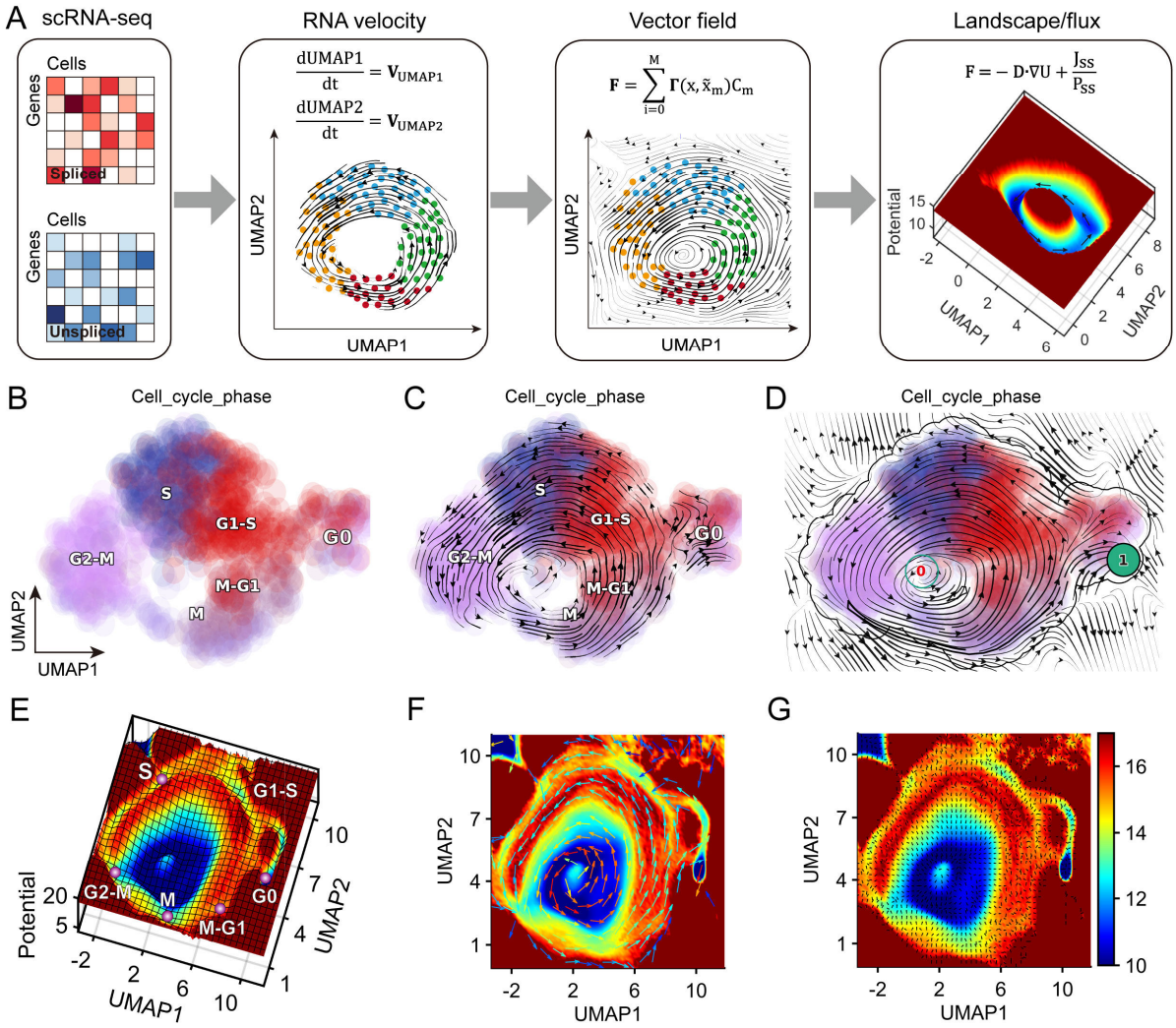


Figure 1. Workflow for analyzing cell cycle landscape-flux from single-cell transcriptomics data

(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 (colored arrow) of cell cycle dynamics landscape in UMAP of U2OS cells. The color of the arrow presents the magnitude of flux, red corresponds to larger, and blue to smaller.

(G) The gradient force (black arrow) of cell cycle dynamics landscape in UMAP of U2OS cells.

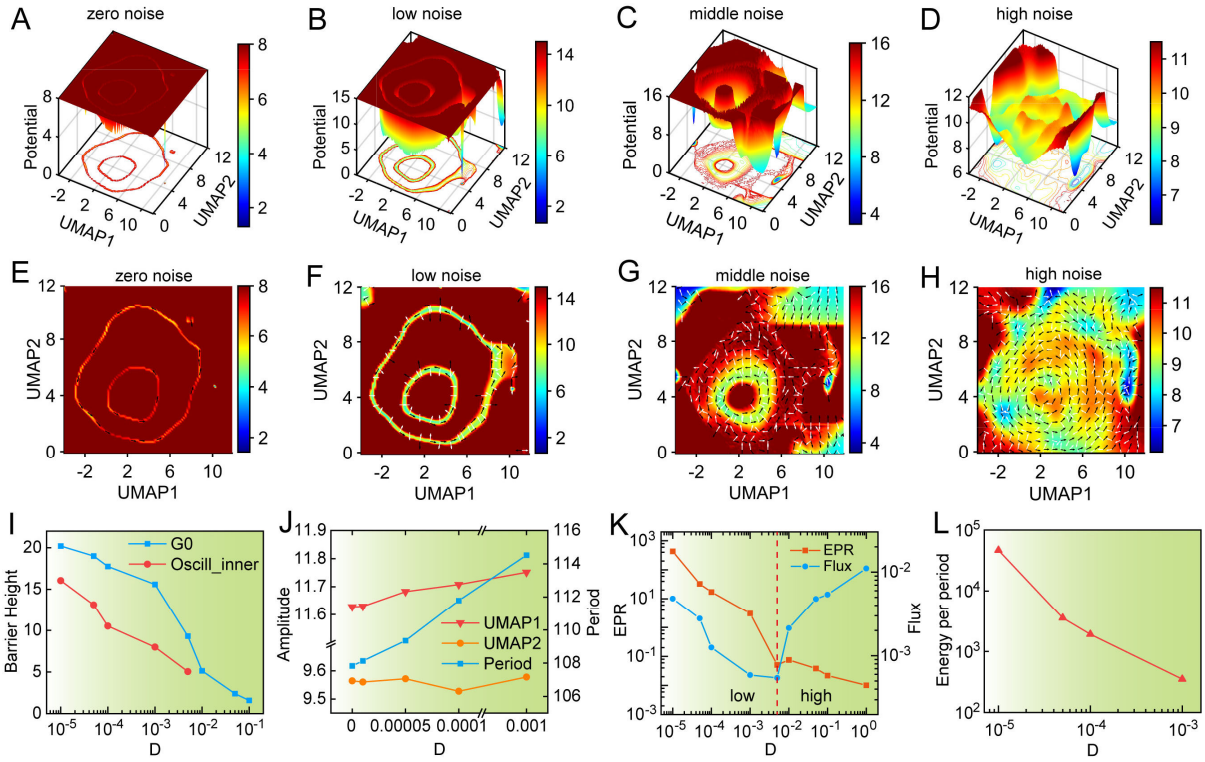


Figure 2. Cell cycle nonequilibrium landscape-flux affected by noise intensity

(A to 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 to 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.

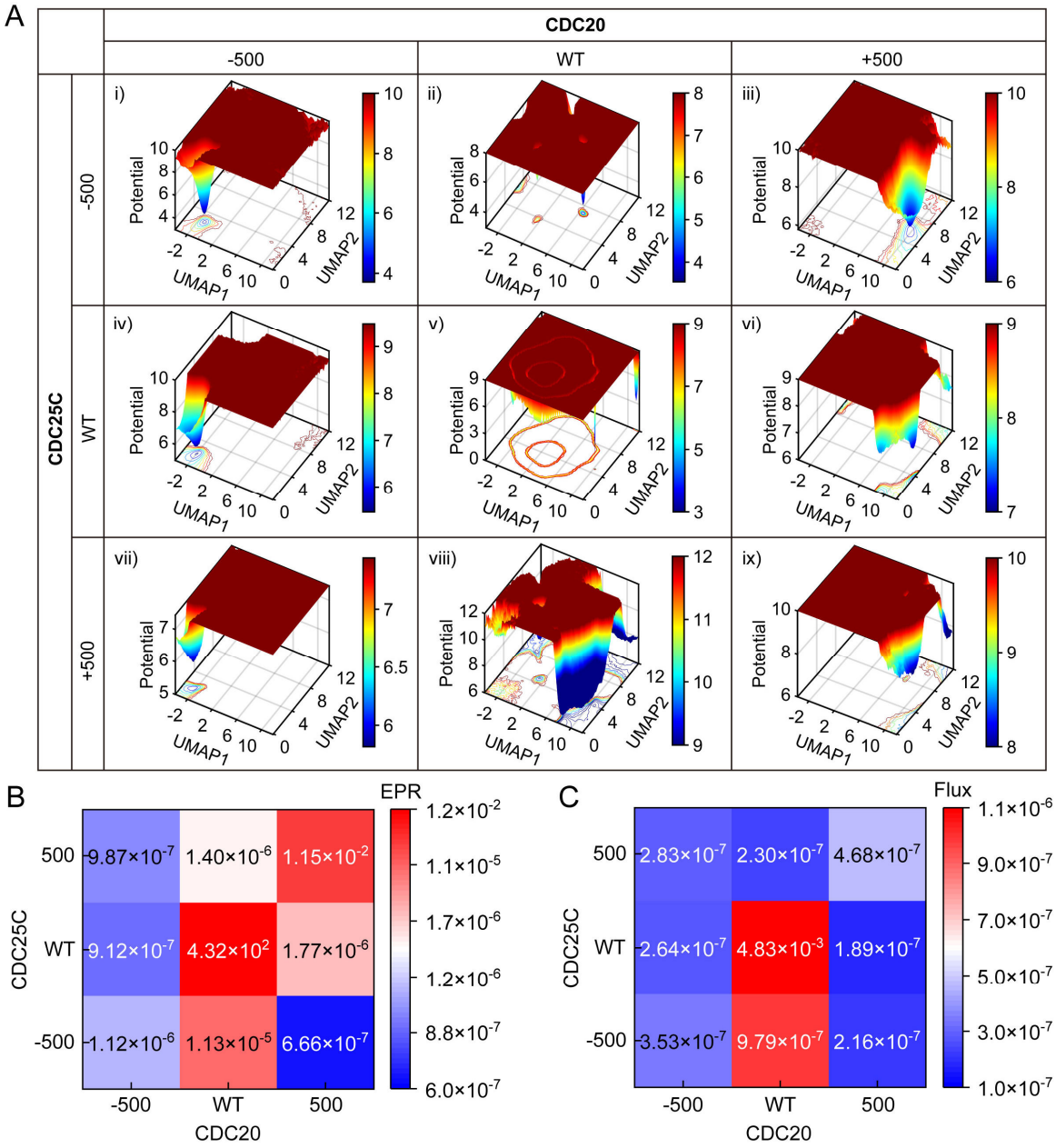


Figure 3. The genetic perturbation alters cell cycle landscape/flux

(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.

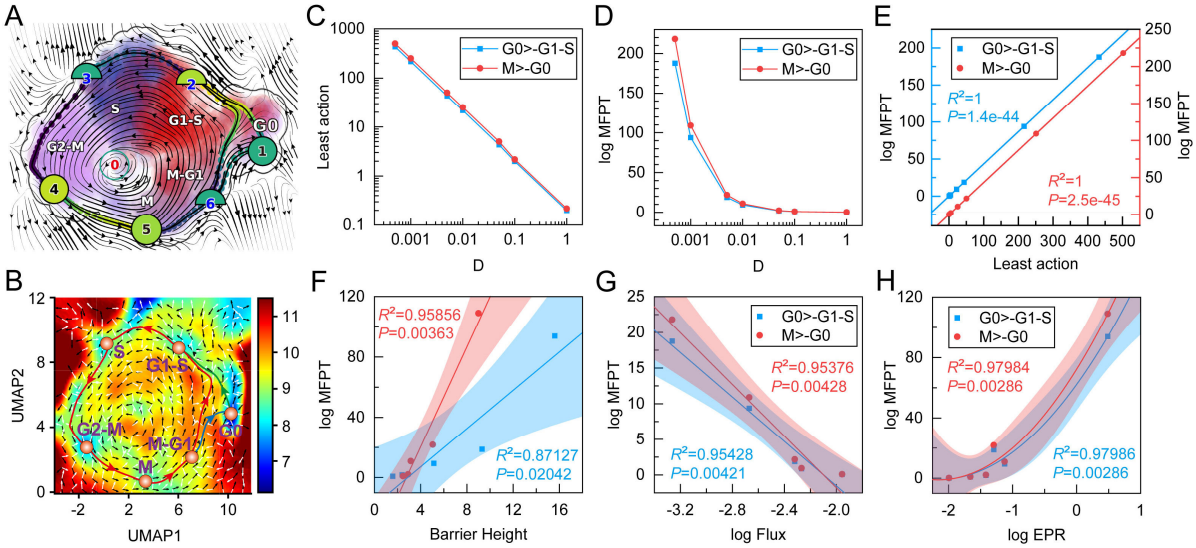


Figure 4. LAPs and MFPT in cell cycle initiation and termination

(A) LAPs between different cell cycle phases in the velocity vector field of U2OS cells. The color of digits in each node reflects the type of fixed point: red, emitting fixed point; black, absorbing fixed point. The color of the numbered nodes corresponds to the confidence of the fixed points. The color of the dots along the paths corresponds to the direction.

(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.

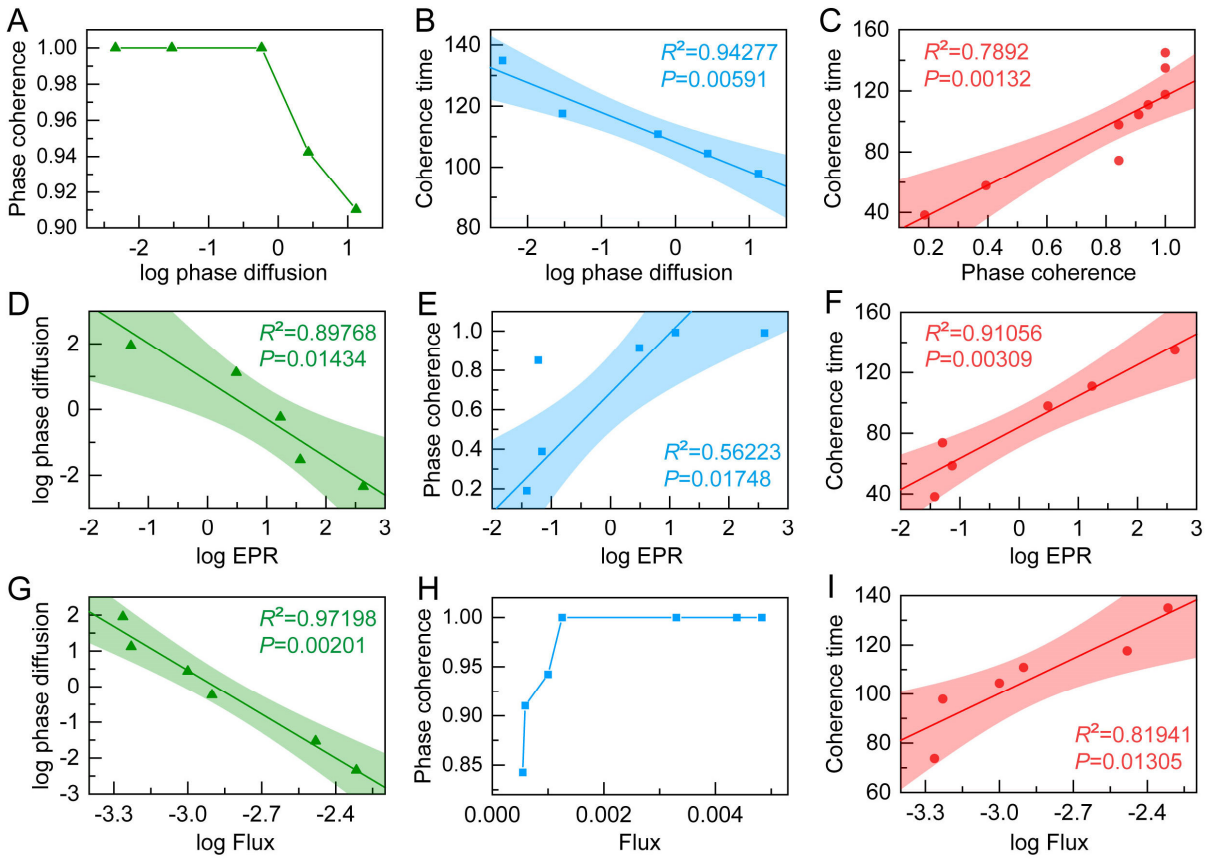


Figure 5. Relationship between phase diffusion, phase coherence, coherence time and EPR and flux

(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.

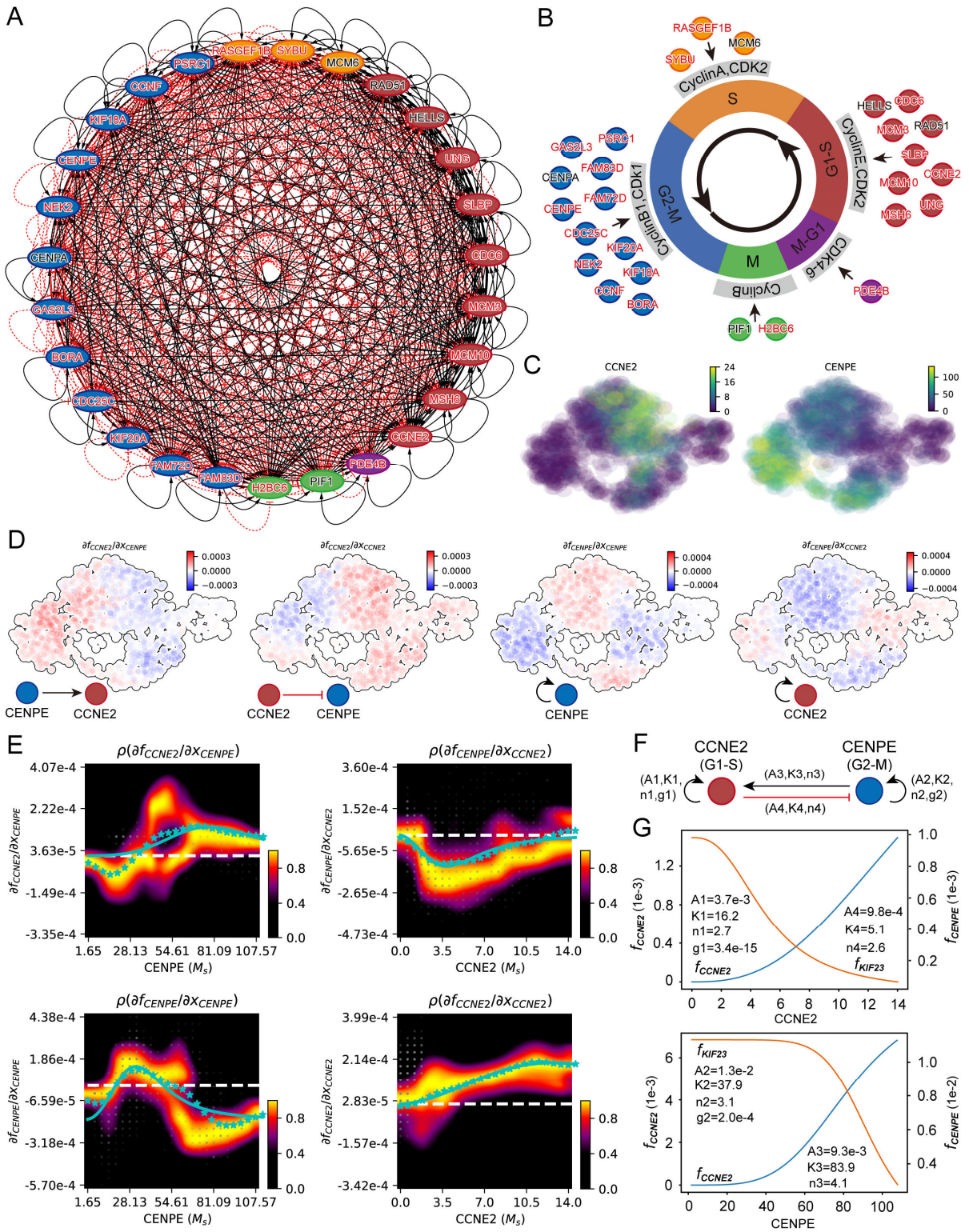


Figure 6. Inference of gene regulatory networks for the cell cycle by single-cell transcriptomics data

(A) The interaction network of genes of each cell cycle phase of U2OS cells. All differentially expressed genes in each cell cycle phase are used to construct network; nodes (genes) 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. Genes with black representing TF and red representing non-TFs.

(B) The diagram for the cell cycle model with key genes 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 expression genes for U2OS cell cycle regulation, related to Figure 6

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 expression genes for RPE1 cell cycle regulation, related to Figure S12

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 Figures 6 and S12

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
	$V_5 = \frac{A_4K_4}{(K_4 + [CCNE2]^{n4})}$	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		
$\frac{d[CCNE2]}{dt} = f_{CCNE2} = V_1 + V_2 - V_3$	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
	$V_2 = \frac{A_4K_4}{(K_4 + [KIF23]^{n4})}$	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

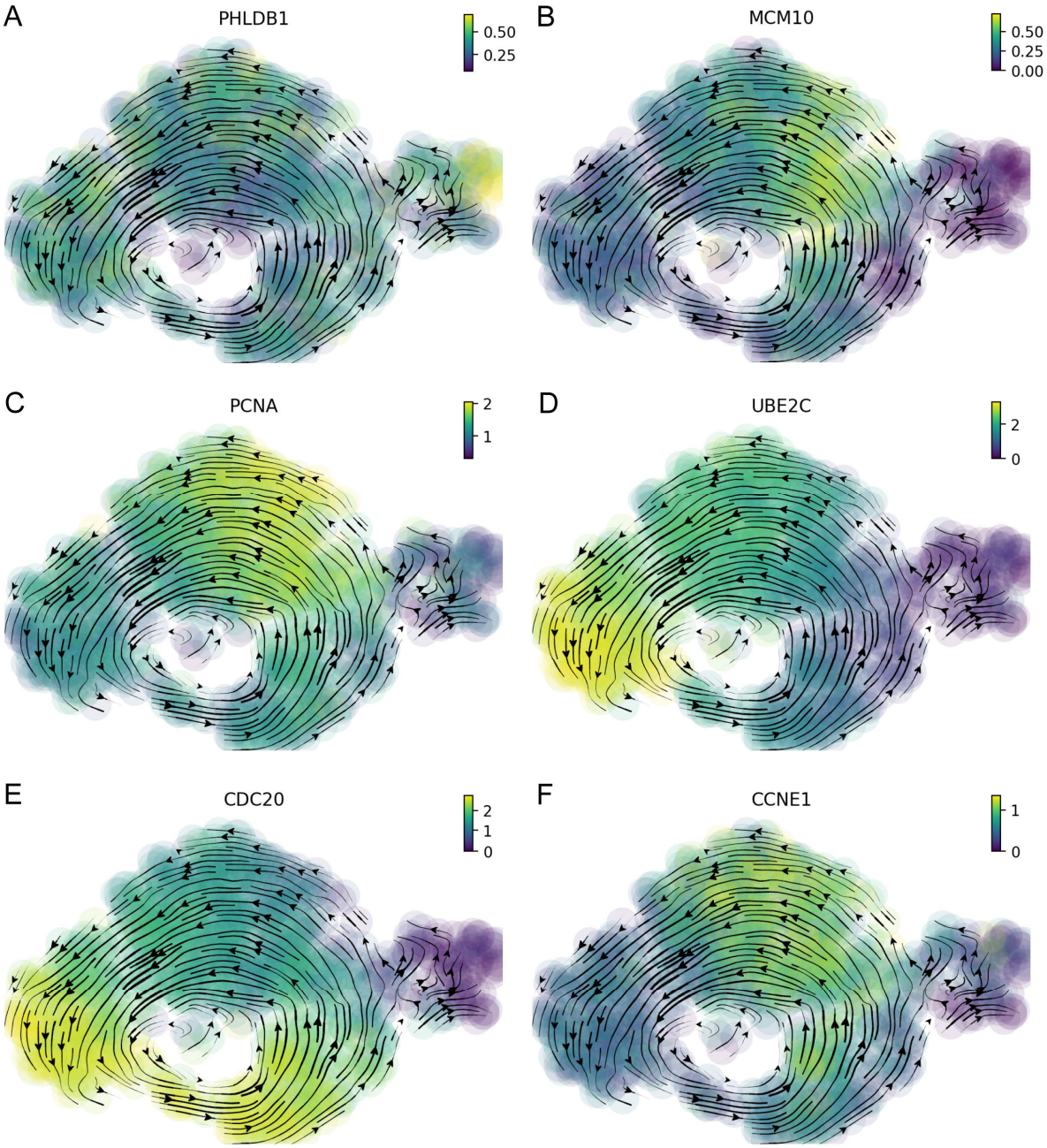


Figure S1. The key gene expression of different cell cycle phases (Related to Figure 1)

(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.

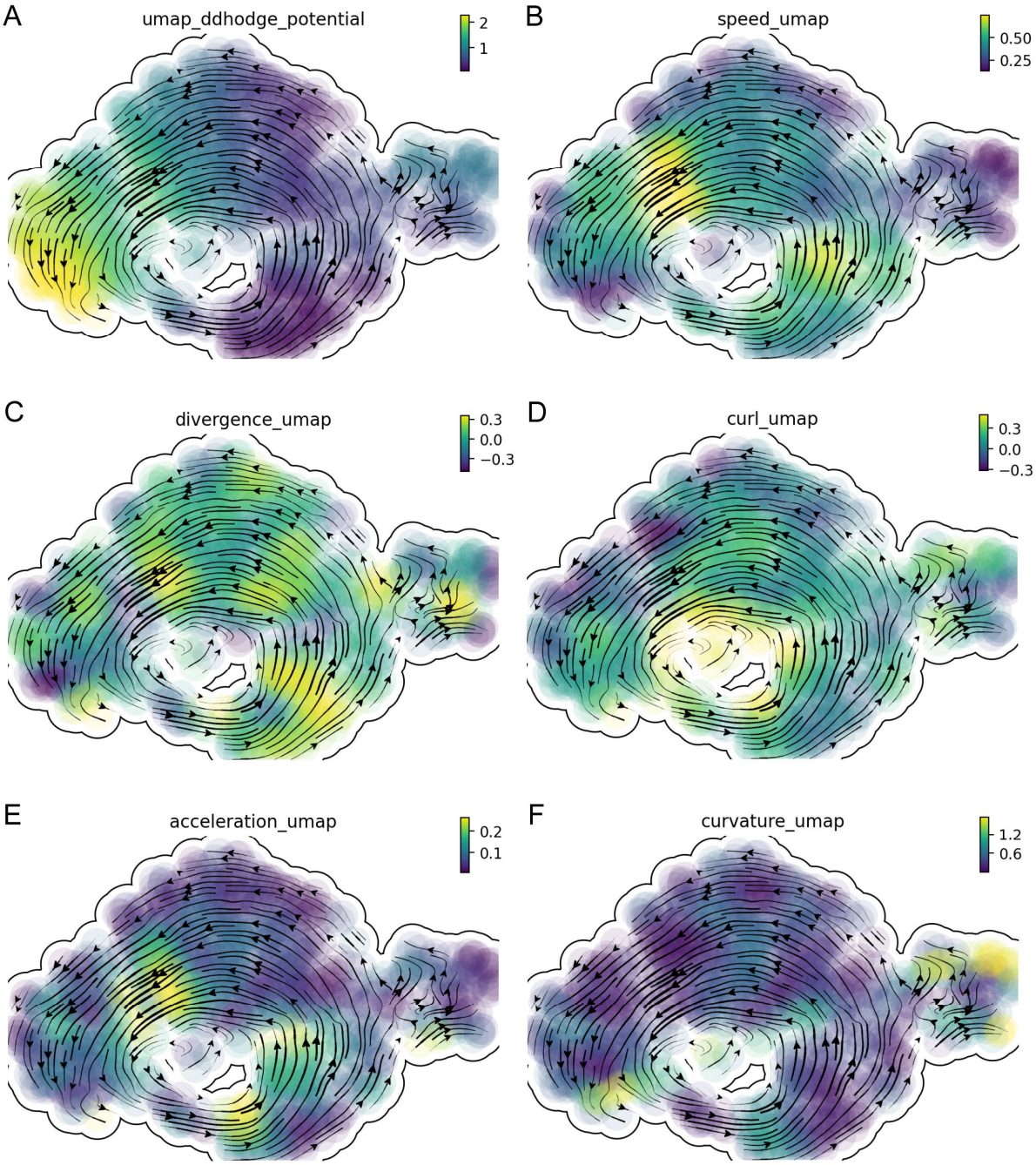


Figure S2. The differential geometric analysis of cell cycle dynamics of U2OS cells (Related to Figure 1)

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

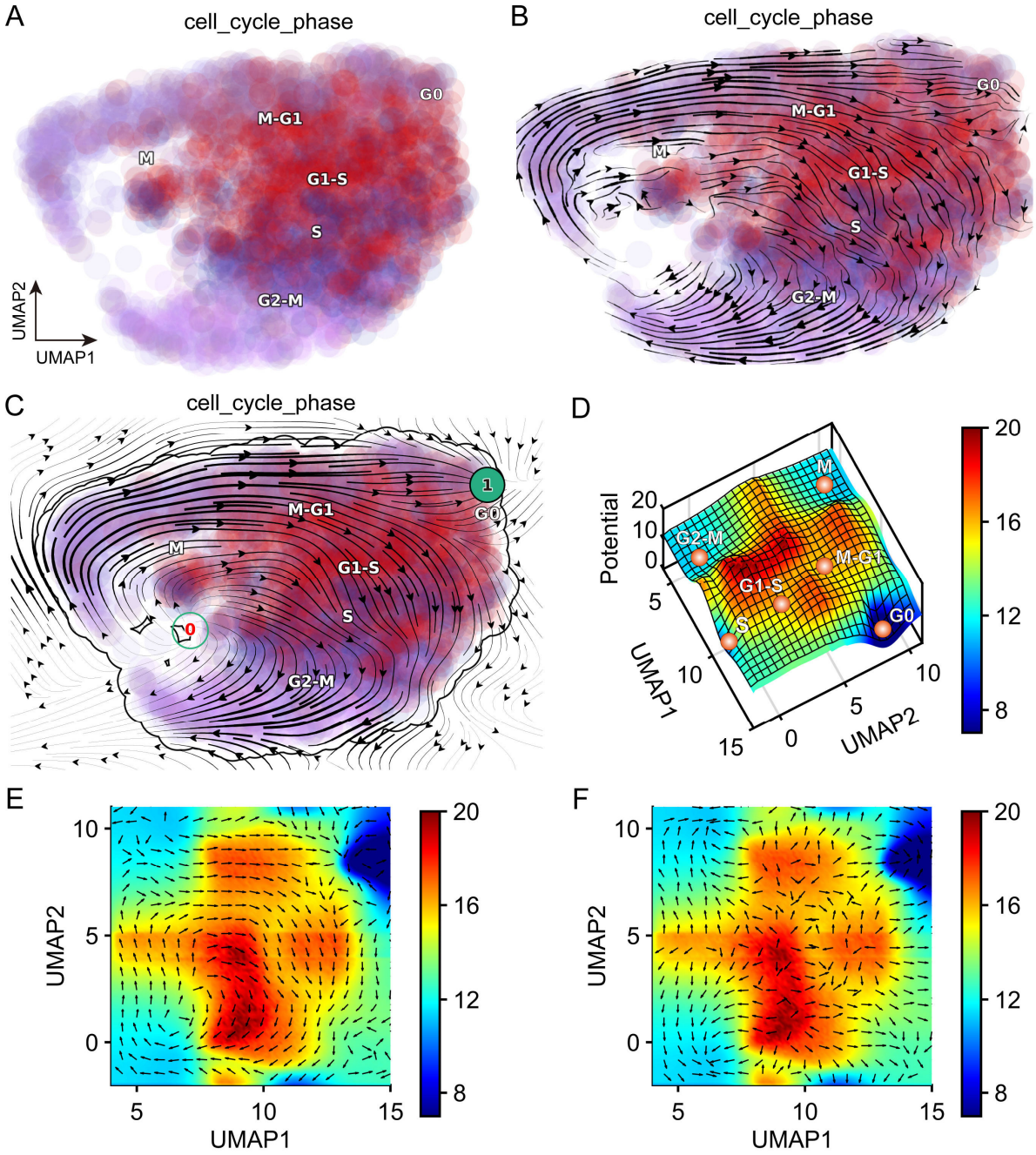
(B) Same as in (A) but for the speed.

(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.



Figutr S3. 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.

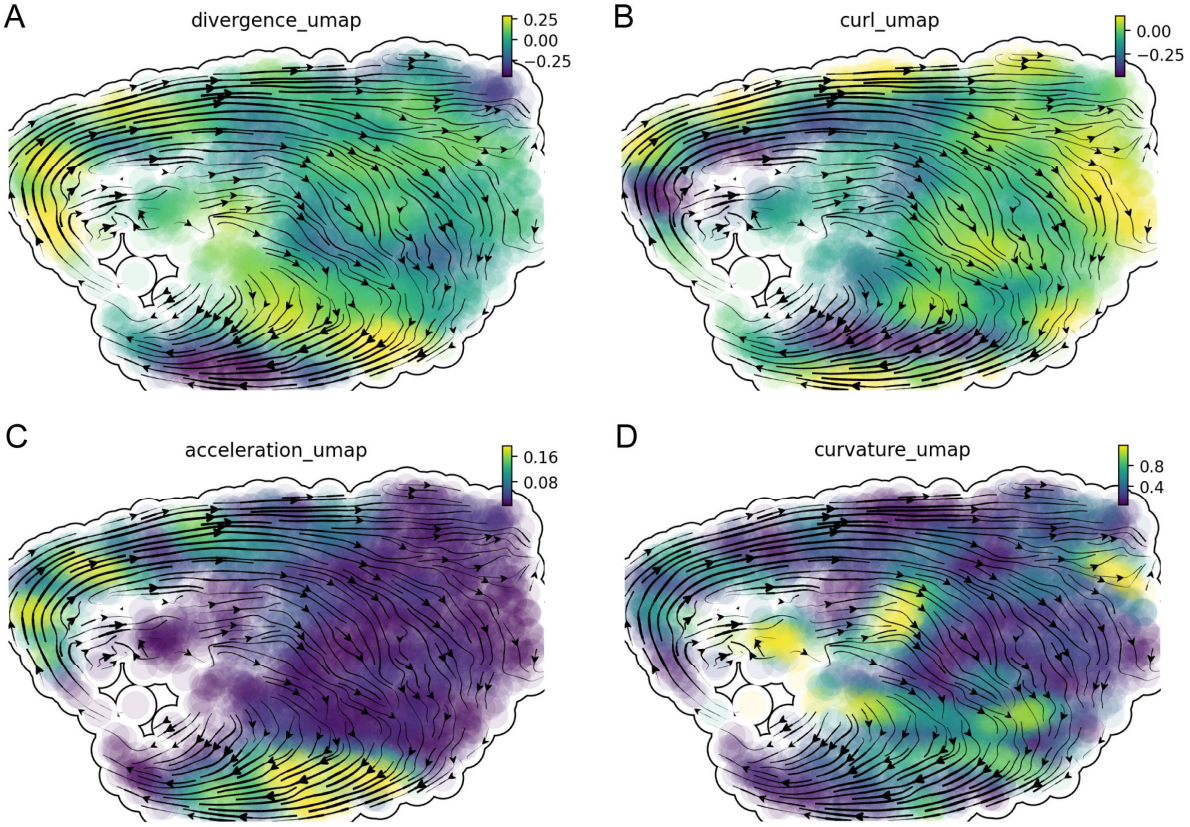


Figure S4. The differential geometric analysis of cell cycle dynamics of RPE1 cell (Related to Figure 1)

(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.

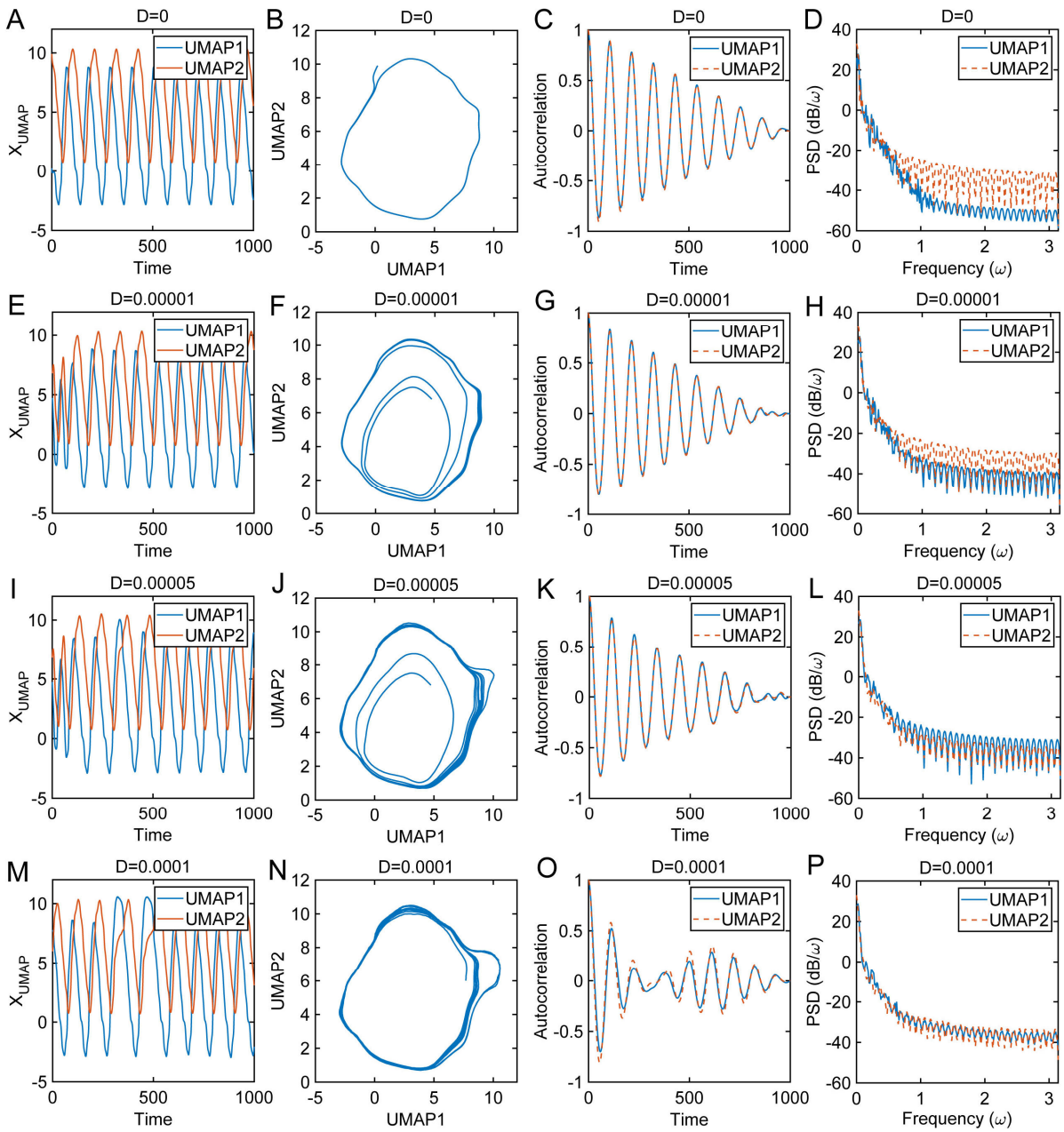


Figure S5. The trajectory, autocorrelation and power spectral density of cell cycle oscillation dynamics with different strength of noise (Related to Figure 2)

(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.

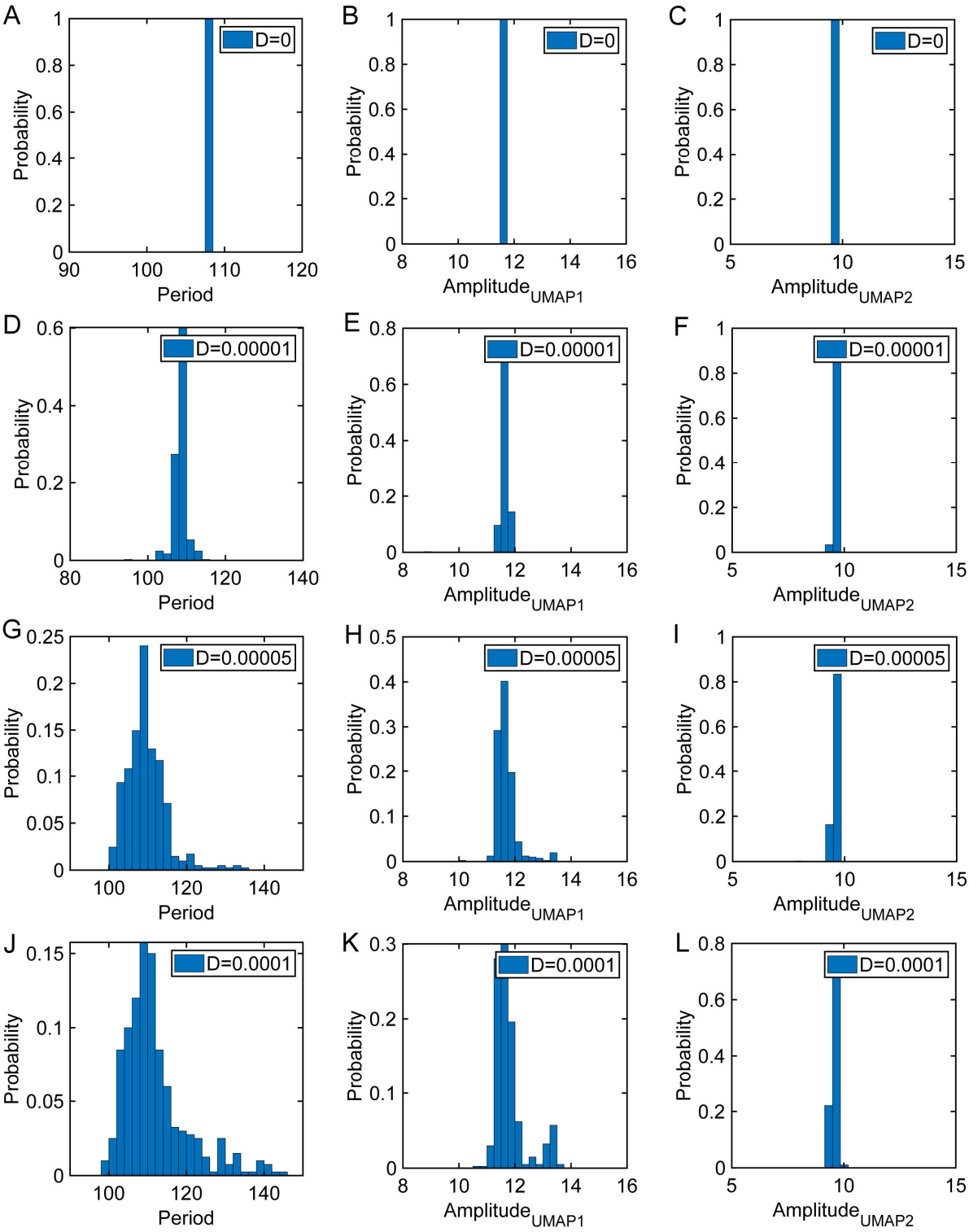

A
B
C
D
E
F
G
H
I
J
K
L
D=0
D=0.00001
D=0.00005
D=0.0001
Probability
Period
Amplitude$_{UMAP1}$
Amplitude$_{UMAP2}$

Figure S6. 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.

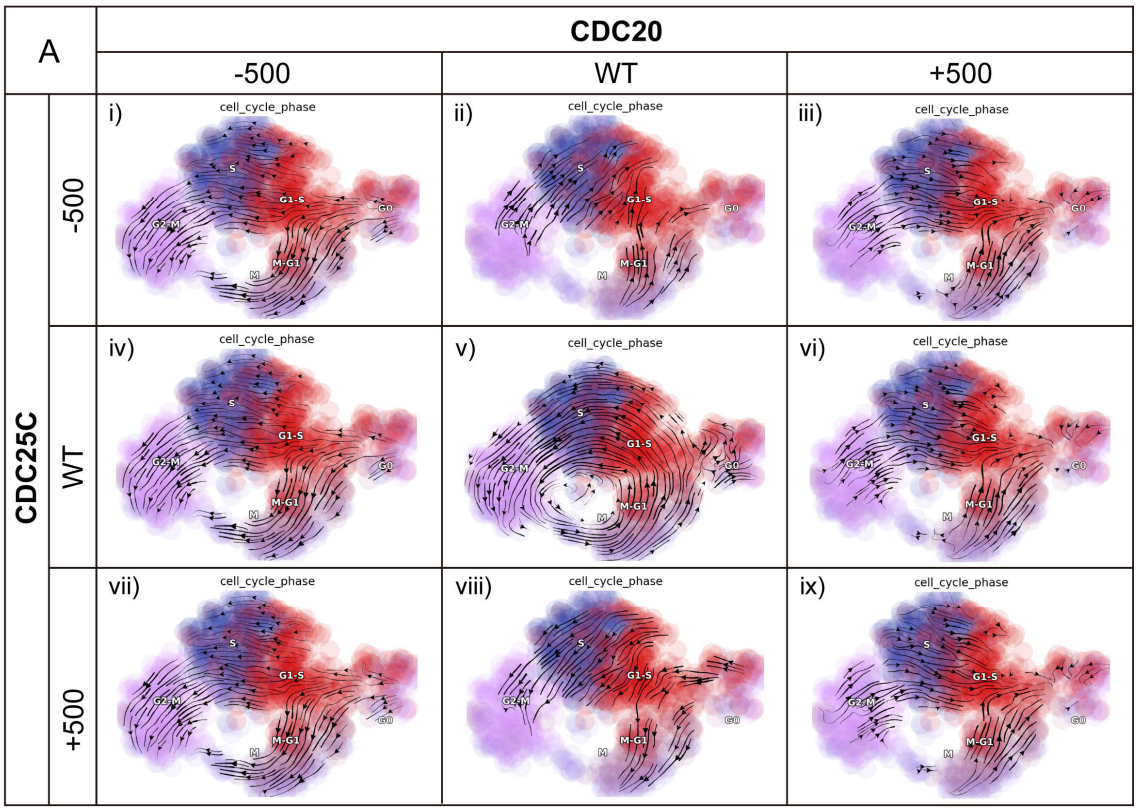
(H) Same as in (B) but for different diffusion coefficient D=0.00005.

(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.


A
CDC20
-500
WT
+500
CDC25C
-500
WT
+500
i)
ii)
iii)
iv)
v)
vi)
vii)
viii)
ix)
cell_cycle_phase
S
G1-S
G0
G2-M
M-G1
M

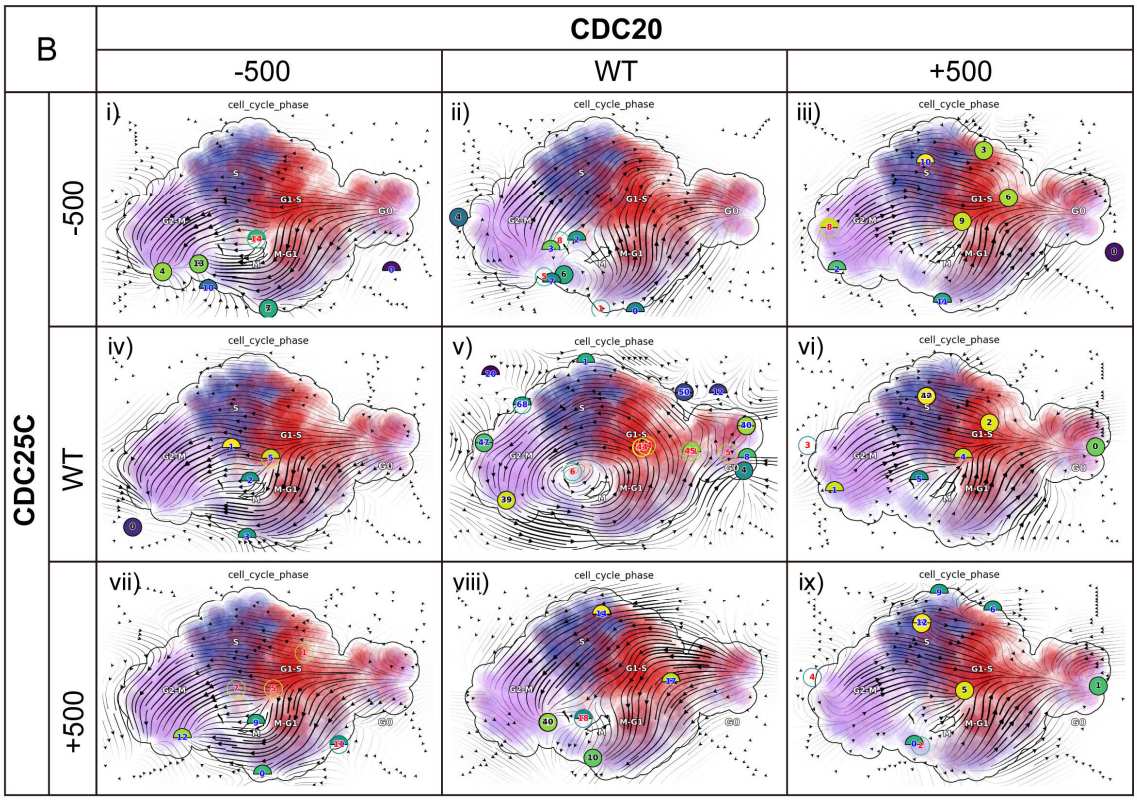

B
CDC20
-500
WT
+500
CDC25C
-500
WT
+500
i)
ii)
iii)
iv)
v)
vi)
vii)
viii)
ix)
cell_cycle_phase
S
G1-S
G0
G2-M
M-G1
M

Figure S7. RNA velocity and reconstructed vector field of cell cycle under genetic perturbation (Related to Figure 3)

(A) *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*.
(B) *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*.

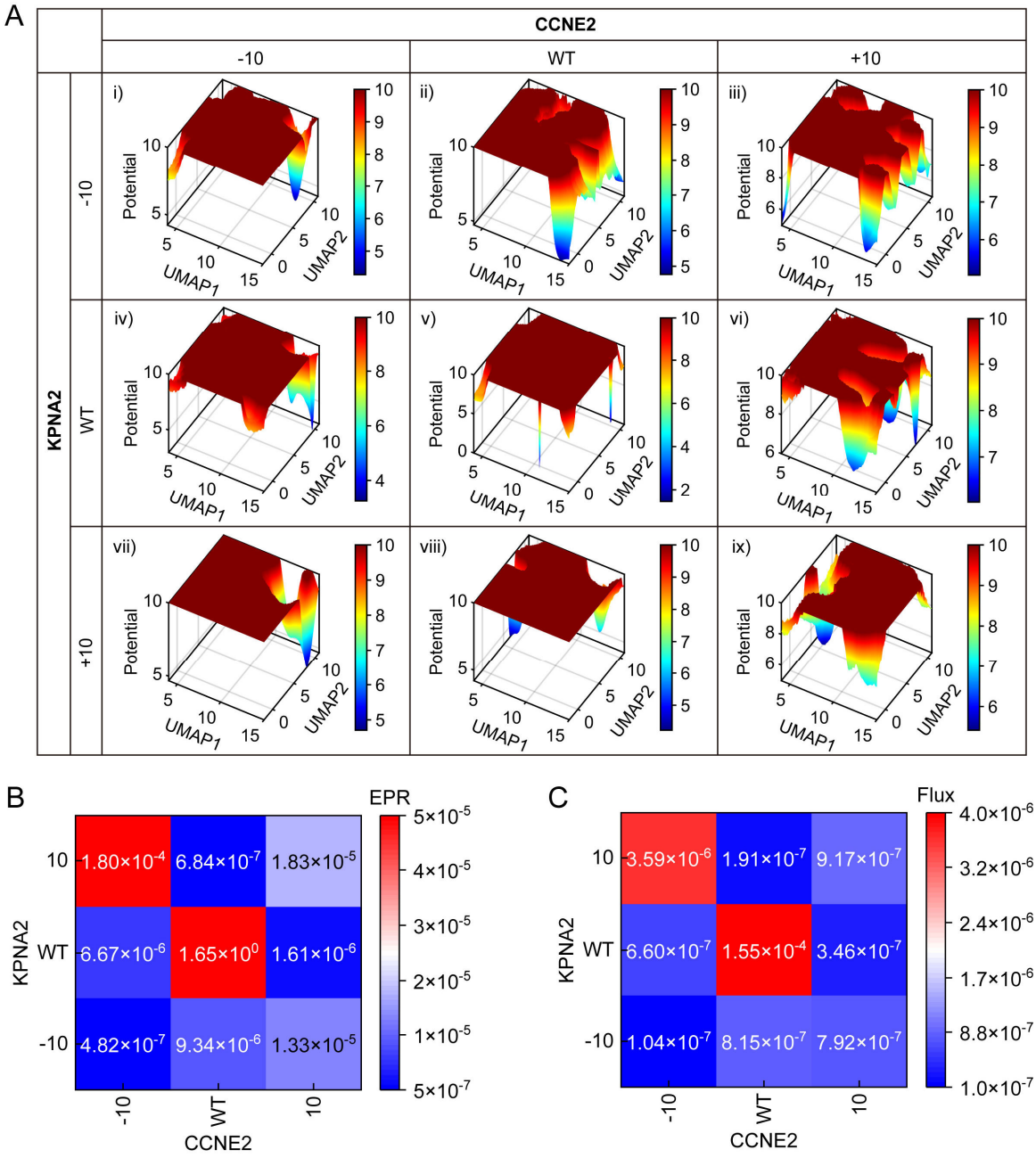


Figure S8. The genetic perturbation alters of RPE1 cell cycle landscape-flux (Related to Figure 3)

(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.

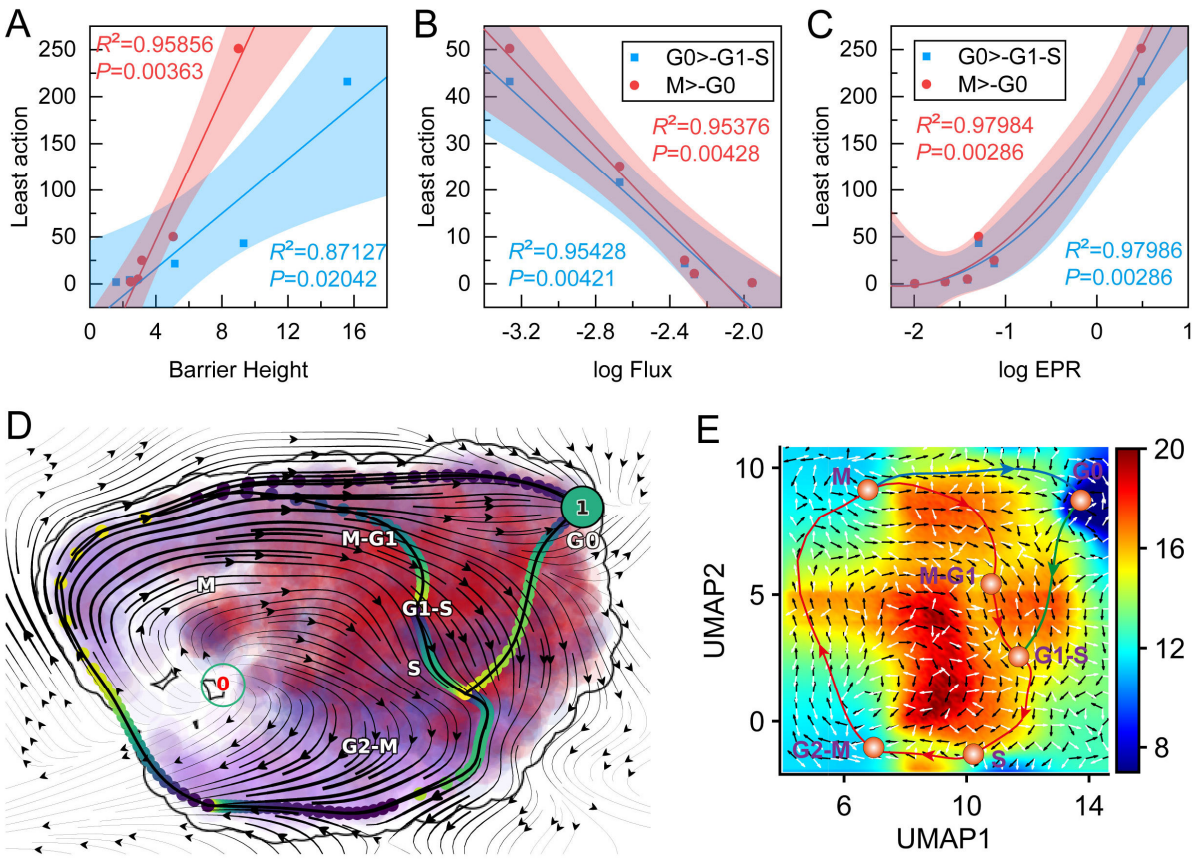


Figure S9. LAPs and MFPT in cell cycle initiation and termination (Related to Figure 4)

(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.

(D) LAPs between different cell cycle phases in the velocity vector field of RPE1 cells. The color of digits in each node reflects the type of fixed point: red, emitting fixed point; black, absorbing fixed point. The color of the numbered nodes corresponds to the confidence of the fixed points. The color of the dots along the paths corresponds to the direction.

(E) 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.

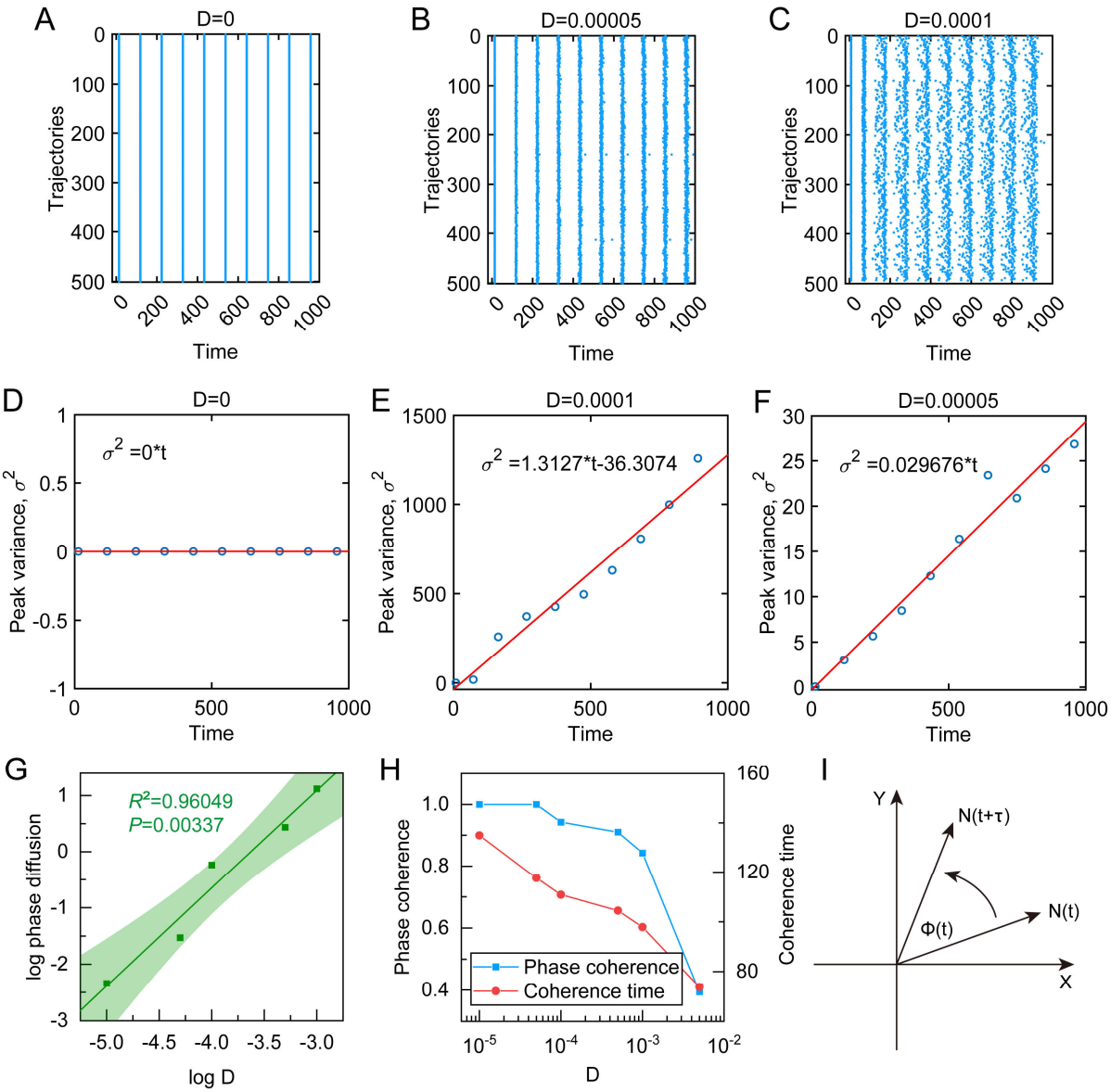


Figure S10. Phase diffusion and phase coherence of cell cycle oscillation with different noise (Related to Figure 5)

(A-C) Raster plot of the peak times for 500 different trajectories starting with the same initial condition at the different diffusion coefficient, D=0 (A), D=0.00005 (B) and D=0.0001 (C).

(D-F) Peak time variance σ^2 goes linearly with the average peak time for the different diffusion coefficient, with the linear coefficient defined as the peak-time diffusion constant, D=0 (D), D=0.00005 (E) and D=0.0001 (F).

(G) 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.

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

(I) Sketch map for the definition of phase coherence.

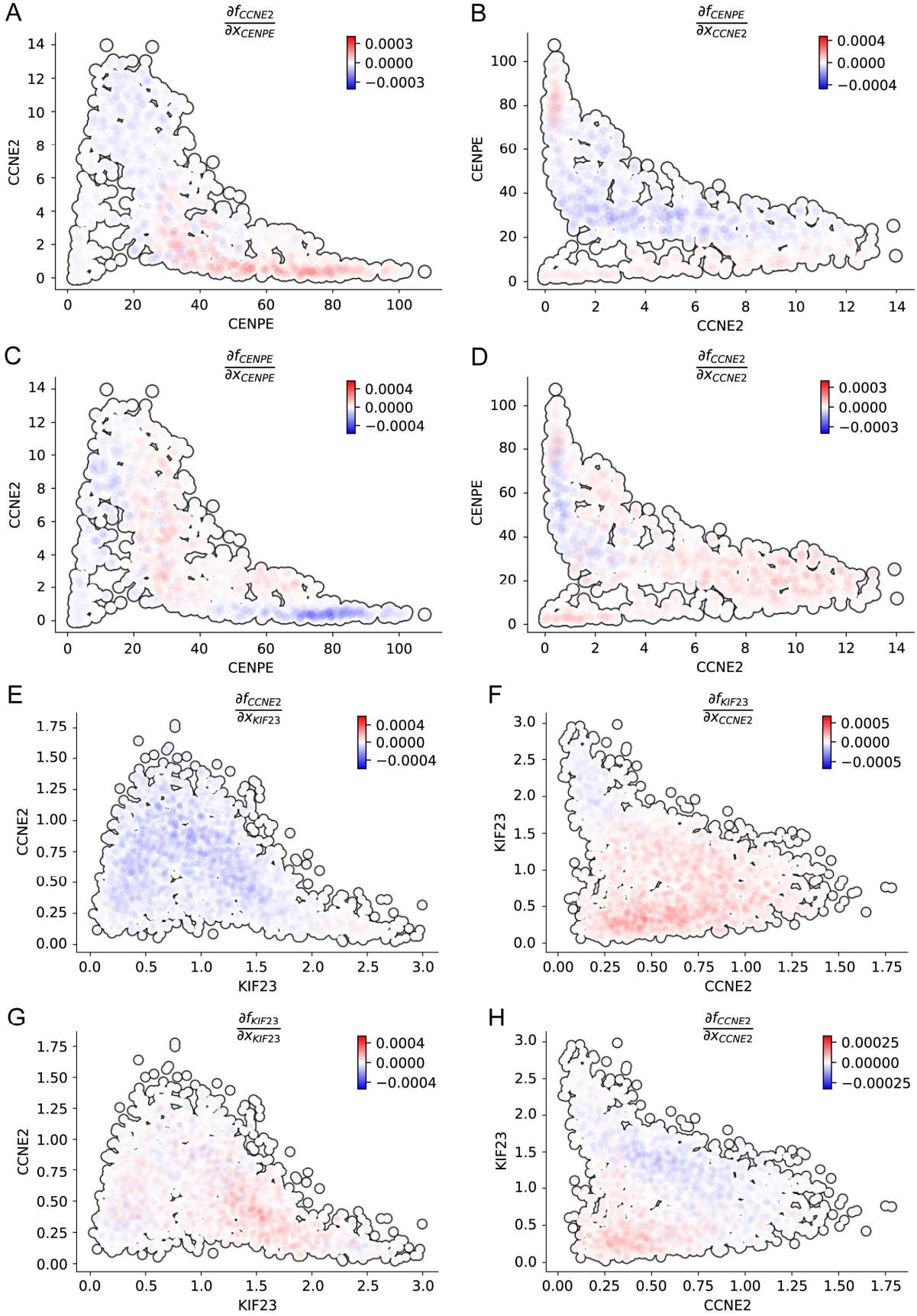

A
$\frac{\partial f_{CCNE2}}{\partial x_{CENPE}}$
0.0003
0.0000
−0.0003
CCNE2
CENPE
B
$\frac{\partial f_{CENPE}}{\partial x_{CCNE2}}$
0.0004
0.0000
−0.0004
CENPE
CCNE2
C
$\frac{\partial f_{CENPE}}{\partial x_{CENPE}}$
0.0004
0.0000
−0.0004
CCNE2
CENPE
D
$\frac{\partial f_{CCNE2}}{\partial x_{CCNE2}}$
0.0003
0.0000
−0.0003
CENPE
CCNE2
E
$\frac{\partial f_{CCNE2}}{\partial x_{KIF23}}$
0.0004
0.0000
−0.0004
CCNE2
KIF23
F
$\frac{\partial f_{KIF23}}{\partial x_{CCNE2}}$
0.0005
0.0000
−0.0005
KIF23
CCNE2
G
$\frac{\partial f_{KIF23}}{\partial x_{KIF23}}$
0.0004
0.0000
−0.0004
CCNE2
KIF23
H
$\frac{\partial f_{CCNE2}}{\partial x_{CCNE2}}$
0.00025
0.00000
−0.00025
KIF23
CCNE2

Figure S11. Jacobian analyses of the cell cycle regulatory interactions in the gene expression space (Related to Figure 6)

(A-D) Jacobian analyses of the U2OS cell cycle regulatory interactions in the CCNE2 and CENPE expression space. Activation from *CENPE* to *CCNE2* in the *CENPE* and *CCNE2* expression space (A), repression from *CCNE2* to *CENPE* in the *CCNE2* and *CENPE* expression space (B), (C) self-activation of *CENPE* in the *CENPE* and *CCNE2* expression space (C), and self-activation of *CCNE2* in the *CCNE2* and *CENPE* expression space (D).

(E-H) Jacobian analyses of the U2OS cell cycle regulatory interactions in the CCNE2 and KIF23 expression space. Repression from *KIF23* to *CCNE2* in the *KIF23* and *CCNE2* expression space (E), activation from *CCNE2* to *KIF23* in the *CCNE2* and *KIF23* expression space (F), self-activation of *KIF23* in the *KIF23* and *CCNE2* expression space (G), and self-activation of *CCNE2* in the *CCNE2* and *KIF23* expression space (H).

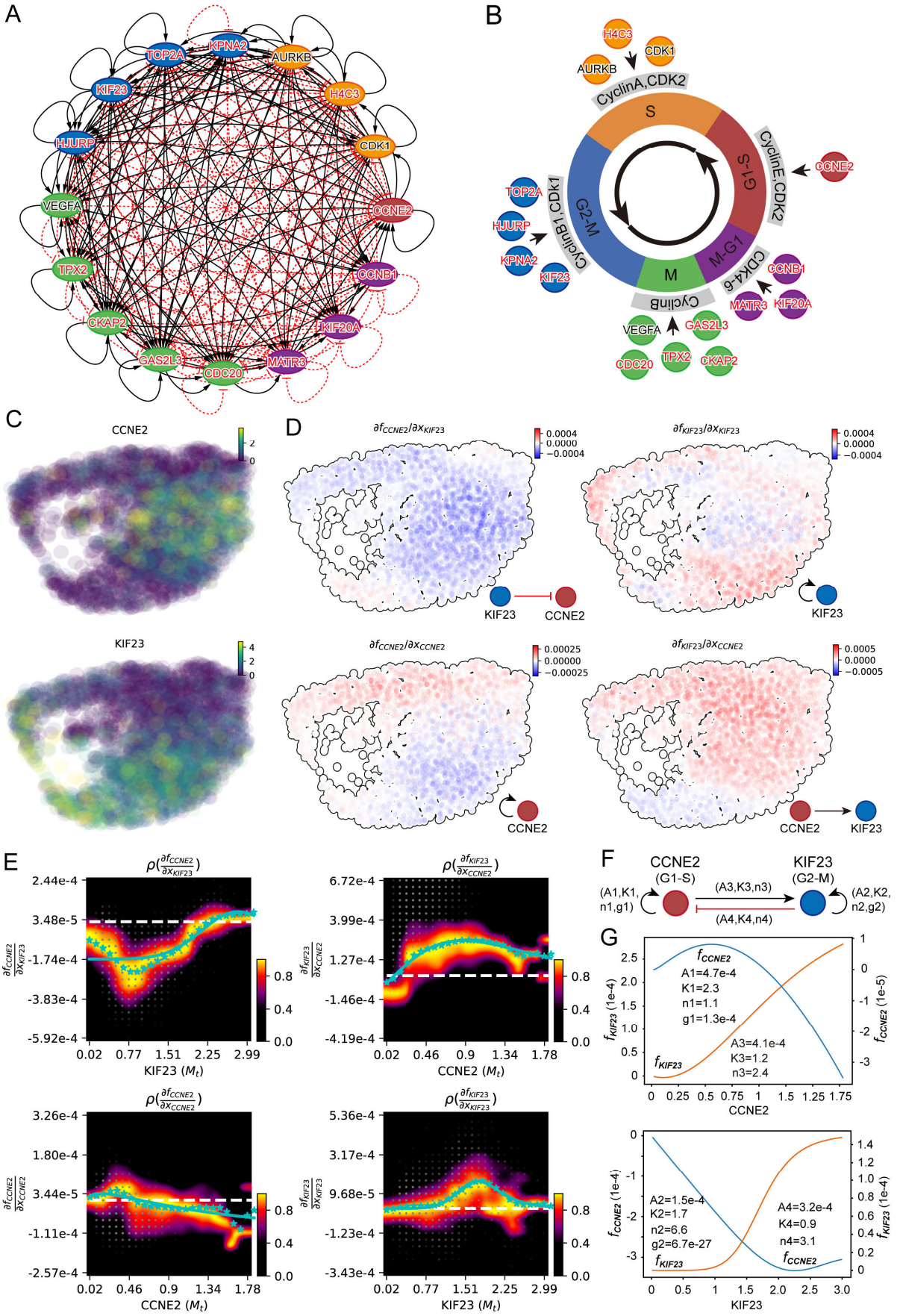

A
TOP2A
KPNA2
AURKB
KIF23
H4C3
HJURP
CDK1
VEGFA
CCNE2
TPX2
CCNB1
CKAP2
KIF20A
GAS2L3
CDC20
MATR3
B
H4C3
CDK1
AURKB
CyclinA,CDK2
S
CyclinE,CDK2
G1-S
CCNE2
TOP2A
HJURP
KPNA2
KIF23
CyclinB1,CDK1
G2-M
M
M-G1
CDK4-6
CCNB1
MATR3
KIF20A
CyclinB
VEGFA
GAS2L3
CDC20
TPX2
CKAP2
C
CCNE2
KIF23
D
$\partial f_{CCNE2}/\partial x_{KIF23}$
$\partial f_{KIF23}/\partial x_{KIF23}$
$\partial f_{CCNE2}/\partial x_{CCNE2}$
$\partial f_{KIF23}/\partial x_{CCNE2}$
E
$\rho(\frac{\partial f_{CCNE2}}{\partial x_{KIF23}})$
$\rho(\frac{\partial f_{KIF23}}{\partial x_{CCNE2}})$
$\rho(\frac{\partial f_{CCNE2}}{\partial x_{CCNE2}})$
$\rho(\frac{\partial f_{KIF23}}{\partial x_{KIF23}})$
KIF23 (M_t)
CCNE2 (M_t)
F
CCNE2
(G1-S)
KIF23
(G2-M)
(A1,K1, n1,g1)
(A3,K3,n3)
(A4,K4,n4)
(A2,K2, n2,g2)
G
f_{CCNE2}
A1=4.7e-4
K1=2.3
n1=1.1
g1=1.3e-4
f_{KIF23}
A3=4.1e-4
K3=1.2
n3=2.4
CCNE2
A2=1.5e-4
K2=1.7
n2=6.6
g2=6.7e-27
A4=3.2e-4
K4=0.9
n4=3.1
KIF23

Figure S12. Inference of gene regulatory networks for the RPE1 cell cycle by single-cell transcriptomics data (Related to Figure 6)

(A) The interaction network of genes of each cell cycle phase of RPE1 cells. All differentially expressed genes in each cell cycle phase are used to construct network; nodes (genes) 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. Genes with black representing TF and red representing non-TFs.
(B) The diagram for the cell cycle model with key genes 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).