PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Cell cycle time series gene expression data encoded as cyclic attractors in Hopfield systems

Anthony Szedlak[1], Spencer Sims[1], Nicholas Smith[2], Giovanni Paternostro[3], Carlo Piermarocchi[1]*

1 Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan, United States of America, **2** Salgomed Inc., Del Mar, California, United States of America, **3** Sanford Burnham Prebys Medical Discovery Institute, La Jolla, California, United States of America

* carlo@pa.msu.edu

Citation: Szedlak A, Sims S, Smith N, Paternostro G, Piermarocchi C (2017) Cell cycle time series gene expression data encoded as cyclic attractors in Hopfield systems. PLoS Comput Biol 13(11): e1005849. https://doi.org/10.1371/journal.pcbi.1005849

Editor: Andrey Rzhetsky, University of Chicago, UNITED STATES

Received: July 24, 2017

Accepted: October 25, 2017

Published: November 17, 2017

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was funded through the National Institutes of Health/National Institute of General Medical Sciences, grant number R01GM122085 (received by CP and GP). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Modern time series gene expression and other omics data sets have enabled unprecedented resolution of the dynamics of cellular processes such as cell cycle and response to pharmaceutical compounds. In anticipation of the proliferation of time series data sets in the near future, we use the Hopfield model, a recurrent neural network based on spin glasses, to model the dynamics of cell cycle in HeLa (human cervical cancer) and *S. cerevisiae* cells. We study some of the rich dynamical properties of these cyclic Hopfield systems, including the ability of populations of simulated cells to recreate experimental expression data and the effects of noise on the dynamics. Next, we use a genetic algorithm to identify sets of genes which, when selectively inhibited by local external fields representing gene silencing compounds such as kinase inhibitors, disrupt the encoded cell cycle. We find, for example, that inhibiting the set of four kinases *AURKB*, *NEK1*, *TTK*, and *WEE1* causes simulated HeLa cells to accumulate in the M phase. Finally, we suggest possible improvements and extensions to our model.

Author summary

Cell cycle—the process in which a parent cell replicates its DNA and divides into two daughter cells—is an upregulated process in many forms of cancer. Identifying gene inhibition targets to regulate cell cycle is important to the development of effective therapies. Although modern high throughput techniques offer unprecedented resolution of the molecular details of biological processes like cell cycle, analyzing the vast quantities of the resulting experimental data and extracting actionable information remains a formidable task. Here, we create a dynamical model of the process of cell cycle using the Hopfield model (a type of recurrent neural network) and gene expression data from human cervical cancer cells and yeast cells. We find that the model recreates the oscillations observed in experimental data. Tuning the level of noise (representing the inherent randomness in gene expression and regulation) to the "edge of chaos" is crucial for the proper behavior

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: NS is an employee of Salgomed Inc., and CP and GP own equity in Salgomed Inc.

of the system. We then use this model to identify potential gene targets for disrupting the process of cell cycle. This method could be applied to other time series data sets and used to predict the effects of untested targeted perturbations.

Introduction

Originally proposed by Conrad Waddington in the 1950s [1] and Stuart Kauffman in the 1970s [2], analysis of biological processes such as cellular differentiation and cancer development using attractor models—dynamical systems whose configurations tend to evolve toward particular sets of states—has gained significant traction over the past decade [3–12]. One such attractor model, the Hopfield model [13], is a type of recurrent artificial neural network based on spin glasses. It was designed with the ability to recall a host of memorized patterns from noisy or partial input information by mapping data directly to attractor states. A great deal of analytical and numerical work has been devoted to understanding the statistical properties of the Hopfield model, including its storage capacity [14], correlated patterns [15], spurious attractors [16], asymmetric connections [17], embedded cycles [18], and complex transition landscapes [19]. Due to its prescriptive, data-driven design, the Hopfield model has been applied in a variety of fields including image recognition [20, 21] and the clustering of gene expression data [22]. It has also been used to directly model the dynamics of cellular differentiation and stem cell reprogramming [23, 24], targeted inhibition of genes in cancer gene regulatory networks [25], and cell cycle across various stages of cellular differentiation [26].

Techniques for measuring large scale omics data, particularly transcriptomic data from microarrays and RNA sequencing (RNA-seq), have become standard, indispensable tools for observing the states of complex biological systems [27–29]. However, analysis of the sheer variety and vast quantities of data these techniques produce requires the development of new mathematical tools. Inference and topological analysis of gene regulatory networks has garnered much attention as a method for distilling meaningful information from large datasets [30–36], but simply analyzing the topology of static networks without a signaling rule (e.g. differential equations, digital logic gates, or discrete maps) fails to capture the nonlinear dynamics crucial to cellular behavior. The non-equilibrium nature of life implies that it can only be truly understood at the dynamical level, necessitating the development of new methods for analyzing time series data. As experimental methods continue to improve, more and more high-resolution time series omics and even multi-omics [37] data sets will inevitably become available. Here, we demonstrate that time series omics data (in this case, transcriptomic data) representing cyclic biological processes can be encoded in Hopfield systems, providing a new model for analyzing the dynamics of, and exploring effects of perturbations to, such systems.

The dynamics of cell cycle (CC)—the process in which a parent cell replicates its DNA and divides into two daughter cells—is both scientifically interesting and therapeutically important, and has been modeled extensively using differential equations, Boolean models, and discrete maps [38–55]. Even relatively simple simulated systems such as an isolated, positively self-regulating gene subject to noise can exhibit rich dynamical behavior [56]; but like many biological processes, the proper functioning of CC requires the decentralized, coordinated action of hundreds of genes. CC thus provides researchers with a convenient case study of self-organization in a noisy environment. CC is also an upregulated process in many forms of cancer [57–60], and control of CC using pharmaceutical compounds such as kinase inhibitors is a critical goal in cancer research. The combinatorics of selectively inhibiting sets of genes makes

exhaustive experimental searches difficult or impractical [61]. However, network-based mathematical models such as the one presented here enable researchers to examine the effects of perturbations to complex systems [62, 63] by testing potential inhibition targets *in silico*. The efficacy of these predictions can then be experimentally validated or invalidated, providing new information and insights to further refine models.

The remainder of this article is structured as follows. In the Models subsection of the Results section we first discuss how periodic genes were identified in the time series gene expression data sets, and how Boolean attractors were extracted from the continuous data (explained in greater detail in the Methods section). We then introduce the Hopfield model and discuss the specific form of the coupling matrix used in this application. We discuss how to interpret the results of Hopfield simulations in the context of gene expression and cells. We also explain the objective function used by the genetic algorithm to identify potential inhibition targets, designed with the intention of disrupting CC. In the Dynamical behavior subsection, we show that this model qualitatively recreates experimental gene expression data, and we demonstrate and analyze some dynamical properties of the delayed Hopfield model, including the role played by noise. We include supplementary videos to emphasize the dynamical nature of this model. Optimal control fields for both unconstrained searches (in which any gene may be inhibited) and searches constrained to kinases are discussed. Finally, we recap our results and suggest possible improvements and generalizations to our methods in the Discussion section.

Results

Model

Periodic gene selection. Microarray and RNA-seq time series data sets were obtained from Eser et al. (*S. cerevisiae*) [64] and Dominguez et al. (HeLa, human cervical cancer) [65]. For consistency and due to its higher resolution, the *S. cerevisiae* data set was chosen to produce all images and movies in this article, but both data sets were analyzed. In order to encode these CC data sets into the Hopfield model, periodic genes needed to be identified, their frequencies and phases computed, and their expression converted from continuous to Boolean form. As detailed in the Methods section, decaying sinusoids were fitted to the trajectory of each gene i, and genes with sufficiently high quality fits were kept. This resulted in 379 periodic genes in *S. cerevisiae* and 722 periodic genes in HeLa cells. Fig 1A shows a heat map of the expression of all periodic genes detected in the Eser data set sorted by their fitted phases, and Fig 1B shows the same genes with the fitted expression curves. These fitted curves were converted from continuous values $x_i(t) \geq 0$ to Boolean values $\xi_i(t) = \pm 1$ (over/underexpressed) as shown in Fig 1C. Finally, one CC period was divided into eight uniformly spaced states $\{\xi_i^\mu\} = \{\xi_i^0, \xi_i^1, \ldots, \xi_i^7\}$ with $\xi_i^\mu = \pm 1$. These states, shown in Fig 1D, were used as the eight attractor patterns in the Hopfield model. $p = 8$ was chosen because the embedded cycle lost stability for $p > 8$ due to the finite capacity of the Hopfield model.

Note that this work is based on the simplifying assumption that cell cycle genes exhibit periodic expression, in line with other studies of cell cycle genes [64–70]. This method cannot detect non-periodic cell cycle effectors, but our goal here is to model cell cycle based solely on time series gene expression data.

The Hopfield model. The Hopfield model [13] is an Ising model whose configuration is defined by N spins $\sigma_i(t)$ at integer time t. The state of each node (gene) takes one of two values, $\sigma_i(t) = \pm 1$ (over/underexpressed). The coupling matrix J_{ij} defines the strength and sign of the signal sent from node j to node i, and its construction is discussed in the following subsection.

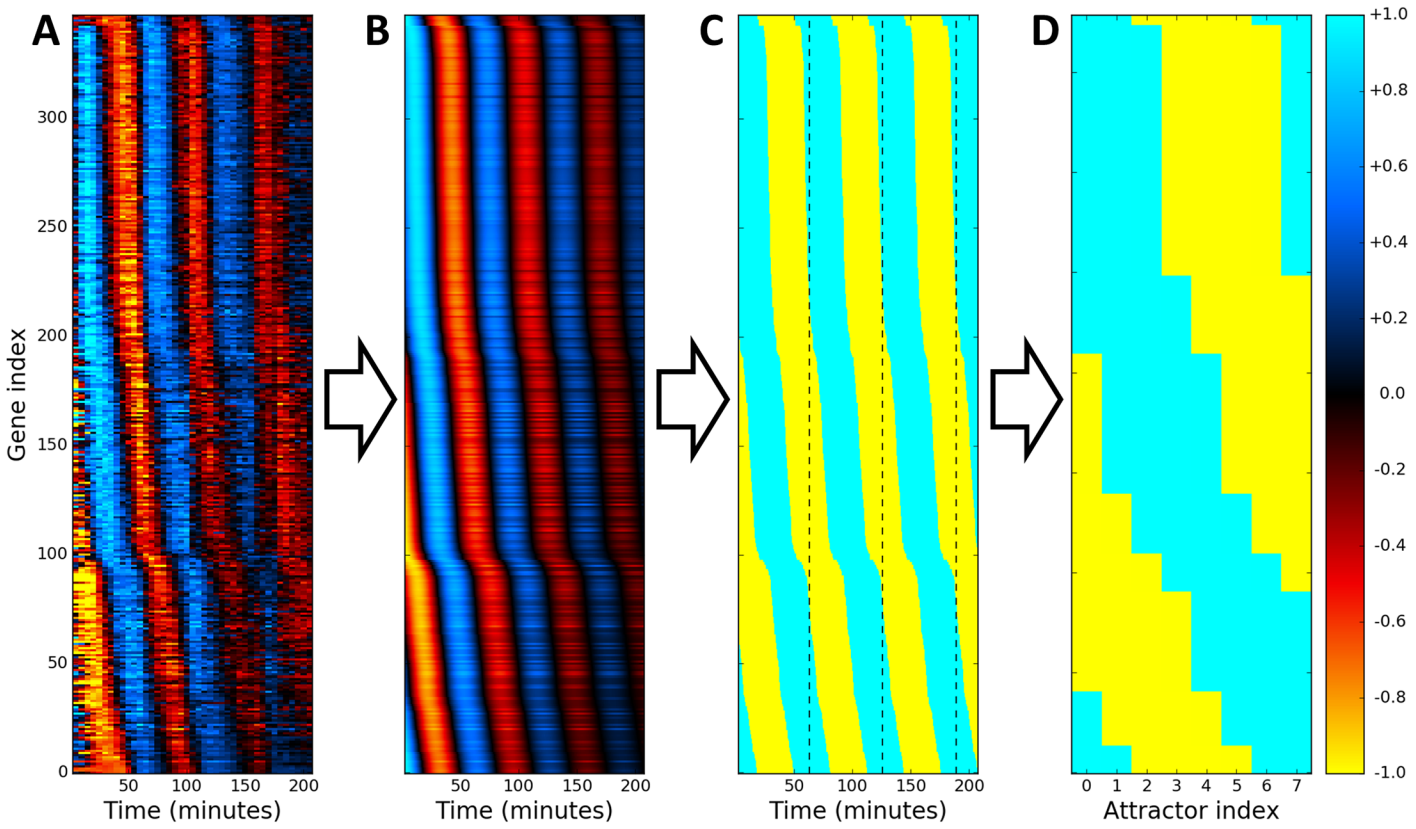

Fig 1. Obtaining attractors from expression data. (A) Heat map of the expression of all detected periodic genes from [64] sorted by their fitted phases. (B) Fitted gene expression. (C) Boolean form of fitted expression, separated into periods by dashed black lines. (D) Final set of $p = 8$ attractors taken from one period.

https://doi.org/10.1371/journal.pcbi.1005849.g001

The total field at node i at time t is given by

$$h_i(t) = \sum_j J_{ij}\sigma_j(t) + h_i^{\text{ext}}, \tag{1}$$

where $\sum_j J_{ij}\,\sigma_j(t)$ is the internal field at node i due to its coupling with all nodes j and h_i^{ext} is an optional external field applied to node i representing the action of therapeutic compounds, e.g. kinase inhibitors. The dynamical update rule is given by

$$\sigma_i(t+1) = \begin{cases} +1 & \text{with probability } \left(1 + e^{-2h_i(t)/T}\right)^{-1} \\ -1 & \text{otherwise} \end{cases}, \tag{2}$$

where the factor of 2 in the exponent is conventional and T is an effective temperature representing the level of noise (not a physical temperature). Biologically, this noise represents the effects of all kinds of biochemical fluctuations present in cells. Note that for $h_i(t) \to \pm\infty$, $\sigma_i(t+1) = \pm 1$; for $T \to \infty$, $\sigma_i(t+1) = \pm 1$ with equal probability; and for $T \to 0$, $\sigma_i(t+1) = \text{sign}\,(h_i(t))$.

The update rule from Eq 2 may be implemented in various ways. The synchronous scheme updates the state of all nodes in the system at every time step, but this is sensible only if the simulated system has a central pacemaker coordinating the activity of all nodes. A more appropriate choice for decentralized systems like gene regulatory networks is the asynchronous

scheme in which the state of a randomly chosen subset of nodes is updated at each time step. Here, we use the asynchronous scheme with update probability 0.2 for each node.

Coupling matrix. In the canonical Hopfield model, the coupling matrix is constructed to store a set of p linearly independent (i.e. distinct) Boolean patterns $\xi_i^\mu = \pm 1$ as point attractors, where $i = 0, 1, \ldots, N-1$ is the node index and $\mu = 0, 1, \ldots, p-1$ is the pattern index. The point attractor coupling matrix J'_{ij} is given by

$$J'_{ij} = \frac{1}{Np} \sum_{\mu\nu} \xi_i^\mu (Q^{-1})_{\mu\nu} \xi_j^\nu , \tag{3}$$

where [15, 71]

$$Q_{\mu\nu} = \frac{1}{N} \sum_i \xi_i^\mu \xi_i^\nu \ . \tag{4}$$

With this coupling matrix and $T = 0$, if at some time t the configuration is given by $\sigma_i(t) = \xi_i^\mu + \delta_i$ for a small perturbation δ_i, then $\lim_{t\to\infty} \sigma_i(t) = \xi_i^\mu$. Note that this formulation means $\pm\xi_i^\mu$ are both attractors of the system.

A simple modification [19] to Eq 3 produces a cyclic attractor coupling matrix $\widetilde{J}_{ij}$, constructed according to

$$\widetilde{J}_{ij} = \frac{1}{Np} \sum_{\mu\nu} \xi_i^{\mathrm{mod}_p(\mu+1)} (Q^{-1})_{\mu\nu} \xi_j^\nu . \tag{5}$$

At $T = 0$, this coupling matrix cyclically maps through the sequence of p patterns

$$\xi_i^0 \to \xi_i^1 \to \ldots \to \xi_i^{p-2} \to \xi_i^{p-1} \to \xi_i^0 \to \xi_i^1 \to \ldots \tag{6}$$

or their negatives. For the remainder of this article, all attractor indexing is understood to be modulo p.

A delayed cyclic Hopfield model may be constructed by combining the point and cyclic attractor matrices into one coupling matrix,

$$J_{ij}(\lambda) = (1-\lambda) J'_{ij} + \lambda \widetilde{J}_{ij} \ , \tag{7}$$

for an adjustable transition strength parameter λ with $0 \le \lambda \le 1$. If $\sigma(t) = \xi_i^\mu$, $\lambda \ll 1$, and $T = 0$, the point attractor term dominates and $\sigma_i(t) = \sigma(t + \Delta t)$ for all $\Delta t = 1, 2, \ldots$. If $T > 0$, however, stochastic fluctuations eventually push the configuration out of the basin of attraction of the μ^{th} attractor and into the $(\mu+1)^{\text{th}}$ basin, then eventually to the $(\mu+2)^{\text{th}}$ basin, and so on. The dynamics of the delayed cyclic Hopfield model are thus governed by noise-induced transitions.

Due to the sinusoidal nature of the gene expression in these CC data sets, however, the attractors are structured such that $\xi_i^\mu = -\xi_i^{\mu+4}$, making $Q_{\mu\nu}$ rank deficient and thus noninvertible. Because the definition of J_{ij} automatically guarantees that if any sequence $\{+\xi_i^\mu\}$ is an attractor, then $\{-\xi_i^\mu\}$ is also an attractor, encoding the sequence

$$\xi_i^0 \to \xi_i^1 \to \xi_i^2 \to \xi_i^3 \to \xi_i^4 (= -\xi_i^0) \tag{8}$$

automatically encodes the sequence

$$\xi_i^4 \to \xi_i^5 \to \xi_i^6 \to \xi_i^7 \to \xi_i^0 (= -\xi_i^4) \ . \tag{9}$$

In this special case of sinusoidal trajectories, the limits of summation in Eqs 3–5 need only run over the first four indices, $\mu = 0, 1, 2, 3$.

Finally, to reflect the fact that real gene regulatory networks are sparse, weak edges were removed by setting all elements of the coupling matrix with $|J_{ij}| < \text{median}(|J|)$ to zero, where $|J|$ is element-wise absolute value. The cyclic attractor was not preserved when using stricter thresholds due to the reduced capacity of diluted Hopfield networks [17].

Biological interpretation of the dynamics. Extracting biological meaning from this model requires defining some convenient coarse-grained quantities. The overlap of the state vector $\sigma_i(t)$ with the μ^{th} pattern is given by

$$m^{\mu}(t) = \frac{1}{N}\sum_i \sigma_i(t)\xi_i^{\mu} \ , \tag{10}$$

where $-1 \leq m^{\mu}(t) \leq +1$. The overlap measures the similarity between the (discretized) experimental and simulated gene expression profiles, and $m^{\mu}(t) = +1$ means there is perfect agreement between the simulated cell's expression and pattern μ.

A single configuration vector $\sigma_i(t)$ represents the expression profile of a single cell. For many cells κ, let $\sigma_{ik}(t)$ be the expression of gene i in cell k. Define

$$m_k^{\mu}(t) = \frac{1}{N}\sum_i \sigma_{ik}(t)\xi_i^{\mu} \tag{11}$$

as the overlap of cell k with attractor μ. Because the microarray and RNA-seq data used here report the gene expression averaged over many cells, it is appropriate to define the population-averaged (i.e. ensemble-averaged) expression,

$$\langle \sigma_i(t) \rangle_{\mathcal{K}} = \frac{1}{\kappa}\sum_{k=0}^{\kappa-1} \sigma_{ik}(t) \, , \tag{12}$$

which has $-1 \leq \langle \sigma_i(t) \rangle_{\mathcal{K}} \leq +1$.

Rather than work with a continuous vector quantity like $m_k^{\mu}(t)$, each cell can simply be identified as being in a discrete phenotypic state at any given time. Define the state of cell k as

$$s_k(t) = \arg\max_{\mu} m_k^{\mu}(t) \ , \tag{13}$$

i.e. the index of the attractor with maximum overlap, which may be interpreted as cell k's phenotype. To better understand population-level dynamics, define the discrete probability distribution $P_{\mu}(t)$ as the fraction of κ cells with $s_k(t) = \mu$; that is, $P_{\mu}(t)$ is the probability that a randomly chosen cell is in state μ at time t. Finally, define the time-averaged distribution of states as

$$\langle P_{\mu} \rangle_{\mathcal{T}} = \frac{1}{\tau}\sum_{t=0}^{\tau-1} P_{\mu}(t) \tag{14}$$

for a window of time τ.

For each data set, T, λ, and the single-node update probability were tuned to the "edge of chaos" [72] such that the cyclic attractor was preserved and the time between transitions was approximately constant, but the system was sensitive enough to perturbations that some targeted inhibitions produced noticeable changes in $\langle P_{\mu} \rangle_{\mathcal{T}}$. See S1 Table for a list of parameters used for each data set.

Gene inhibition optimization. In this application, the goal is to identify perturbations that halt or retard the encoded cyclic attractor. A standard genetic algorithm (GA; explained in S1 Text) was employed to identify an optimal control field h_i^{opt} that maximized a given objective function $f(h_i^{\text{ext}})$,

$$h_i^{\text{opt}} = \arg\max_{h_i^{\text{ext}}} f(h_i^{\text{ext}}) \ , \tag{15}$$

where h_i^{ext} is the control vector given by

$$h_i^{\text{ext}} = \begin{cases} -\infty & \text{if gene } i \text{ is targeted} \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

for a fixed number of targets (nonzero elements) n_{targ}. Only negative control fields are used here to simulate the effects of targeted gene inhibition from pharmaceutical compounds. The objective function used here is $\langle P_\mu \rangle_{\mathcal{T}}$, meaning that the optimal control field maximizes the time-averaged number of cells occupying a particular attractor state μ. This search was conducted across all attractors μ to determine the controllability of each attractor state. Note that while the numerical value of $\langle P_\mu \rangle_{\mathcal{T}}$ depends on the choice of parameters such as T, λ, and the single-node update probability, varying these parameters by ±5% gave comparable results for the optimal sets of controllers.

Dynamical behavior

Fig 2 shows the time evolution of $s_k(t)$ for a single simulated cell using the attractors derived from [64]. As expected, the system progresses cyclically through the eight attractor states. The duration of each cycle varies somewhat due to the stochasticity in the update rule from Eq 2.

Although the gene expression for each simulated cell k has $\sigma_{ik}(t) = \pm 1$, the population-averaged expression has $-1 \leq \langle \sigma_i(t) \rangle_{\mathcal{K}} \leq +1$, and for many cells initially synchronized with $\sigma_{ik}(0) = \xi_i^0$ for all k, $\langle \sigma_i(t) \rangle_{\mathcal{K}}$ successfully recovers the experimentally observed decaying sinusoidal gene expression. Fig 3 shows a comparison between the experimental expression $x_i(t)$ from the Eser data set and the mean simulated expression $\langle \sigma_i(t) \rangle_{\mathcal{K}}$ with $\kappa = 50$ for $i = SLD2$,

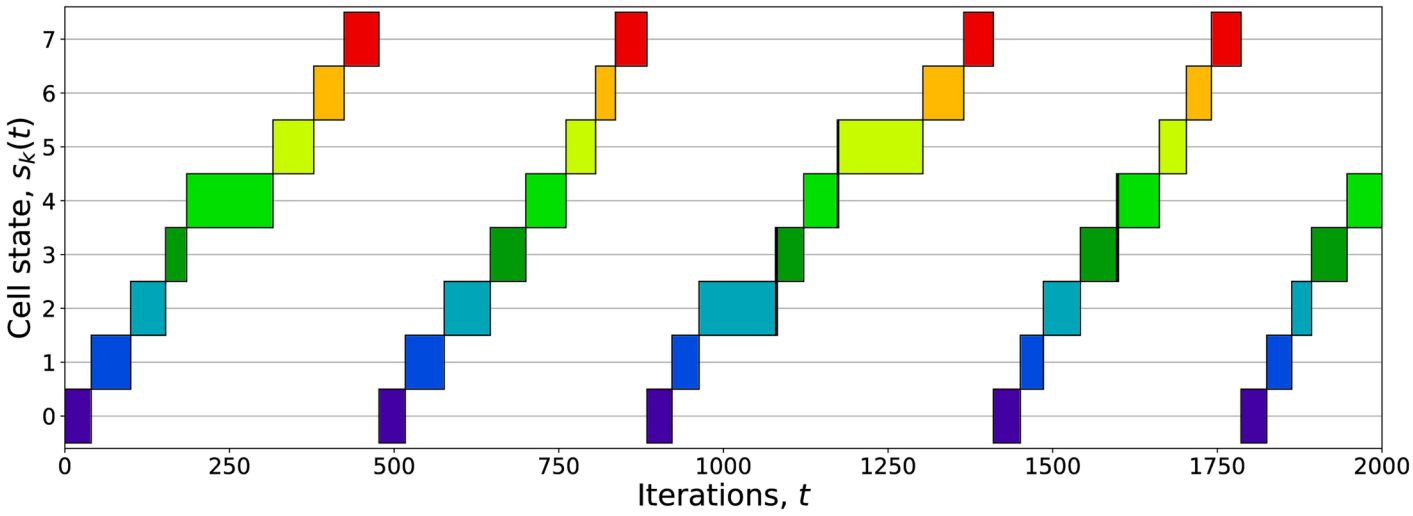

Fig 2. Unperturbed cell state versus time. Boxes indicate $s_k(t)$, i.e. the index of the attractor with maximum overlap at time t. The system began with the configuration $\sigma_i(0) = \xi_i^0$ and was allowed to evolve according to the Hopfield signaling rules with zero external field, mapping cyclically through the set of eight attractors. The pattern and cycle durations vary due to the system's stochasticity.

https://doi.org/10.1371/journal.pcbi.1005849.g002

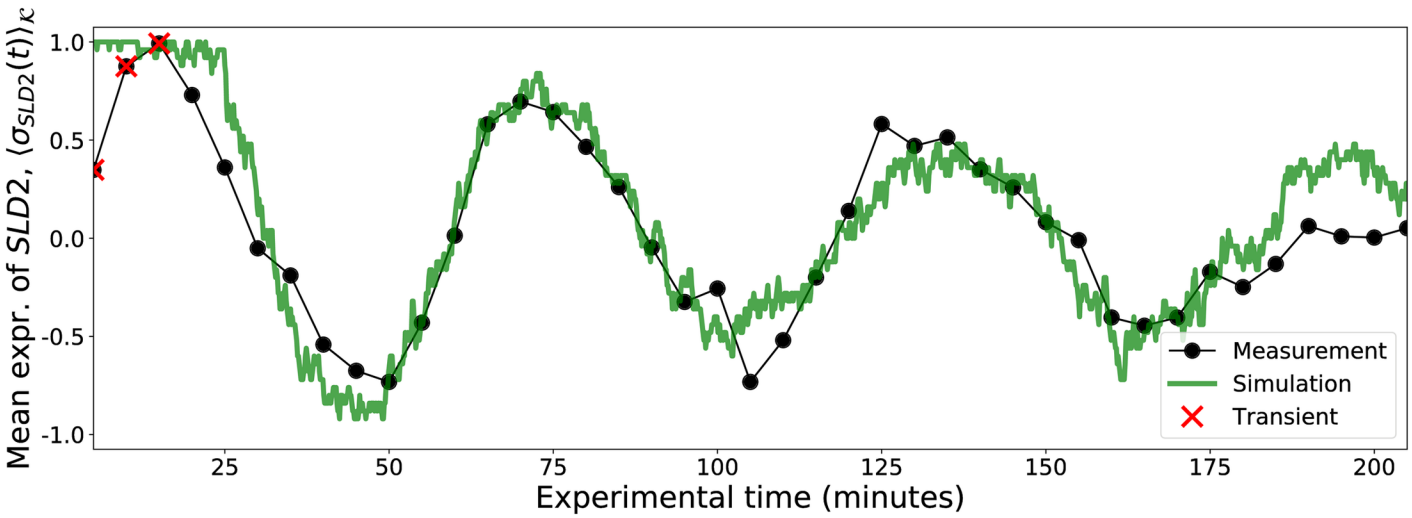

Fig 3. Measured and simulated gene expression. The measured expression of the *S. cerevisiae* gene *SLD2* from [64] was scaled to the range [−1, +1] and is shown as a black beaded curve, and the population-averaged expression of the same gene as defined in Eq 12 for $\kappa = 50$ cells is shown in green, with the t axis rescaled by eye to match experimental time. Transient points (red X's) were ignored when fitting Eq 17.

https://doi.org/10.1371/journal.pcbi.1005849.g003

one of the genes responsible for initiating DNA replication in *S. cerevisiae* [73, 74]. The simulation time t was rescaled by eye to align the simulated and experimental curves (so that 1 minute corresponds to approximately 7.5 iterations).

Trajectories can be visualized by projecting them onto the first two principal components (PCs) of the attractor configurations. Fig 4 shows the eight attractors as stars, and a single cell

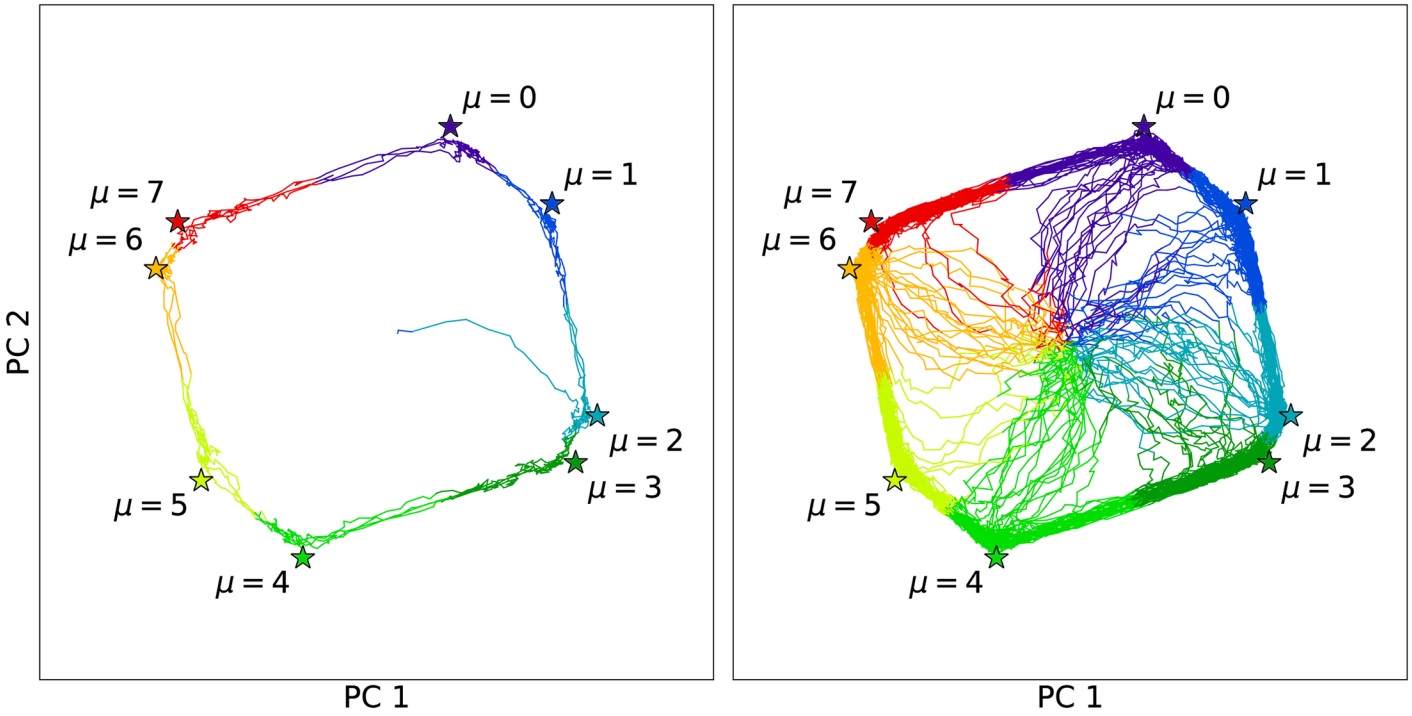

Fig 4. Principal component projection of unperturbed cell trajectories. The simulated single cell (left panel) and 100 cells (right panel) began with random initial states (projected near the center of the plot), but quickly settled into the encoded cycle. Line segments were colored according to $s_k(t)$, i.e. which of the eight attractors (labeled stars) had maximum overlap at time t.

https://doi.org/10.1371/journal.pcbi.1005849.g004

trajectory (left panel) and 100 cell trajectories (right panel) with random initial states as curves with line segments colored according to $s_k(t)$ (as computed in the full N-dimensional space). Although the cells begin nearly equidistant from all ξ_i^μ, they quickly relax into encoded cycle. S1 Video shows an animation of a system of $\kappa = 50$ cells with random initial conditions projected onto the same PCs, where cells (circles) are colored according to $s_k(t)$. As with the cells shown in Fig 4, all initially random configurations eventually converge to the cycle. S2 Video shows an animation of $\kappa = 50$ cell trajectories with $\sigma_{ik}(0) = \xi_i^0$. As time progresses, the phases of the initially synchronized cells slowly decohere because cells stochastically and independently transition between attractors due to the finite temperature in Eq 2.

S3 Video demonstrates the effect of temperature on the dynamics. 50 cells were given random initial states, and the temperature was increased and decreased in steps. Cells rarely escape the eight attractor states for $T = 0.045$, and one cell becomes stuck near the center in a spurious attractor (unintentional metastable states that arise from the model's nonlinearity). At $T = 0.06$, fluctuations allow the cells to transition somewhat regularly through the encoded cycle, and the cell trapped in the spurious attractor eventually escapes and joins the cycle. At $T = 0.09$ the cells begin to noticeably diverge from the eight attractor states, but still collectively display a net clockwise flow. The noise is too great for the cells to follow the cycle for $T = 0.15$, but lowering the temperature again returns cells to the cycle. This illustrates the fact that the cycle is preserved only for intermediate temperatures: cells become "frozen" in intended or spurious attractor states at low temperatures, but at high temperatures the noise is too great and the couplings between genes become irrelevant to the dynamics.

Optimal control fields

The GA was used to identify some effective combinations of gene targets that slowed progress through the cyclic attractor for varying numbers of targets, n_{targ}. The GA found, for example, that inhibiting the set of eight *S. cerevisiae* genes *HEK2*, *PRR1*, *QRI1*, *RFC4*, *STB1*, *TDA7*, *VPS17*, and *ZIM17* was sufficient to trap $\sim$95% of cells in the $\mu = 7$ state. The effects of this control field on the time evolution of $P_\mu(t)$ for $\kappa = 50$ and $\kappa = 5000$ are shown in Fig 5. Cells were given random, independent initial states at $t = -200$ (not shown), quickly settling into the cyclic attractor with evenly distributed phases so that $P_\mu(0 \leq t < 200) \approx 1/8$. The control field was activated at $t = 200$, causing the cells to accumulate in the $\mu = 7$ state. The field was then disabled at $t = 1000$, allowing the cells to resume cycling with initially synchronized phases, as shown by the sequence of oscillations in $P_\mu(t > 1000)$. The stochastic nature of the transitions causes the cells' phases to slowly spread so that $P_\mu(t \to \infty) \approx 1/8$, eventually returning the system to a desynchronized state.

The effects of this control field can also be visualized using a PC projection as shown in Fig 6 and S4 Video. The same set of $\kappa = 50$ trajectories from Fig 5 was projected onto the attractors' PCs, with cells colored according to $s_k(t)$. The control field manages to fix most cells near the $\mu = 7$ state, but as shown in the $t = 910$ panel in Fig 6, fluctuations occasionally push individual cells out of the $\mu = 7$ basin and back into the cycle.

Further searches were constrained to inhibiting between 1 and 38 kinases in HeLa cells and between 1 and 12 kinases in *S. cerevisiae*. For nearly all numbers of targets, the GA found $\mu = 2$ (M phase) to be the most controllable attractor in HeLa cells and $\mu = 7$ (S phase) in *S. cerevisiae*, as shown in S1 and S2 Figs, respectively. In HeLa cells, for example, an optimal score of $\langle P_2 \rangle_{\mathscr{T}} \approx 0.230$ was obtained for $n_{\text{targ}} = 4$ (0.55% of the 722 periodic genes) when inhibiting the genes *AURKB*, *NEK1*, *TTK* (all involved in the formation and operation of the spindle apparatus), and *WEE1* (a G2/M checkpoint regulator involved with DNA damage repair). The effects of inhibiting *AURKB* and *TTK* are well documented. *AURKB* is targetable by the kinase

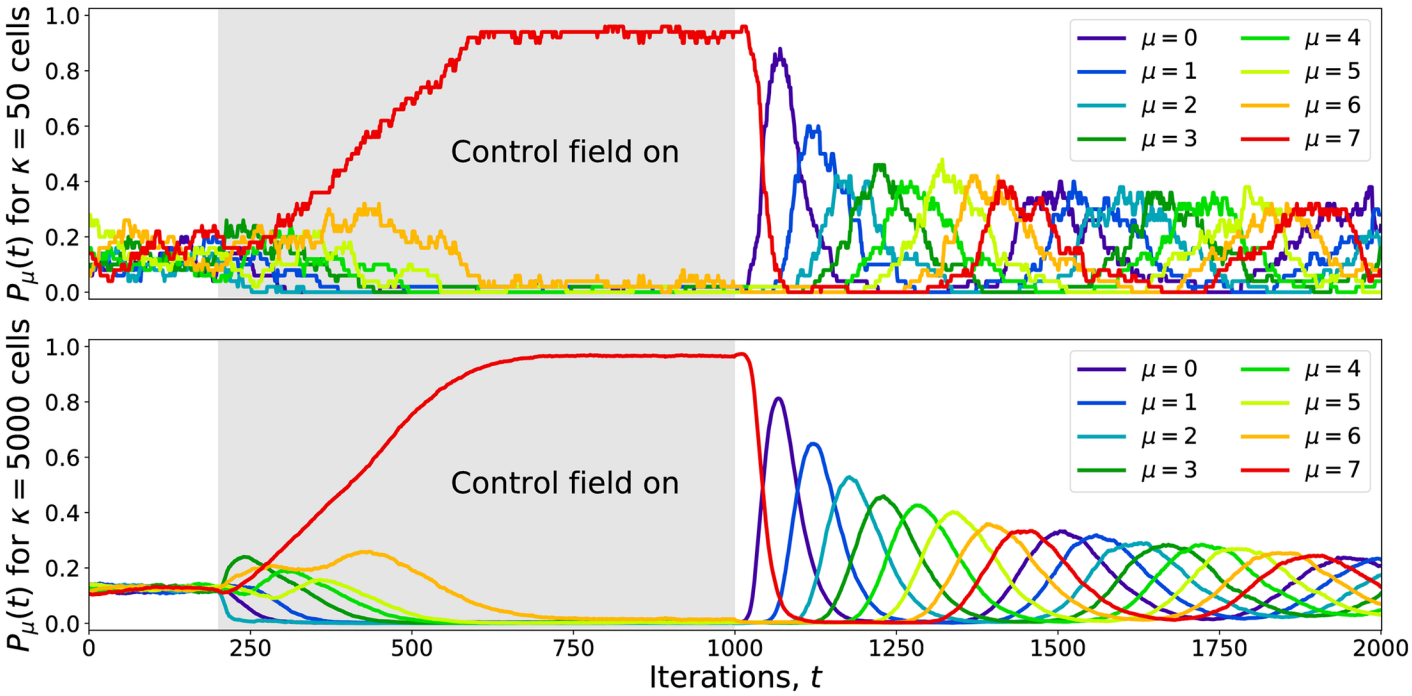

Fig 5. Cell state synchronization by targeted inhibition for 50 and 5000 cells. Cells were initialized with random states at $t = -200$ (not shown) and allowed to relax into the cyclic attractor so that $P_\mu(0 \leq t < 200) \approx 1/8$. A set of eight genes was inhibited with an external control field over the range $200 \leq t < 1000$, fixing most cells near the $\mu = 7$ state. After removing the control field, the cells resumed moving through the cycle with initially synchronized phases that slowly broaden. Eventually the system returns to a desynchronized state, $P_\mu(t \to \infty) \approx 1/8$.

https://doi.org/10.1371/journal.pcbi.1005849.g005

inhibitor barasertib, found to be effective in combating small-cell lung cancer *in vitro* and in xenografts [75]. *TTK* (also known as *MPS1*) is targetable by the kinase inhibitors Mps1-IN-1 and AZ3146. Mps1-IN-1 has been found to be an effective tool to inhibit proliferation in an array of cancer cell lines, including HeLa cells [76]. AZ3146 has been effective against hepatocellular carcinoma (liver cancer) *in vitro* [77] and pancreatic ductal adenocarcinoma *in vitro* [78]. siRNAs and shRNAs have also been used to inhibit TTK in pancreatic ductal adenocarcinoma [78].

Discussion

Above we presented a delayed cyclic Hopfield model designed to store CC time series gene expression data from synchronized *S. cerevisiae* and HeLa cells, and the behaviors of both individual cells and populations of cells were studied. The dynamics of populations of cells successfully recreated the experimental gene expression data, including the slow decoherence of initially synchronized cells due to the stochastic transitions between attractors. Optimal control fields that freeze or stall the cyclic attractor by inhibiting only a small number of genes were identified. These predictions could be experimentally tested using kinase inhibitors, micro RNAs, short interfering RNAs, or knockout studies and confirmed simply by counting the number of cells as a function of time compared to a control with no inhibitors.

Admittedly, there are several limitations to this model. Additionally, although using the temporal ordering of the time series gene expression samples provides more information about potentially causal relationships than static samples, the Hopfield model is ultimately an effective model that builds gene-gene couplings from pairwise correlations in gene expression, thereby capturing direct, indirect, and spurious relationships between genes. Independently

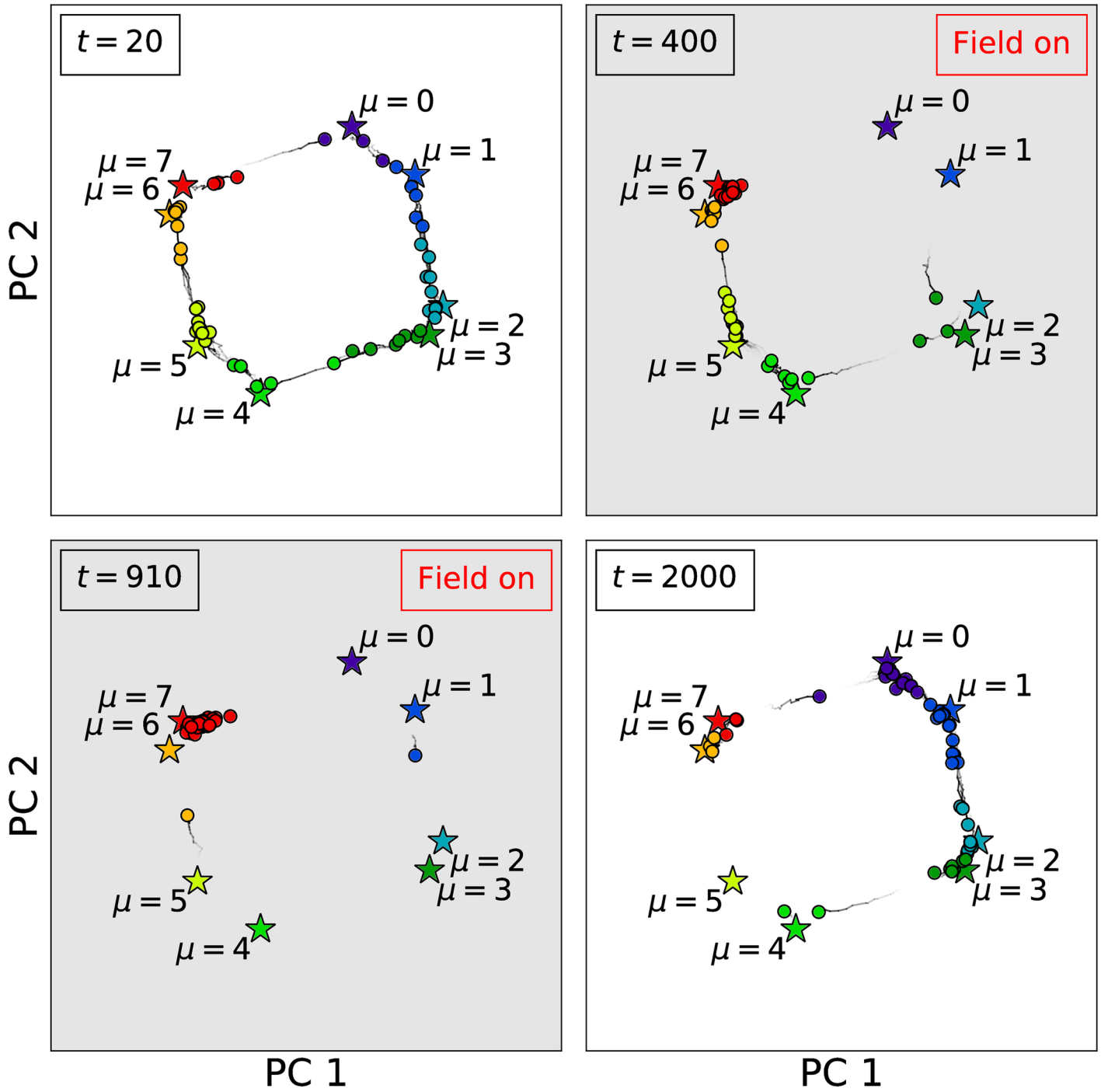

Fig 6. Principal component projection of 50 cell trajectories. The trajectories used to make the $\kappa = 50$ panel of Fig 5 were projected onto the the first two principal components (PCs) of the attractor array ξ_i^μ (labeled stars). Cells (circles) are colored according to the closest attractor as computed by Eq 13. When the external field is activated, most cells become trapped in the $\mu = 7$ state, although occasionally cells break from the group and complete another circuit before becoming trapped again. After the external field is removed, the cells eventually return to a desynchronized state. See S4 Video for an animation of these trajectories.

https://doi.org/10.1371/journal.pcbi.1005849.g006

derived network information with experimentally confirmed molecular regulatory interactions could perhaps be used to refine the construction of the coupling matrix. Other data sets that explore a more diverse set of expression states other than a simple cycle could also reduce the number of spurious couplings. We are currently developing these kinds of non-cyclic models and time series data sets in a separate project.

Our approach can be generalized and improved in many ways. This incarnation of the model causes simulated cells to continuously undergo CC with no G0 (resting) phase. Adding a relatively stable G0 attractor between the M and G1 phases could cause cells to pause between cycles. A GA search could then be conducted to find the best sets of inhibition targets to freeze cells in the G0 state, or to find the best sets of targets to stimulate entry into CC, mimicking the effects of environmental signals such as growth factors.

We chose to discretize the continuous gene expression data using a traditional two-state model, which assumes that each gene is either fully activated or fully deactivated. Using a multi-level Hopfield model [79] could better reflect the continuous nature of gene expression data and potentially improve the search results. This model can also incorporate additional omics information, e.g. proteomics and metabolomics, simply by increasing the number of nodes in the system. We plan to explore this option as more multi-omics time series data sets become available. Single-cell experimental techniques and analytical tools are also rapidly improving in quality, decreasing in cost, and gaining in popularity [80–82], and using techniques like pseudo-temporal ordering [83] could allow the Hopfield model to encode single-cell RNA-seq data as well.

Although the above simulated populations of cells exhibit intriguing dynamical and statistical properties, they behave as completely homogeneous, non-interacting particles. The importance of cell-cell communication and interactions in populations of cells has been demonstrated in a variety of systems including bacterial quorum sensing [84] and community spatial patterning [85], neuron synchronization in circadian rhythm [86], and various forms of cancer [87–91]. As with many nonlinear systems, even seemingly minor changes can produce dramatically different outcomes. More complex extensions to our model could incorporate cell-cell communication by, for example, adding couplings between known signaling molecules and receptors between different cells, and could even allow for interactions between heterogeneous cell types. This would increase the computational complexity of the model, but could better reflect the underlying biology.

Methods

Gene expression fitting

In order to encode these CC data sets into the Hopfield model, periodic genes needed to be identified, their frequencies and phases computed, and their expression converted from continuous to Boolean form. SciPy's Trust Region Reflective method [92] was used to identify genes i with periodic expression $x_i(t)$ by fitting to the form

$$x_i(t) = a_i e^{-b_i t} \cos(\omega_i t - \phi_i) + x_{i0} \tag{17}$$

for amplitude a_i, decay rate b_i, angular frequency ω_i, phase ϕ_i, and asymptotic mean expression x_{i0}. The first several time points were ignored to avoid fitting the parameters of Eq 17 to chemically perturbed (transient) states. A gene was labeled periodic if the maximum relative uncertainty in its parameters from the fit,

$$r_i^{\max} = \max\left\{\frac{\delta x_{i0}}{x_{i0}}, \frac{\delta a_i}{a_i}, \frac{\delta b_i}{b_i}, \frac{\delta \omega_i}{\omega_i}, \frac{\delta \phi_i}{2\pi}\right\}, \tag{18}$$

was less than the thresholds defined in S1 Table. (Note that because the HeLa data covers only a little more than one period, many of the b_i's were approximately zero. This resulted in extremely large $r_i^{\max}$ despite the remaining variables having small uncertainty. The $\delta b_i/b_i$ term was thus ignored for the HeLa data.) Once all frequencies $\{\omega_i\}$ for periodic genes were computed, the frequency was fixed to the mean frequency $\bar{\omega}$ and the fits were recomputed for each periodic gene using the form

$$x_i(t) = a_i e^{-b_i t} \cos(\bar{\omega} t - \phi_i) + x_{i0} \ , \tag{19}$$

thus producing the final set of continuous phases $\{\phi_i\}$. Fig 1A shows a heat map of the expression of all periodic genes detected in the Eser data set sorted by their fitted phases, and Fig 1B shows the same genes with the fitted expression curves. These fitted curves were converted from continuous values $x_i(t) \geq 0$ to Boolean values $\xi_i(t) = \pm 1$ (over/underexpressed) by assigning

$$\xi_i(t) = \ \text{sign}\,(x_i(t) - x_{i0}) \tag{20}$$

as shown in Fig 1C. Finally, one CC period was divided into eight uniformly spaced states $\{\xi_i^\mu\} = \{\xi_i^0, \xi_i^1, \ldots, \xi_i^7\}$. These states, shown in Fig 1D, were used as attractors in the Hopfield model.

Determining cell cycle phase

The approximate CC phase for each attractor μ in the HeLa data set was determined using a set of marker genes [65, 93] and over-representation analysis using a hypergeometric distribution to calculate p-values with the Benjamini-Hochberg procedure [94] to correct for multiple hypothesis testing. All genes i with $\xi_i^{\mu-1} = -1$ and $\xi_i^\mu = +1$ were used as the μ^{th} input set, and all detected cyclic genes were used as the background. $\mu = 2$ showed the activation of the G2/M checkpoint marker *CCNA2* and was enriched for the gene ontology (GO) term cell division ($p = 3.1 \times 10^{-2}$). $\mu = 3$ showed the activation of the G2, G2/M, and M markers *BUB1*, *CCNB1*, *NEK2*, and *PLK1*, and was enriched for signaling by Rho GTPases, cell division, midbody, mitotic prometaphase, mitotic nuclear division, and small GTPase mediated signal transduction ($p = 3.7 \times 10^{-3}$); M phase ($p = 6.2 \times 10^{-3}$); resolution of sister chromatid cohesion ($p = 7.6 \times 10^{-3}$); mitotic anaphase, spindle pole, and RHO GTPase effectors ($p = 3.0 \times 10^{-2}$); and mitotic metaphase and anaphase ($p = 3.8 \times 10^{-2}$). $\mu = 6$ expressed the late G1, G1/S, and S markers *CCNE1*, *CDC6*, *E2F1*, *MCM2*, and *SLBP*, and was enriched for DNA replication initiation ($p = 5.9 \times 10^{-4}$); G1/S-specific transcription, G1/S transition, DNA replication pre-initiation, M/G1 transition, and assembly of the pre-replicative complex ($p = 4.3 \times 10^{-2}$); and G1/S transition of mitotic cell cycle ($p = 4.4 \times 10^{-2}$). $\mu = 7$ showed the activation of the G1/S and S markers *CDK2* and *CDKN2AIP*. The database yeastgenome.org [95] was used to determine the CC phases for the Eser *S. cerevisiae* data set. $\mu = 0$ was enriched for the GO term DNA replication ($p = 2.02 \times 10^{-12}$), indicating an attractor in the S phase of CC. $\mu = 2$ was enriched for mitotic spindle organization ($p = 2.3 \times 10^{-3}$) indicating the beginning of mitosis in *S. cerevisiae*. $\mu = 6$ from Eser was enriched for the GO term pre-replicative complex assembly involved in nuclear cell cycle DNA replication ($p = 3.0 \times 10^{-5}$), indicating an attractor at the end of G1 phase as the cells prepare for DNA replication.

Supporting information

S1 Video. 50 cell trajectories with random initial conditions. Data was projected onto the first two principle components of the attractor array ξ_i^μ. Attractors are shown as stars, and cells

are shown as circles. Cell colors are assigned using $s_k(t)$ as measured in the full N-dimensional space. All cells k were given random initial conditions $\sigma_{ik} = \pm 1$ with equal probability for all i and k, but eventually converge to the cyclic attractor.
(AVI)

S2 Video. 50 cell trajectories with identical initial conditions. See the caption of S1 Video for an explanation of the projection and colors. All cells k were initially synchronized with $\sigma_{ik}(t) = \xi_i^0$, but progress through the cycle stochastically, causing the distribution of $s_k(t)$ to broaden.
(AVI)

S3 Video. Effects of temperature on 50 cell trajectories. See the caption of S1 Video for an explanation of the projection and colors. All cells were given initial random states, and the temperature was increased and decreased in steps as shown in the top panel.
(AVI)

S4 Video. Principal component projection of 50 cells being synchronized. See the caption of S1 Video for an explanation of the projection and colors. This video is an animation of the trajectories used in Figs 5 and 6.
(AVI)

S1 Text. Explanation of genetic algorithm.
(PDF)

S1 Table. List of parameters used for each data set.
(XLSX)

S2 Table. Enrichment results for each attractor from the Eser data set.
(XLSX)

S3 Table. Enrichment results for each attractor from the Dominguez data set.
(XLSX)

S1 Fig. GA's optimal control results versus the number of inhibited kinases for HeLa cells. Each point indicates the best measured score (mean fraction of cells) for the given attractor and number of targets. $\mu = 2$ (M phase) is generally the most controllable attractor in HeLa cells.
(PNG)

S2 Fig. GA's optimal control results versus the number of inhibited kinases for *S. cerevisiae*. Each point indicates the best measured score (mean fraction of cells) for the given attractor and number of targets. $\mu = 7$ (S phase) is generally the most controllable attractor in *S. cerevisiae*.
(PNG)

S1 File. Raw yeast time series gene expression data. Taken from [64].
(CSV)

S2 File. Raw HeLa time series gene expression data. Taken from [65].
(XLSX)

Acknowledgments

We would like to thank George Mias (Michigan State University) and Yunyi Kang (Sanford Burnham Prebys Medical Discovery Institute) for helpful discussions and recommendations.

Author Contributions

Conceptualization: Anthony Szedlak, Giovanni Paternostro, Carlo Piermarocchi.

Formal analysis: Anthony Szedlak, Spencer Sims, Nicholas Smith.

Funding acquisition: Giovanni Paternostro, Carlo Piermarocchi.

Investigation: Anthony Szedlak.

Project administration: Giovanni Paternostro, Carlo Piermarocchi.

Supervision: Giovanni Paternostro, Carlo Piermarocchi.

Writing – original draft: Anthony Szedlak.

References

1. Waddington CH (2014) The Strategy of the Genes, volume 20. Routledge.
2. Kauffman S (1971) Differentiation of malignant to benign cells. Journal of Theoretical Biology 31: 429–451. https://doi.org/10.1016/0022-5193(71)90020-8 PMID: 5556142
3. Huang S, Eichler G, Bar-Yam Y, Ingber DE (2005) Cell fates as high-dimensional attractor states of a complex gene regulatory network. Physical Review Letters 94: 128701. https://doi.org/10.1103/PhysRevLett.94.128701 PMID: 15903968
4. Alvarez-Buylla ER, Chaos Á, Aldana M, Benítez M, Cortes-Poza Y, et al. (2008) Floral morphogenesis: stochastic explorations of a gene network epigenetic landscape. PLOS ONE 3: e3626. https://doi.org/10.1371/journal.pone.0003626 PMID: 18978941
5. Huang S, Ernberg I, Kauffman S (2009) Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. In: Seminars in Cell & Developmental Biology. Elsevier, volume 20, pp. 869–876.
6. Demicheli R, Coradini D (2010) Gene regulatory networks: a new conceptual framework to analyse breast cancer behaviour. Annals of Oncology 22: 1259–1265. https://doi.org/10.1093/annonc/mdq546 PMID: 21109571
7. Choi M, Shi J, Jung SH, Chen X, Cho KH (2012) Attractor landscape analysis reveals feedback loops in the p53 network that control the cellular response to DNA damage. Sci Signal 5: ra83–ra83.
8. Yuan R, Zhu X, Radich JP, Ao P (2016) From molecular interaction to acute promyelocytic leukemia: Calculating leukemogenesis and remission from endogenous molecular-cellular network. Scientific Reports 6.
9. Pusuluri ST (2017) Exploring Neural Network Models with Hierarchical Memories and Their Use in Modeling Biological Systems. Ph.D. thesis, Ohio University.
10. Udyavar AR, Wooten DJ, Hoeksema M, Bansal M, Califano A, et al. (2017) Novel hybrid phenotype revealed in small cell lung cancer by a transcription factor network model that can explain tumor heterogeneity. Cancer Research 77: 1063–1074. https://doi.org/10.1158/0008-5472.CAN-16-1467 PMID: 27932399
11. Wooten DJ, Quaranta V (2017) Mathematical models of cell phenotype regulation and reprogramming: Make cancer cells sensitive again! Biochimica et Biophysica Acta (BBA)-Reviews on Cancer 1867: 167–175. https://doi.org/10.1016/j.bbcan.2017.04.001
12. Yuan R, Zhu X, Wang G, Li S, Ao P (2017) Cancer as robust intrinsic state shaped by evolution: a key issues review. Reports on Progress in Physics 80: 042701. https://doi.org/10.1088/1361-6633/aa538e PMID: 28212112
13. Hopfield JJ (1982) Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences 79: 2554–2558. https://doi.org/10.1073/pnas.79.8.2554
14. Peretto P (1988) On learning rules and memory storage abilities of asymmetrical neural networks. Journal de Physique 49: 711–726. https://doi.org/10.1051/jphys:01988004905071100
15. Kanter I, Sompolinsky H (1987) Associative recall of memory without errors. Physical Review A 35: 380. https://doi.org/10.1103/PhysRevA.35.380
16. Amit DJ, Gutfreund H, Sompolinsky H (1985) Storing infinite numbers of patterns in a spin-glass model of neural networks. Physical Review Letters 55: 1530. https://doi.org/10.1103/PhysRevLett.55.1530 PMID: 10031847

17. Derrida B, Gardner E, Zippelius A (1987) An exactly solvable asymmetric neural network model. Europhysics Letters 4: 167. https://doi.org/10.1209/0295-5075/4/2/007
18. Nishimori H, Nakamura T, Shiino M (1990) Retrieval of spatio-temporal sequence in asynchronous neural network. Physical Review A 41: 3346. https://doi.org/10.1103/PhysRevA.41.3346
19. Personnaz L, Guyon I, Dreyfus G (1986) Collective computational properties of neural networks: New learning mechanisms. Physical Review A 34: 4217. https://doi.org/10.1103/PhysRevA.34.4217
20. Nasrabadi NM, Li W (1991) Object recognition by a Hopfield neural network. IEEE Transactions on Systems, Man, and Cybernetics 21: 1523–1535. https://doi.org/10.1109/21.135694
21. Cote S, Tatnall A (1997) The Hopfield neural network as a tool for feature tracking and recognition from satellite sensor images. International Journal of Remote Sensing 18: 871–885. https://doi.org/10.1080/014311697218809
22. Maetschke SR, Ragan MA (2014) Characterizing cancer subtypes as attractors of Hopfield networks. Bioinformatics 30: 1273–1279. https://doi.org/10.1093/bioinformatics/btt773 PMID: 24407221
23. Lang AH, Li H, Collins JJ, Mehta P (2014) Epigenetic landscapes explain partially reprogrammed cells and identify key reprogramming genes. PLOS Comput Biol 10: e1003734. https://doi.org/10.1371/journal.pcbi.1003734 PMID: 25122086
24. Jang S, Choubey S, Furchtgott L, Zou LN, Doyle A, et al. (2017) Dynamics of embryonic stem cell differentiation inferred from single-cell transcriptomics show a series of transitions through discrete cell states. eLife 6: e20487. https://doi.org/10.7554/eLife.20487 PMID: 28296635
25. Szedlak A, Paternostro G, Piermarocchi C (2014) Control of asymmetric Hopfield networks and application to cancer attractors. PLOS ONE 9: e105842. https://doi.org/10.1371/journal.pone.0105842 PMID: 25170874
26. Hannam R, Annibale A, Kuehn R (2017) Cell reprogramming modelled as transitions in a hierarchy of cell cycles. Journal of Physics A: Mathematical and Theoretical.
27. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics 10: 57–63. https://doi.org/10.1038/nrg2484 PMID: 19015660
28. Marguerat S, Bähler J (2010) RNA-seq: from technology to biology. Cellular and Molecular Life Sciences 67: 569–579. https://doi.org/10.1007/s00018-009-0180-6 PMID: 19859660
29. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols 7: 562–578. https://doi.org/10.1038/nprot.2012.016 PMID: 22383036
30. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, et al. (2006) ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics 7: S7.
31. Hecker M, Lambeck S, Toepfer S, Van Someren E, Guthke R (2009) Gene regulatory network inference: data integration in dynamic models—a review. Biosystems 96: 86–103. https://doi.org/10.1016/j.biosystems.2008.12.004 PMID: 19150482
32. Lee I, Blom UM, Wang PI, Shim JE, Marcotte EM (2011) Prioritizing candidate disease genes by network-based boosting of genome-wide association data. Genome Research 21: 1109–1121. https://doi.org/10.1101/gr.118992.110 PMID: 21536720
33. Marbach D, Costello JC, Küffner R, Vega NM, Prill RJ, et al. (2012) Wisdom of crowds for robust gene network inference. Nature Methods 9: 796–804. https://doi.org/10.1038/nmeth.2016 PMID: 22796662
34. Cahan P, Li H, Morris S, LummertzdaRocha E, Daley G, et al. (2014) CellNet: network biology applied to stem cell engineering. Cell 158: 903–915. https://doi.org/10.1016/j.cell.2014.07.020 PMID: 25126793
35. Ong E, Szedlak A, Kang Y, Smith P, Smith N, et al. (2015) A scalable method for molecular network reconstruction identifies properties of targets and mutations in acute myeloid leukemia. Journal of Computational Biology 22: 266–288. https://doi.org/10.1089/cmb.2014.0297 PMID: 25844667
36. Szedlak A, Smith N, Liu L, Paternostro G, Piermarocchi C (2016) Evolutionary and topological properties of genes and community structures in human gene regulatory networks. PLOS Computational Biology 12: e1005009. https://doi.org/10.1371/journal.pcbi.1005009 PMID: 27359334
37. Chen R, Mias GI, Li-Pook-Than J, Jiang L, Lam HY, et al. (2012) Personal omics profiling reveals dynamic molecular and medical phenotypes. Cell 148: 1293–1307. https://doi.org/10.1016/j.cell.2012.02.009 PMID: 22424236
38. Takahashi M (1966) Theoretical basis for cell cycle analysis i. labelled mitosis wave method. Journal of Theoretical Biology 13: 202–211. https://doi.org/10.1016/0022-5193(66)90017-8
39. Takahashi M (1968) Theoretical basis for cell cycle analysis: Ii. further studies on labelled mitosis wave method. Journal of theoretical biology 18: 195–209. PMID: 5647131

40. Lloyd D, Lloyd AL, Olsen LF (1992) The cell division cycle: a physiologically plausible dynamic model can exhibit chaotic solutions. Biosystems 27: 17–24. https://doi.org/10.1016/0303-2647(92)90043-X PMID: 1391688

41. Panetta JC, Adam J (1995) A mathematical model of cycle-specific chemotherapy. Mathematical and computer modelling 22: 67–82. https://doi.org/10.1016/0895-7177(95)00112-F

42. Swierniak A, Polanski A, Kimmel M (1996) Optimal control problems arising in cell-cycle-specific cancer chemotherapy. Cell proliferation 29: 117–139. https://doi.org/10.1046/j.1365-2184.1996.00995.x PMID: 8652742

43. Chiorino G, Metz J, Tomasoni D, Ubezio P (2001) Desynchronization rate in cell populations: mathematical modeling and experimental data. Journal of Theoretical Biology 208: 185–199. https://doi.org/10.1006/jtbi.2000.2213 PMID: 11162063

44. Tyson JJ, Novak B (2001) Regulation of the eukaryotic cell cycle: molecular antagonism, hysteresis, and irreversible transitions. Journal of theoretical biology 210: 249–263. https://doi.org/10.1006/jtbi.2001.2293 PMID: 11371178

45. Basse B, Baguley BC, Marshall ES, Joseph WR, van Brunt B, et al. (2003) A mathematical model for analysis of the cell cycle in cell lines derived from human tumors. Journal of mathematical biology 47: 295–312. https://doi.org/10.1007/s00285-003-0203-0 PMID: 14523574

46. Alarcon T, Byrne H, Maini P (2004) A mathematical model of the effects of hypoxia on the cell-cycle of normal and cancer cells. Journal of theoretical biology 229: 395–411. https://doi.org/10.1016/j.jtbi.2004.04.016 PMID: 15234206

47. Chen KC, Calzone L, Csikasz-Nagy A, Cross FR, Novak B, et al. (2004) Integrative analysis of cell cycle control in budding yeast. Molecular biology of the cell 15: 3841–3862. https://doi.org/10.1091/mbc.E03-11-0794 PMID: 15169868

48. Li F, Long T, Lu Y, Ouyang Q, Tang C (2004) The yeast cell-cycle network is robustly designed. Proceedings of the National Academy of Sciences of the United States of America 101: 4781–4786. https://doi.org/10.1073/pnas.0305937101 PMID: 15037758

49. Steuer R (2004) Effects of stochasticity in models of the cell cycle: from quantized cycle times to noise-induced oscillations. Journal of theoretical biology 228: 293–301. https://doi.org/10.1016/j.jtbi.2004.01.012 PMID: 15135028

50. Fauré A, Naldi A, Chaouiya C, Thieffry D (2006) Dynamical analysis of a generic boolean model for the control of the mammalian cell cycle. Bioinformatics 22: e124–e131.

51. Ferrell JE, Pomerening JR, Kim SY, Trunnell NB, Xiong W, et al. (2009) Simple, realistic models of complex biological processes: positive feedback and bistability in a cell fate switch and a cell cycle oscillator. FEBS letters 583: 3999–4005. https://doi.org/10.1016/j.febslet.2009.10.068 PMID: 19878681

52. Kar S, Baumann WT, Paul MR, Tyson JJ (2009) Exploring the roles of noise in the eukaryotic cell cycle. Proceedings of the National Academy of Sciences 106: 6471–6476. https://doi.org/10.1073/pnas.0810034106

53. Barik D, Baumann WT, Paul MR, Novak B, Tyson JJ (2010) A model of yeast cell-cycle regulation based on multisite phosphorylation. Molecular Systems Biology 6: 405. https://doi.org/10.1038/msb.2010.55 PMID: 20739927

54. Ferrell JE, Tsai TYC, Yang Q (2011) Modeling the cell cycle: why do certain circuits oscillate? Cell 144: 874–885. https://doi.org/10.1016/j.cell.2011.03.006 PMID: 21414480

55. Luo X, Xu L, Wang J (2017) Funneled potential and flux landscapes dictate the stabilities of both the states and the flow: Fission yeast cell cycle. Advanced online publication.

56. Assaf M, Roberts E, Luthey-Schulten Z, Goldenfeld N (2013) Extrinsic noise driven phenotype switching in a self-regulating gene. Physical Review Letters 111: 058102. https://doi.org/10.1103/PhysRevLett.111.058102 PMID: 23952448

57. Evan GI, Vousden KH (2001) Proliferation, cell cycle and apoptosis in cancer. Nature 411: 342. https://doi.org/10.1038/35077213 PMID: 11357141

58. Malumbres M, Barbacid M (2009) Cell cycle, CDKs and cancer: a changing paradigm. Nature Reviews Cancer 9: 153. https://doi.org/10.1038/nrc2602 PMID: 19238148

59. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144: 646–674. https://doi.org/10.1016/j.cell.2011.02.013 PMID: 21376230

60. Diaz-Moralli S, Tarrado-Castellarnau M, Miranda A, Cascante M (2013) Targeting cell cycle regulation in cancer therapy. Pharmacology & Therapeutics 138: 255–271. https://doi.org/10.1016/j.pharmthera.2013.01.011

61. Feala JD, Cortes J, Duxbury PM, Piermarocchi C, McCulloch AD, et al. (2010) Systems approaches and algorithms for discovery of combinatorial therapies. Wiley Interdisciplinary Reviews: Systems Biology and Medicine 2: 181–193. PMID: 20836021

62. Csermely P, Agoston V, Pongor S (2005) The efficiency of multi-target drugs: the network approach might help drug design. Trends in Pharmacological Sciences 26: 178–182. https://doi.org/10.1016/j.tips.2005.02.007 PMID: 15808341

63. Ágoston V, Csermely P, Pongor S (2005) Multiple weak hits confuse complex systems: a transcriptional regulatory network as an example. Physical Review E 71: 051909. https://doi.org/10.1103/PhysRevE.71.051909

64. Eser P, Demel C, Maier KC, Schwalb B, Pirkl N, et al. (2014) Periodic mRNA synthesis and degradation co-operate during cell cycle gene expression. Molecular Systems Biology 10: 717. https://doi.org/10.1002/msb.134886 PMID: 24489117

65. Dominguez D, Tsai YH, Gomez N, Jha DK, Davis I, et al. (2016) A high-resolution transcriptome map of cell cycle reveals novel connections between periodic genes and cancer. Cell Research. https://doi.org/10.1038/cr.2016.84 PMID: 27364684

66. Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, et al. (1998) Comprehensive identification of cell cycle–regulated genes of the yeast saccharomyces cerevisiae by microarray hybridization. Molecular biology of the cell 9: 3273–3297. https://doi.org/10.1091/mbc.9.12.3273 PMID: 9843569

67. Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, et al. (2002) Identification of genes periodically expressed in the human cell cycle and their expression in tumors. Molecular biology of the cell 13: 1977–2000. https://doi.org/10.1091/mbc.02-02-0030 PMID: 12058064

68. De Lichtenberg U, Jensen LJ, Fausbøll A, Jensen TS, Bork P, et al. (2004) Comparison of computational methods for the identification of cell cycle-regulated genes. Bioinformatics 21: 1164–1171. https://doi.org/10.1093/bioinformatics/bti093 PMID: 15513999

69. Rustici G, Mata J, Kivinen K, Lió P, Penkett CJ, et al. (2004) Periodic gene expression program of the fission yeast cell cycle. Nature genetics 36: 809. https://doi.org/10.1038/ng1377 PMID: 15195092

70. Kelliher CM, Leman AR, Sierra CS, Haase SB (2016) Investigating conservation of the cell-cycle-regulated transcriptional program in the fungal pathogen, cryptococcus neoformans. PLoS genetics 12: e1006453. https://doi.org/10.1371/journal.pgen.1006453 PMID: 27918582

71. Kohonen T (1974) An adaptive associative memory principle. IEEE Transactions on Computers 100: 444–445. https://doi.org/10.1109/T-C.1974.223960

72. Kauffman SA, Johnsen S (1991) Coevolution to the edge of chaos: coupled fitness landscapes, poised states, and coevolutionary avalanches. Journal of Theoretical Biology 149: 467–505. https://doi.org/10.1016/S0022-5193(05)80094-3 PMID: 2062105

73. Muramatsu S, Hirai K, Tak YS, Kamimura Y, Araki H (2010) CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol ϵ, and GINS in budding yeast. Genes & Development 24: 602–612. https://doi.org/10.1101/gad.1883410

74. Bruck I, Kaplan DL (2014) The replication initiation protein Sld2 regulates helicase assembly. Journal of Biological Chemistry 289: 1948–1959. https://doi.org/10.1074/jbc.M113.532085 PMID: 24307213

75. Helfrich BA, Kim J, Gao D, Chan DC, Zhang Z, et al. (2016) Barasertib (azd1152), a small molecule aurora b inhibitor, inhibits the growth of sclc cell lines in vitro and in vivo. Molecular cancer therapeutics 15: 2314–2322. https://doi.org/10.1158/1535-7163.MCT-16-0298 PMID: 27496133

76. Kwiatkowski N, Jelluma N, Filippakopoulos P, Soundararajan M, Manak MS, et al. (2010) Small-molecule kinase inhibitors provide insight into mps1 cell cycle function. Nature chemical biology 6: 359–368. https://doi.org/10.1038/nchembio.345 PMID: 20383151

77. Liu X, Liao W, Yuan Q, Ou Y, Huang J (2015) Ttk activates akt and promotes proliferation and migration of hepatocellular carcinoma cells. Oncotarget 6: 34309. https://doi.org/10.18632/oncotarget.5295 PMID: 26418879

78. Stratford JK, Yan F, Hill RA, Major MB, Graves LM, et al. (2017) Genetic and pharmacological inhibition of ttk impairs pancreatic cancer cell line growth by inducing lethal chromosomal instability. PloS one 12: e0174863. https://doi.org/10.1371/journal.pone.0174863 PMID: 28380042

79. Zurada JM, Cloete I, Van der Poel E (1996) Generalized Hopfield networks for associative memories with multi-valued stable states. Neurocomputing 13: 135–149. https://doi.org/10.1016/0925-2312(96)00086-0

80. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, et al. (2014) Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science 344: 1396–1401. https://doi.org/10.1126/science.1254257 PMID: 24925914

81. Scialdone A, Natarajan KN, Saraiva LR, Proserpio V, Teichmann SA, et al. (2015) Computational assignment of cell-cycle stage from single-cell transcriptome data. Methods 85: 54–61. https://doi.org/10.1016/j.ymeth.2015.06.021 PMID: 26142758

82. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, et al. (2017) Massively parallel digital transcriptional profiling of single cells. Nature Communications 8: 14049. https://doi.org/10.1038/ncomms14049 PMID: 28091601

83. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, et al. (2014) Pseudo-temporal ordering of individual cells reveals dynamics and regulators of cell fate decisions. Nature Biotechnology 32: 381.

84. Waters CM, Bassler BL (2005) Quorum sensing: cell-to-cell communication in bacteria. Annu Rev Cell Dev Biol 21: 319–346. https://doi.org/10.1146/annurev.cellbio.21.012704.131001 PMID: 16212498

85. Momeni B, Brileya KA, Fields MW, Shou W (2013) Strong inter-population cooperation leads to partner intermixing in microbial communities. eLife 2: e00230. https://doi.org/10.7554/eLife.00230 PMID: 23359860

86. Kalsbeek A, Merrow M, Roenneberg T, Foster R (2012) Suprachiasmatic nucleus: cellular clocks and networks. The Neurobiology of Circadian Timing 199: 129. https://doi.org/10.1016/B978-0-444-59427-3.00029-0

87. Skog J, Würdinger T, Van Rijn S, Meijer DH, Gainche L, et al. (2008) Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. Nature Cell Biology 10: 1470–1476. https://doi.org/10.1038/ncb1800 PMID: 19011622

88. Hong BS, Cho JH, Kim H, Choi EJ, Rho S, et al. (2009) Colorectal cancer cell-derived microvesicles are enriched in cell cycle-related mRNAs that promote proliferation of endothelial cells. BMC Genomics 10: 556. https://doi.org/10.1186/1471-2164-10-556 PMID: 19930720

89. Renzulli JF, Del Tatto M, Dooner G, Aliotta J, Goldstein L, et al. (2010) Microvesicle induction of prostate specific gene expression in normal human bone marrow cells. The Journal of Urology 184: 2165–2171. https://doi.org/10.1016/j.juro.2010.06.119 PMID: 20850816

90. Del Tatto M, Ng T, Aliotta JM, Colvin GA, Dooner MS, et al. (2011) Marrow cell genetic phenotype change induced by human lung cancer cells. Experimental Hematology 39: 1072–1080. https://doi.org/10.1016/j.exphem.2011.08.008 PMID: 21864488

91. Tetta C, Ghigo E, Silengo L, Deregibus MC, Camussi G (2013) Extracellular vesicles as an emerging mechanism of cell-to-cell communication. Endocrine 44: 11–19. https://doi.org/10.1007/s12020-012-9839-0 PMID: 23203002

92. Jones E, Oliphant T, Peterson P, et al. (2001–). SciPy: Open source scientific tools for Python. URL http://www.scipy.org/.

93. Whitfield ML, George LK, Grant GD, Perou CM (2006) Common markers of proliferation. Nature reviews Cancer 6: 99. https://doi.org/10.1038/nrc1802 PMID: 16491069

94. Hochberg Y, Benjamini Y (1990) More powerful procedures for multiple significance testing. Statistics in Medicine 9: 811–818. https://doi.org/10.1002/sim.4780090710 PMID: 2218183

95. Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, et al. (2011) Saccharomyces Genome Database: the genomics resource of budding yeast. Nucleic Acids Research 40: D700–D705.