BIOINFORMATICS ORIGINAL PAPER Vol. 22 no. 20 2006, pages 2539–2546
doi:10.1093/bioinformatics/btl432

Systems biology

Few crucial links assure checkpoint efficiency in the yeast cell-cycle network

Gautier Stoll[1], Jacques Rougemont[3] and Félix Naef[1,2,3,*]

[1]Swiss Institute of Experimental Cancer Research, ISREC, NCCR Molecular Oncology, CH-1066 Epalinges, Switzerland, [2]Ecole Polytechnique Fédérale de Lausanne, EPFL, School of Life Sciences, CH-1015 Lausanne, Switzerland and [3]Swiss Institute of Bioinformatics, SIB, CH-1015 Lausanne, Switzerland

Received on May 10, 2006; revised on July 27, 2006; accepted on August 3, 2006
Advance Access publication August 7, 2006
Associate Editor: Thomas Lengauer

ABSTRACT

Motivation: The ability of cells to complete mitosis with high fidelity relies on elaborate checkpoint mechanisms. We study S- and M-phase checkpoint responses *in silico* in the budding yeast with a stochastic dynamical model for the cell-cycle. We aim to provide an unbiased functional classification of network interactions that reflect the contribution of each link to checkpoint efficiency in the presence of cellular fluctuations.

Results: We developed an algorithm BNetDyn to compute stochastic dynamical trajectories for an input gene network and its structural perturbations. User specified output measures like the mutual information between trigger and output nodes are then evaluated on the stationary state of the Markov process. Systematic perturbations of the yeast cell-cycle model by Li *et al.* classify each link according to its effect on checkpoint efficiencies and stabilities of the main cell-cycle phases. This points to the crosstalk in the cascades downstream of the SBF/MBF transcription activator complexes as determinant for checkpoint optimality; a finding that consistently reflects recent experiments. Finally our stochastic analysis emphasizes how dynamical stability in the yeast cell-cycle network crucially relies on backward inhibitory circuits next to forward induction.

Availability: C++ source code and network models can be downloaded at http://www.vital-it.ch/Software/

Contact: felix.naef@isrec.ch

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 INTRODUCTION

The cell-cycle progression has the particularity of triggering fairly abrupt transitions between successive phases rather than following a smooth phase trajectory (Ingolia and Murray, 2004). The transitions are supervised by a system of checkpoints that allow intervention at critical cell-cycle phases according to both external and internal alarm signals (Cobb *et al*., 2004). By comparing across species, it appears that the minimal and invariable skeleton of cell-cycle oscillators consists of a negative feedback circuits similar to the circadian pacemakers (Barkai and Leibler, 2000). Such designs use an activator consisting of an activated cyclin dependent kinase (Cdk)/cyclin complex that induces its own repressor: the anaphase promoting complex (APC) which counteracts Cdk activity through proteolytic degradation of the cyclin (Ingolia and Murray, 2004). As a result the activity level of Cdk/cyclin raises until the complex is degraded and the system is reset to low Cdk/cyclin level characteristic of the pre-mitotic gap phase G1. In addition, positive feedback loops that control Cdk activity levels are mediated through Cdc25 and Wee1. Such loops were shown to induce bistability resulting in abrupt changes of the Cdk/cyclin activity at mitotic entry (Pomerening *et al*., 2003). This simplicity together with abundant genetic and biochemical data (Tyers, 2004) have made the *Saccharomyces cerevisiae* cell cycle an attractive test ground for mathematical modeling. Model of ordinary differential equations that implement chemical kinetics studied both the quantitative and qualitative dynamical behavior of the yeast cell-cycle network (YCC) in wild-type and mutants (Cross, 2003; Ingolia and Murray, 2004; Novak *et al*., 1998, 2001). Such models could recapitulate observed mutant phenotypes and predict novel characteristics that were validated experimentally (Cross *et al*., 2002). Checkpoints were studied in *Schizosaccharomyces pombe* using continuous models and bifurcation diagrams in which stable steady states were interpreted as the different phases of the cell-cycle (Novak and Tyson, 2003).

However, several basic hypotheses underlying chemical kinetics are usually not satisfied in the cellular environment. In fact, in a biological system such as the cell-cycle, a correct description at the microscopic level should include stochastic fluctuations in numbers of molecules, non-homogeneity of the medium (McAdams and Arkin, 1997; van Kampen, 1992) (ref. 11 p.171-2). Therefore, a microscopically detailed model has enough unknown parameters and such complexity that it will tend to lose its predictive capacity. It is not obvious why simple networks of effective chemical reactions can give correct predictions; one possibility may be the network wiring of interacting proteins is determinant in biological systems, rather than the choice of the dynamics applied. However this cannot hold in full generality (Guet *et al*., 2002; Samoilov *et al*., 2005). For most biological pathways wiring diagrams are still derived from genetic data without further quantitative microscopic details. Consequently many successful theoretical studies focused on the qualitative dynamical behavior of models, e.g. by studying bifurcation diagrams (Chen *et al*., 2004; Novak *et al*., 1998, 2001).

For the YCC, an effective model that does not implement explicitly chemical kinetics uses a discrete description of gene activities with a minimal number of free parameters (Li *et al*., 2004).

*To whom correspondence should be addressed.

This approach applies a deterministic, discrete time dynamics in which proteins assume Boolean states ('on' or 'off' activities) interpreted as concentration levels, phosphorylation states, or presence in active complexes. This simplification finds some justification in the abundance of positive feedback circuits (Brandman *et al*., 2005; Pomerening *et al*., 2005) leading to sharp switches, as discussed above in the case of the cyclins. In the Boolean YCC model (Li *et al*., 2004), the state of each gene is updated according to the states of its parent nodes at the previous time step, via a simple threshold-based rule (or generalized 'OR' function). Despite its simplicity this model showed distinct dynamical characteristics, notably the presence of a super fixed point (large attractor) corresponding to the G1 stationary state. In addition the dynamical landscape indicated that trajectories from random initial conditions to G1 tend to have short transients before ending onto a common chain of states representing the canonical sequence through the cycle: from G1 to S to M and back to the stationary G1 phase. Because of the unusually large fraction (86%) of initial conditions ending in this stationary G1 states, this model was characterized as having a super fixed point. Although the model is deterministic, this suggests that the system can accommodate for fluctuations that would occur in the transition between states (error correction). Moreover, these properties were relatively insensitive to structural modifications in the network topology induced by link addition or removal, indicating that robustness in dynamical behavior followed from the specific wiring (topology) of the yeast cell-cycle network.

Here we study *in silico* the efficiency of S- and M-phase checkpoints by quantifying their ability to halt the cell-cycle in the proper cellular states. For example we would like to know which network links are most important in maintaining checkpoint function in a noisy cellular environment (Elowitz *et al*., 2002). To address this question, we formulate a stochastic generalization of the network by Li *et al*. (2004) and quantify how perturbations modify checkpoint responses. This provides an unbiased functional classification of links reflecting their effect on checkpoint efficiencies and stabilities of the main cell-cycle phases. Finally we predict putative interactions that enhance these properties and discuss design principles revealed by these predictions.

2 SYSTEMS AND METHODS

2.1 Stochastic dynamical model

We consider Boolean networks where each node assumes a value 0 or 1 (on or off). A network of N nodes is represented by an N by N adjacency matrix A, in which an activating link between node i and node j is represented by a positive entry $A_{ij} = 1$ and an inhibiting link by $A_{ij} = -1$. Self-inhibitory and self-activating links, $A_{ii} = \pm 1$, are also possible. A network state consists of a Boolean vector S representing the states of each node. The full phase space contains 2^N states.

In the absence of noise, the temporal evolution of the state variable is taken as in Li *et al*. (2004): the state at the next time-step $S(t+1)$ is given in terms of the current state $S(t)$ by

$$\begin{aligned} S_i(t+1) &= 1 \quad \text{if} \quad \sum_j A_{ij}S_j(t) > 0 \\ S_i(t+1) &= 0 \quad \text{if} \quad \sum_j A_{ij}S_j(t) < 0 \\ S_i(t+1) &= S_i(t) \quad \text{if} \quad \sum_j A_{ij}S_j(t) = 0. \end{aligned}$$

Thus nodes are updated according to a thresholded summation of their positive and negative inputs. Moreover the state is unchanged when the inputs sum to zero. Biochemical networks must be able to buffer environmental and intrinsic noise sources (Elowitz *et al*., 2002; McAdams and Arkin, 1997). To mimic such stochastic events in the Boolean context we allow nodes to flip their state randomly (Shmulevich *et al*., 2002) instead of following the deterministic updates (Supplementary Fig. S1B). Other alternatives to noiseless synchronized Boolean dynamics have considered various desynchronization schemes (Chaves *et al*., 2005; Klemm and Bornholdt, 2005; Koch *et al*., 2005; Sanchez and Thieffry, 2003; Thomas and Kaufman, 2001) which our noise implementation also partially simulates. Gradually increasing the noise strength changes between a regime dominated by the deterministic dynamics and one where transitions between states are fully random, independent of both the dynamical rules and the network topology. To model the noise, we introduce a finite probability at each time step that a node flips its state randomly instead of following the deterministic rule. This node flipping probability (NFP) is such that the probability of a stochastic update at one of the N node (without specifying which one) is $N{\cdot}$NFP. In the presence of noise, the set of attractors considered by Li *et al*. (2004) is replaced by a stationary state (example in Supplementary Material, Fig. S1). The noise is applied to all nodes, which guarantees uniqueness of the stationary state. Our stochastic model thus defines a Markov process with a unique stationary state which defines a probability distribution $p(S)$ over the state space. In the Yeast cell-cycle model (and in the toy model in the Supplementary Material), some nodes (e.g. the checkpoints) are considered as triggers and are not updated during the dynamical evolution. In that case, one stationary state is computed for each trigger state (Fig. 2A).

2.2 Input–output characteristics

To quantify input–output relationships between a set of input states $x \in X$ and output states $y \in Y$, we simulate the joint probability by applying a Markov Chain Monte Carlo method to the stochastic network model. Depending on the problem, X and Y together may not span the entire state space. In that case $p(x, y)$ is a marginal of $p(S)$. For example to measure checkpoint efficiency, X will be taken as the checkpoint states. We then compute entropies $H(X)$ and $H(Y)$, conditional entropies $H(X\,|\,Y)$ and $H(Y\,|\,X)$ and mutual information $I(X, Y)$ (cf. Supplementary Material).

2.3 Structural network perturbations

Two types of structural network perturbations are considered: the links are (1) removed or (2) added from the wild-type (unperturbed) network. Stationary states are computed for each perturbed network. To compare two stationary distributions, typically one from the wild-type network p and the other from a perturbation p', we use the probability distance measure

$$\Delta(X,Y) = \frac{1}{2}\sum_{x,y} |\,p(x,y) - p'(x,y)\,|$$

which takes values in the interval [0,1]. This measure discriminates between perturbations with behavior close to the wild-type model (small Δ) from others that compromise the biological relevance of the model.

3 IMPLEMENTATION

3.1 Yeast cell-cycle model

We study the YCC model (Li *et al*., 2004) built around the master cell-cycle regulator Cdc28 and its most important functional partners. In the simplest version of the model, cell-cycle checkpoints are by-passed resulting in 11 dynamical nodes some of which represent multiple proteins. All cyclins (Cln3, Cln1,2, Clb5,6 and Clb1,2) form complexes with Cdc28 and the latter is not explicitly part of the model. The model recapitulates the following sequence of events: (1) Re-entry into the cell cycle is triggered by activation

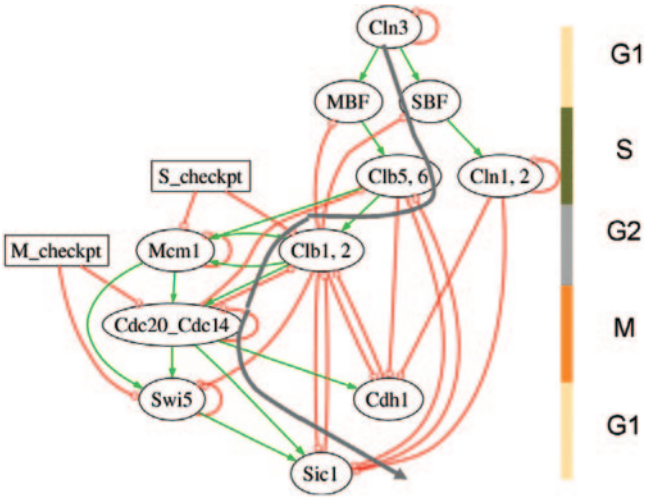

Fig. 1. Yeast cell-cycle network of (Li *et al*., 2004). Oval nodes represent either proteins or protein complexes (as for the cyclins Cln3, Cln1,2, Clb5,6 and Clb1,2); squared boxes represent checkpoints. Red link indicate inhibitory interaction (for example through ubiquitination) or decay (self-degradation); green link indicate activation, either by transcriptional induction or by posttranscriptional activation (for example through phosphorylation or complex formation). The gray path indicates the sequential activation of nodes in the original model [see Table 2 in Li *et al*. (2004)]: cell-cycle reentry is characterized by activation of Cln3 (top) which then propagates down along the G1-S-G2-M-G1 states (cf. right vertical bar) and ends in the G1 stationary state (bottom).

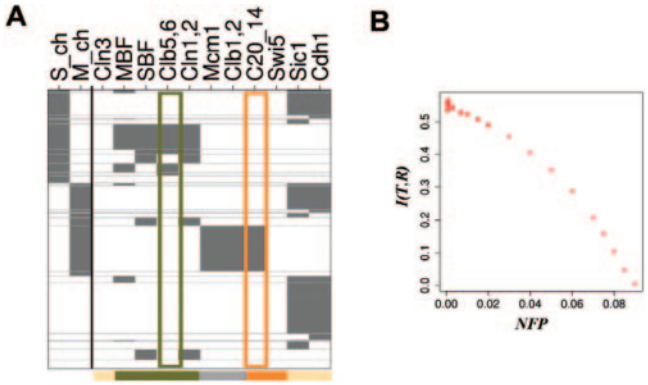

Fig. 2. Stationary state of the yeast-cell network in the presence of noise and global checkpoint responses. (**A**) Representation of the stationary state for NFP = 0.0005. Each Boolean vector S is represented by a horizontal band with active nodes in gray and inactive nodes in white. The thickness of a band is proportional to its probability in the stationary state. The S and M-phase checkpoints are in the two first columns followed by the dynamical nodes ordered according to the cell-cycle phases, colored bar at the bottom follows Figure 1. The colored boxes indicate the S and M-phase markers defining the state variables (SV). (**B**) Mutual information in function of NFP. The limit for NFP→0 exists (data not shown). The spread of values for fixed NFP represents sampling errors as obtained from multiple independent runs and emphasizes that sampling errors are largest for small probabilities. The model correctly predicts that the information I drops to zero at the maximal noise NPF = 1/11 value ($q = 1$). In both the panels $3 \times 500\,000$ iterations were used.

of Cln3 (2) the SBF and MBF transcription factor complexes are active in early S-phase followed by the Clb5,6 and Cln1,2 in late S-phase; (3) these induce the G2 markers Clb1,2 and the transcription factor complex Mcm1/SFF; (4) mitotic entry is hallmarked by the activation of the anaphase promoting complex APC/C through binding of Cdc20, which is later replaced by Cdh1 following the activation of Cdc14 (anaphase marker); (5) mitotic exit coincides with the degradation of Cdc20/14 and S/M phase cyclins after the activation of Swi5/Sic1/Cdh1 effectors. For convenience Cdc20 and Cdc14 are fused into a single node effectively collapsing early M-phase events. We used the S- and M-phase checkpoints (Fig. 1) as trigger nodes. The intra-S checkpoint slows down DNA replication in response to DNA damage during S-phase. Biochemically the activation of the Mec1-Rad53-Cdc5 cascade slows down the progression of replication forks (Cobb *et al*., 2004) and is represented in this model by blocking G2 entry or activation of Clb1, 2 and Mcm1 (Fig. 1A in Li *et al*., 2004). Owing to the fusion of the Cdc20 and Cdc14 nodes, the spindle checkpoint (blocking of the metaphase to anaphase transition through the inhibition of Cdc20/APC complex via Mps1–Bub1/3–Mad1/2/3 cascade) is merged with the DNA damage checkpoint (activation of Pds1 in response to DNA damage during the separation of the sister chromatids) and effectively termed the M-phase checkpoint. The latter is implemented by blocking Cdc14 (direct interaction) plus its main effector the transcription factor Swi5 (Fig. 1).

As explained in the results we will study two aspects of the cell-cycle model: (1) the checkpoint efficiency (CE) and (2) the phase tightness (PT). For this we split the set of dynamical nodes (all nodes except the checkpoints) into two groups: one set contains markers for the canonical cell-cycle phases and the second consisting of all remaining nodes. As markers we have chosen Clb5,6 and the composite node Cdc20/14 which are markers of S-phase and M-phase entry, respectively. These nodes define the following states: (Clb5,6; Cdc20/14) = (on, off) corresponds to the S/G2 phases (these two phases are very short, one state each in the unperturbed model as defined in Li *et al*. (2004, Table 2); (off, on) characterizes the late M-Phase (four last M-phase states); (off, off) defines G1 (five states) while the (on, on) state occurs in a single state right at M-phase entry.

3.2 Terminology

Yeast cell-cycle (YCC). SBF and MBF are transcription factors that activate gene expression during the G1/S transition of the cell cycle in yeast. SBF is a heterodimer of Swi4 and Swi6, and MBF is a heterodimer of Mbpl and Swi6. APC denotes the anaphase promoting complex.

Modeling. NFP is the node flipping probability; CE the checkpoint efficiency PT the phase tightness (PT). The static S and M checkpoint states are called triggers (T). The states of the 11 dynamic nodes are denoted (R) in Fig 2. Dynamical nodes are grouped into state variables (SV) and other nodes (O) (cf. Fig. 3).

3.3 Simulations

BNetDyn computes entropies, conditional entropies and mutual information for a given network (specified in the GraphViz dot format). For this the program estimates joint probabilities $p(x,y)$ by generating trajectories of the Markov process. For Figure 2, the cell-cycle was simulated for $3 \times 1\,000\,000$ time steps (3 is the number of checkpoint states) for different values of NFP. For

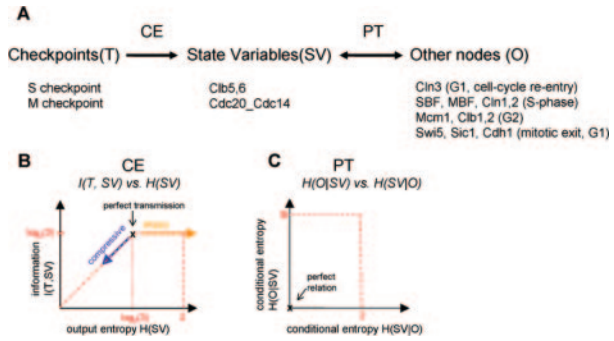

Fig. 3. Scheme used to study the dynamical response to perturbations. (**A**) The relationship between checkpoints (T) and state variables (SV) measures checkpoint efficiency (CE). The correlation between SV and the remaining nodes (O) determines a measure of phase tightness (PT) (cf. Methods). (**B**) To compare CE across networks, we represent $I(T, SV)$ versus $H(SV)$ so that the optimal case (a one-to-one relation) sits at the intersection marked by 'x'. Departure from optimality happens via two (possibly mixed) modes. The sloppy direction (orange) corresponds to relations where several outputs coexist for the same input. The compressive direction (blue) indicates that several inputs are mapped to the same output. The accessible region is delimited by the dashed lines, given by the three checkpoint states and four SV states. Any network in the accessible region can be assigned a sloppiness and compression according to the coordinates defined by the orange and blue vectors. (**C**) PT is represented using conditional entropies so that the ideal network is located at the origin. The accessible region is delimited on the x-axis by the number of state variables (2) and on the y-axis by the number of other dynamical nodes (9).

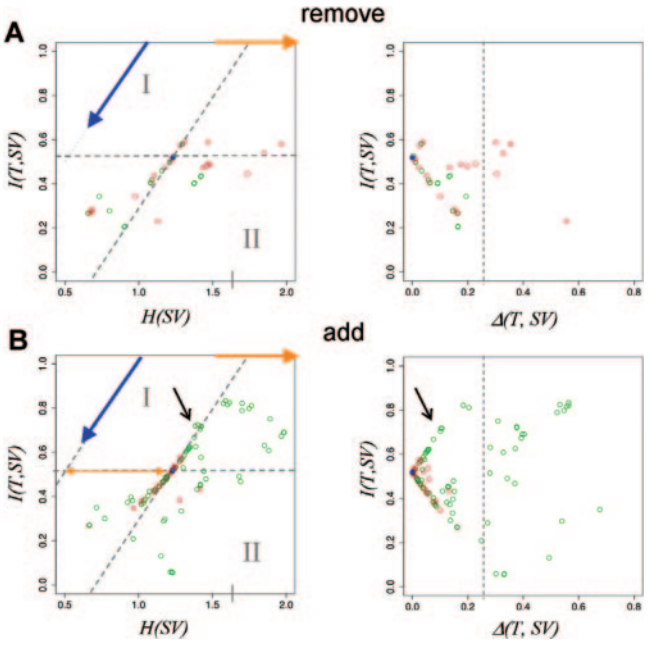

Fig. 4. Perturbation fingerprint for the checkpoints efficiency (CE). (**A**) Link removal (38 possibilities). Red (respectively green) dots indicate networks obtained from removal of one inhibitory (respectively activating) link; the wild-type network is in blue. Left: $I(T,SV)$ versus $H(SV)$. Range for $I(T,SV)$ is [0,1.58] and [0,2] for $H(SV)$. The dotted lines run parallel to the sloppy and compressive axes trough the wild-type model and delimit region (I: stabilizing links) and (II: destabilizing links). Right: $I(T,SV)$ versus $\Delta(T,SV)$ shows distance to the wild-type stationary state; y-axis is shared with the left panel. Wild-type network has $\Delta(T,SV) = 0$ by definition. Links to the right of the dashed vertical line are those whose removal generate non-biological stationary states, notice these are all inhibitory. (**B**) Link addition (162 possibilities, 81 activating/81 inhibitory). Left: $I(T,SV)$ versus $H(SV)$. Range as in A. Most perturbations move parallel to the compressive 'blue' axis; orange arrow indicates the sloppiness direction. Right: $I(T,SV)$ versus $.\Delta(T,SV)$; y-axis is shared with the left panel. In contrast to A, necessary links are by and large activating. Globally it is hard to improve the CE by perturbation; best candidates are among added activating links (black arrows). Left panels: dashed lines pass through the wild-type values. Right panels: dashed lines indicates $\Delta(T,SV) = 0.25$ (cf. 3.4).

Figures 4 and 5, networks were simulated for $3 \times 500\,000$ time steps. BNetDyn can also generate perturbed networks and evaluate distance probabilities $\Delta(x,y)$. The choice of trajectory length assures that the conditional entropies had errors of $\sim$1%. Figure 3 illustrates the method for producing Figures 4 and 5. The C++ program is available at http://www.vital-it.ch/Software/ along with the commands and input files used in this article.

3.4 Classification of network links from perturbation phenotypes

Necessary or *toxic* links are identified from networks with large $\Delta(I,SV)$ in the CE analysis (cutoff was set to 0.25), or large $\Delta(SV,O)$ in the PT (cutoff was set to 0.28). The cutoffs were fixed from a natural separation in the bimodal densities for Δ. *Stabilizing* links were defined as those whose removal would make both conditional entropies larger than the wild-type values (augmented by 1% to take into account estimated simulation errors). Such links are identifiable from the regions (II) in Figures 4 and 5. Removal of *destabilizing* links decreases both conditional entropies below wild-type values minus 1%. These links correspond to regions (I) in Figures 4 and 5. All other links are neutral.

4 RESULTS AND DISCUSSION

4.1 A probabilistic model recapitulates the yeast cell-cycle

Our goal is to provide an unbiased functional classification of links in the YCC network based on their contribution to checkpoint efficiencies and stability of the main cell-cycle phases. We start from the YCC model of Li *et al*. built around the master cell-cycle regulator Cdc28 and add S- and M-phase checkpoints (Fig. 1, Methods). The model describes the negative feedback module Cdk/cyclin→APC⊣ Cdk/cyclin using a dozen of key cell-cycle regulators providing forward induction and backward inhibition mediated mainly by ubiquitin dependent proteolysis (Supplementary Material). In the absence of noise and checkpoints this deterministic model induces a wave of activity propagating from cell-cycle reentry at G1 to S, G2, M and ending in the stationary G1 phase (Li *et al*., 2004) (Fig. 1).

One essential concern is the stability of the cell-cycle phases in a model that implements checkpoints and stochastic fluctuations. A useful cell-cycle model must have the property that the main cell-cycle phases coincide with the probable states, i.e. it would be awkward that the G1 fixed point evaporates upon dynamical perturbations such as noise or desynchronization. To determine

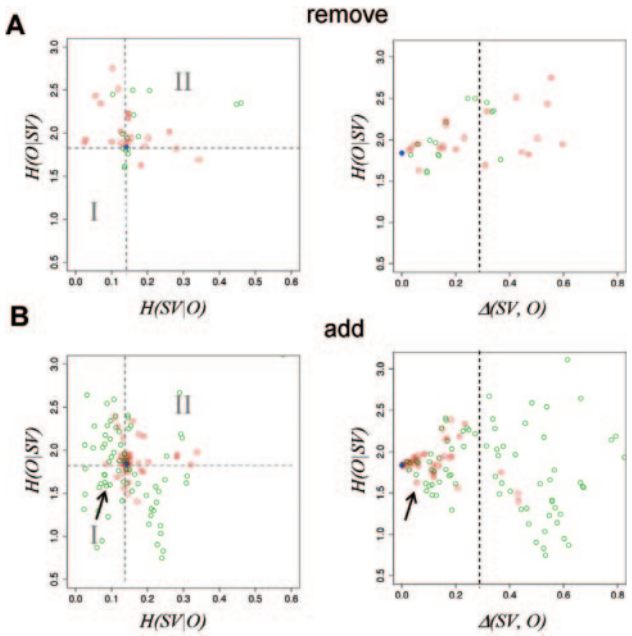

Fig. 5. Perturbation fingerprint for the phase tightness (PT). (**A**) Link removal. Colors are as in Figure 4. Left: $H(O \mid SV)$ versus $H(SV \mid O)$. Range for $H(O \mid SV)$ is [0,9] and [0,2] for $H(SV,O)$. Only four perturbations decrease both conditional entropies simultaneously (region I). Right: $H(O \mid SV)$ versus $\Delta(SV,O)$ shows distance to the wild-type stationary state; *y*-axis is shared with the left panel. Inhibitory links are dominant among necessary links whose removal generate non-biological stationary states (located right of the dashed vertical line). (**B**) Link addition. Left: $H(O \mid SV)$ versus $H(SV,O)$. A total of 36 added links increase both conditional entropies (region II). Range as in A. Right: $H(O \mid SV)$ versus $\Delta(SV,O)$; *y*-axis is shared with the left panel. In contrast to A necessary links are by and large activating. Candidate links that improve PT are among added links (arrows). Left panels: dashed lines pass through the wild-type values. Right panels: dashed lines indicate $\Delta(SV,O) = 0.28$ (cf. 3.4).

relevant noise strength, we recall that the dynamical backbone representing the canonical cell-cycle sequence is a chain of 13 Boolean states (Li *et al*., 2004). Hence the probability of completing a cell cycle without random perturbation is $q = (1 - N{\cdot}\mathrm{NFP})^{13}$ where $N = 11$ is the number of nodes in the model. Fixing NFP $= 0.0005$ leads to $q = 93\%$ and defines a regime where the cell-cycle model dominates over the fluctuations and most cell-cycles are completed after onset of Cln3. This choice is further supported by simulations (Fig. 2A) showing that the most visited Boolean states (the large horizontal bands) are G1 (Sic1 and Cdh1 active), the S-phase (SBF, MBF, Clb5,6 and Cln1,2 active) and a state characteristic of the G2/early M phases (Mcm1, Clb1,2 and Cdc20_14 active). In the latter state, the fact that Cdc20_Cdc14 is active although it is repressed by the M-phase checkpoints is a consequence of the specific update rules and the double activation by Mcm1 and Clb1,2. We have preferred to keep the original rules and thus have an M-phase checkpoint that effectively prevents activation of the mitotic exit genes. To summarize, the simplest addition of both checkpoints and noise induces a dynamical landscape that consistently toggles between S, M or G1 depending on the three independent checkpoint states 'S-checkpoint=on, M-checkpoint=off', 'S=off, M=on', 'S=off, M=off'. We verified that this behavior is robust when NFP is increased to 0.001 (the probability for an unperturbed cell-cycle is then 85%, cf. Section 4.4). Some of the smaller bands are difficult to interpret biologically and could describe novel biological states although it is more likely that these reflect model incompleteness or limitations of the modeling approach.

To evaluate the stationary state in further details, for example how the noise strength affects the property that the G1 phase is also visited when the S-phase checkpoint is on (Fig. 2A, top third), we first study the mutual information between the S and M checkpoint states (henceforth denoted by T for triggers) and the remaining 11 nodes in the network (O). Since the S and M checkpoints are taken as mutually exclusive they represent 3 states, and hence generate an entropy $H(T) = \log_2(3) \approx 1.58$ bits. This is the maximal possible information $I(T,R)$ (cf. Methods, Supplemental Material). We find that I monotonically decreases to zero, as expected (Fig. 2B). It appears that $I(T,R)$ never approaches its theoretical maximum and remains below $\sim$0.5 bits. Inspection of the stationary state for NFP $= 0.0005$ (Fig. 2A) allows identifying several reasons for low I. First, most of the smaller horizontal bands in Fig. 2A decrease information as they are not obviously correlated with checkpoint states. Second, for active S- or M-phase checkpoints, the system is not fully arrested in the corresponding phases; for example the G1 state is also frequently visited when either checkpoint is on (Fig. 2A). This is biologically plausible for several reasons, first we do not expect the S checkpoint to halt cells when the cell cycle has already passed the S-phase and entered G2. In other words, checkpoints do not attract backwards with respect to the cell-cycle progression. A similar scenario repeats for the M-phase checkpoint, however with stronger efficiency since the M-phase occurs later so that the fraction of states after the checkpoint is smaller than for the S-phase. Also partial efficiency is a property of many checkpoints and reflects variability in cellular signaling (Colman-Lerner *et al*., 2005). For example adaptation in the M-phase checkpoint has been described in *Saccharomyces cerevisiae* (Andreassen *et al*., 2003). We find therefore reasons to tolerate partially leaky checkpoints.

4.2 Checkpoint efficiency fingerprint reveals minimal sloppiness in wild-type model

The above analysis emphasized two properties. The first is the ability of the checkpoints to control cell-cycle progression and is called checkpoint efficiency (CE); it is related to the broad band structure in the stationary state representation (Fig. 2A). The second termed phase tightness (PT) is related to the fine structure in the smaller bands and reflects the susceptibility of the G1, S and M phases to external noise. To quantify both we introduce the S and M-phase entry markers Clb5,6 and Cdc20/14 as state variables (SV). CE is related to the correlation between T and SV and PT measures how tightly the SV determine the remaining nodes (O) using conditional entropies. Our analysis scheme is outlined in Fig. 3; the technical details for the computation of CE and PT are given in the Methods and Supplement. Loss and gain of function mutants are implemented *in silico* by respectively removing existing links or adding putative interactions from the original network. Input-output relationships are then quantified as sloppy and compressive. CE measures to which extent several outputs coexist

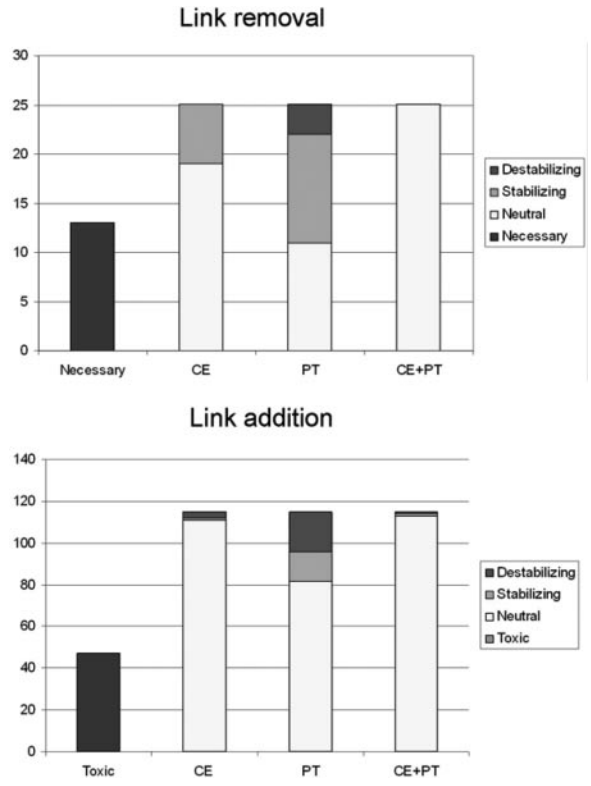

Fig. 6. Link classification according to functional categories. Top: Removed links are necessary, neutral, stabilizing or destabilizing depending on their effect on CE, PT or both. In the latter category, all links are either necessary or neutral. More than two-thirds of the 30% necessary links are inhibitory (Supplementary Table 1). Bottom: Added links can be toxic, neutral, stabilizing or destabilizing. Nearly all toxic links are activating (Supplementary Table 2).

for the same input; the second indicates whether different inputs are mapped to the same output states.

When evaluating the CE for all modified networks, we find that the sloppiness is stiff to both types of perturbations as indicated by the nearly one-dimensional accumulation of points parallel to the compressive direction (Fig. 4). In other words, the wild-type network has minimal sloppiness as no networks are found in region I in Fig. 4A and B. This suggests that evolution towards less leaky checkpoints is difficult with the applied perturbations and that this property is hardwired in the cell-cycle network. On the other hand few networks, mostly the ones with added connections, are less compressive than wild type (arrow in Fig. 4B). To also monitor which perturbations dramatically change the cell-cycle progression we measure the similitude between the perturbed and the wild-type stationary states using the distance function Δ between two stationary states (Methods). When removing links, it is striking that those leading to a biologically poor stationary state are all inhibiting; on the other hand added links with that property are mostly activating (Fig. 4, right panels). This highlights the general principle that negative feedback systems are stabilizing while positive feedback generates instabilities. In summary the CE fingerprint shows that it is globally hard to improve the CE by perturbations, nevertheless, we find a few putative link additions that slightly decrease checkpoint leakage while preserving the relevant stationary state (Fig. 4B, arrows).

4.3 Phase tightness is robust to link additions

The phase tightness fingerprint (Fig. 5) shares a common property with CE: it is difficult to improve PT by removing links from the wild-type model. Namely, there are at most four links that marginally increase PT beyond the wild-type level (region I in Fig. 5A). Moreover, link additions are mostly neutral (Fig. 6, Supplementary Table 2): either close to wild-type or out of regions I and II (Fig. 5B). This indicates robustness of the model with respect to link additions. Only a handful (fourteen) of mainly activating additions increases PT while maintaining relevant stationary states (Fig. 5B right, arrow). Finally added links that disrupt the biological states are predominantly activating while necessary links are inhibitory.

4.4 Dynamical analysis provides functional network annotation

We can now summarize the effects of perturbations on CE and PT together. First, we find that link removal from the wild-type model leads to poor cell-cycle models (with large Δ for $\sim$30% of the 38 existing links. Such links are termed necessary (defined in Methods) and are in majority (9/13) inhibitory. Among the remaining links we distinguish neutral links that induce only weak modifications when removed, and stabilizing links as those that contribute positively either to the CE or the PT (Fig. 6, Supplementary Table 1). Importantly we find no link that simultaneously destabilizes CE and PT, indicating consistency in our dynamical quantification as it would be difficult to understand such a link from an evolutionary perspective. Thus, considering CE and PT jointly, we find that all wild-type links are either necessary or neutral. Second, link additions mimicking gain of function mutations lead to neutral phenotypes in $\sim$50% of all possibilities, while $\sim$25% dramatically disrupt cell-cycle progression and are therefore termed toxic (c.f. Methods). Among these the large majority is activating reflecting the general destabilizing potential of newly created positive feedback circuits. One single added link, the inhibition of Cdc20_Cdc14 by Swi5, increases both the CE and PT.

These classifications allow us to re-annotate the original model (Fig. 7 right). First, none of the mutual activations or inhibitions (four instances, e.g. Mcm1$\leftrightarrow$ Clb1,2) are necessary, indicating that the phenotypic relevance of these network motifs might only appear under more dramatic perturbations. Also, we find no parallel activating links to be necessary, e.g. the links activating Swi5 or Sic1. Interestingly both early and late S-phase exits seem to be crucial as indicated by the double backward inhibition from Clb1,2 to the S-phase regulators SBF and MBF or the degradation of Clb5,6 by Cdc20. Finally, 5 out of 13 necessary interactions are self-degradations probably emphasizing the importance of gauging overall activity levels in the network, particularly in the presence of noise when nodes can auto-activate. Three links that enhance CE or PT are found. First the inhibition of Cdc20_Cdc14 by Swi5 is the only link that contributes positively to both CE and PT. Inspection of the stationary state for this perturbation confirms that it is mainly the M-phase checkpoint that becomes less leaky, as expected from the position of the link in the network. The two activating links SBF$\rightarrow$MBF and Cln1,2$\rightarrow$MBF are the next candidates since they stabilize PT while not changing the

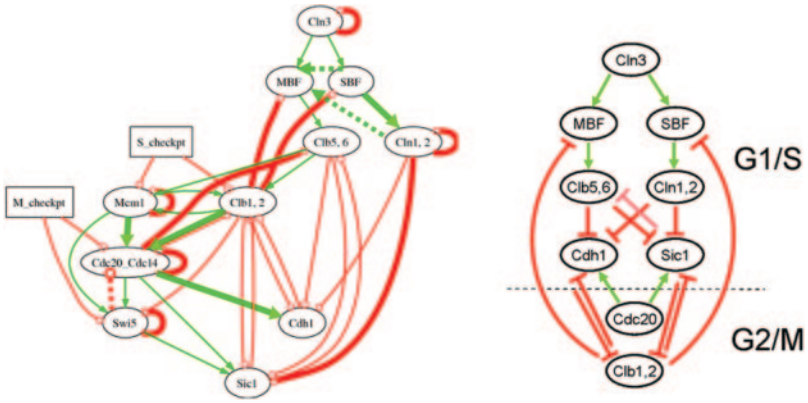

Fig. 7. Right: The thick lines show existing necessary connections, i.e. those whose removal dramatically alters the biological response. Dashed links are hypothetical connections that enhance CE and PT. Predicted connection between the SBF and MBF branches points to crucial redundancy in the targets for these regulators. Left: Symmetry in the G1/S module. This representation of all nodes in the G1/S phases and intra G1/S links emphasizes the redundancy in this sector of the model. Symmetric coupling to and from the G2/M module is also shown. With the exception of the one inhibitory connection (in pink) the nodes Sic1 and Cdh1 could be exchanged without modifying the intra G1/S connectivity structure. Even the coupling of these nodes to the G2/M module would remain symmetric except for the input from the Swi5.

CE within the simulations errors. Furthermore the hypothetical induction Cln1,2→MBF reappears as contributing to both CE and PT when the analysis starts with the original network augmented by the inhibition of Cdc20_Cdc14 by Swi5. This inhibition can be interpreted as a refinement of the Cdc20_Cdc14→Swi5 cascade that brings the induction of Swi5 closer to a toggle switch. Importantly we verified that all three links are robustly predicted using the same criteria over a range of noise strength *NFP* ranging from 0.0005 (this gives a probability of 93% to complete an unperturbed cell-cycle) to 0.001 (probability for unperturbed cell-cycle is 85%).

4.5 Crosstalk in the G1/S module increases noise tolerance

The predicted interaction between the SBF and MBF pathway through SBF→MBF or Cln1,2→MBF is more subtle. For example this crosstalk consistently reflects well-known and yet mechanistically elusive redundancy in the SBF and MBF targets (Bean *et al*., 2005). Therefore the hypothetical links between the two cascades induces at the modeling level what is known to occur biologically, namely that many MBF and SBF (at the G1/S transition) targets are shared so that both regulators must be disrupted to prevent S-phase entry (Bean *et al*., 2005; Koch and Nasmyth, 1994). Our predictions do not necessarily suggest direct interaction but could reflect an effective influence through biochemical intermediates. Figure 7 (right) emphasizes further design principles in this network. For example, many links are doubled, e.g. Cln1,2 blocks both Sic1 and Cdh1, which makes some pairs of nodes occupy unusually symmetric positions in the model. Assessing symmetry by the number of connections which are different between the original model and a model were a pair of nodes was swapped, we find that, SBF–MBF and Sic1–Cdh1 are the two most symmetric pairs, both occurring in the G1/S sub-module. The high degree of regularity in this portion of the model is emphasized (Fig. 7 left) by the 'parallel' SBF and MBF paths. Interestingly the predicted crosstalk increases the symmetry at the top: the less symmetrical part of the module; notice that the bottom (the Sic1 and Cdh1 nodes) is already highly symmetrical due to the cross inhibitions. One possibility is that these parallel paths reflect duplication events, possibly of pairs or triplets of nodes at once, in the G1/S module defined here as the Cln3, SBF, MBF, Clb5,6, Cln1,2, Sic1, Cdh1 proteins.

5 CONCLUSION/OUTLOOK

We studied checkpoint efficiencies in the yeast cell-cycle with a tractable stochastic discrete dynamical model. By reducing the full complexity of dynamical landscape to a few relevant variables, we could screen a large number of *in silico* generated mutants. We then tested the dynamical and structural robustness to identify fragile links in the network, i.e. those that are necessary for proper cell-cycle progression, as well as putative additions that enhanced the checkpoint efficiency and phase tightness. Interestingly the number of stabilizing links was small, suggesting that the wild-type model has optimality properties by design. This finding may also indicate that the cell-cycle network has been mapped to sufficient accuracy to be amenable to the kind of analysis performed. Nevertheless we found few putative additions suggesting that crosstalk in G1/S module can introduce increased dynamical tightness. Our results suggest that the YCC architecture reflects a dynamical evolutionary process in which circuits were stabilized in some portions of the network, notably the G1/S sub-module mainly through redundancy, while other portions remain more fragile. Besides revealing design principles in genetic networks we believe that such approaches will also be valuable to suggest directions for new experimental investigations.

ACKNOWLEDGEMENTS

The authors thank Mirko Bischofberger and Lorenzo Cerutti for insightful comments on the manuscript. This work was supported by the NCCR Molecular Oncology program from the Swiss National Science Foundation. The simulations were computed on the Itanium2 Intel/HP cluster at the Vital-IT facilities of the Swiss Institute of Bioinformatics.

Conflict of Interest: none declared.

REFERENCES

Andreassen,P.R. *et al*. (2003) G2 and spindle assembly checkpoint adaptation, and tetraploidy arrest: implications for intrinsic and chemically induced genomic instability. *Mutat Res*., **532**, 245–253.

Barkai,N. and Leibler,S. (2000) Circadian clocks limited by noise. *Nature*, **403**, 267–268.

Bean,J.M. *et al*. (2005) High functional overlap between MluI cell-cycle box binding factor and Swi4/6 cell-cycle box binding factor in the G1/S transcriptional program in *Saccharomyces cerevisiae*. *Genetics*, **171**, 49–61.

Brandman,O. *et al*. (2005) Interlinked fast and slow positive feedback loops drive reliable cell decisions. *Science*, **310**, 496–498.

Chaves,M. *et al*. (2005) Robustness and fragility of Boolean models for genetic regulatory networks. *J. Theor. Biol*., **235**, 431–449.

Chen,K.C. *et al*. (2004) Integrative analysis of cell cycle control in budding yeast. *Mol. Biol. Cell*., **15**, 3841–3862.

Cobb,J.A. *et al*. (2004) Redundancy, insult-specific sensors and thresholds: unlocking the S-phase checkpoint response. *Curr. Opin. Genet. Dev*., **14**, 292–300.

Colman-Lerner,A. *et al*. (2005) Regulated cell-to-cell variation in a cell-fate decision system. *Nature*, **437**, 699–706.

Cross,F.R. (2003) Two redundant oscillatory mechanisms in the yeast cell cycle. *Dev. Cell*, **4**, 741–752.

Cross,F.R. *et al*. (2002) Testing a mathematical model of the yeast cell cycle. *Mol. Biol. Cell*., **13**, 52–70.

Elowitz,M.B. *et al*. (2002) Stochastic gene expression in a single cell. *Science*, **297**, 1183–1186.

Guet,C.C. *et al*. (2002) Combinatorial synthesis of genetic networks. *Science*, **296**, 1466–1470.

Ingolia,N.T. and Murray,A.W. (2004) The ups and downs of modeling the cell cycle. *Curr. Biol*., **14**, R771–R777.

Klemm,K. and Bornholdt,S. (2005) Stable and unstable attractors in Boolean networks. *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys*., **72**, 055101.

Koch,C. and Nasmyth,K. (1994) Cell cycle regulated transcription in yeast. *Curr. Opin. Cell. Biol*., **6**, 451–459.

Koch,I. *et al*. (2005) Application of Petri net theory for modelling and validation of the sucrose breakdown pathway in the potato tuber. *Bioinformatics*, **21**, 1219–1226.

Li,F. *et al*. (2004) The yeast cell-cycle network is robustly designed. *Proc. Natl Acad. Sci. USA*, **101**, 4781–4786.

McAdams,H.H. and Arkin,A. (1997) Stochastic mechanisms in gene expression. *Proc. Natl Acad. Sci. USA*, **94**, 814–819.

Novak,B. *et al*. (1998) Mathematical model of the fission yeast cell cycle with checkpoint controls at the G1/S, G2/M and metaphase/anaphase transitions. *Biophys. Chem*., **72**, 185–200.

Novak,B. *et al*. (2001) Mathematical model of the cell division cycle of fission yeast. *Chaos*, **11**, 277–286.

Novak,B. and Tyson,J.J. (2003) Modelling the controls of the eukaryotic cell cycle. *Biochem. Soc. Trans*., **31**, 1526–1529.

Pomerening,J.R. *et al*. (2003) Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2. *Nat. Cell. Biol*., **5**, 346–351.

Pomerening,J.R. *et al*. (2005) Systems-level dissection of the cell-cycle oscillator: bypassing positive feedback produces damped oscillations. *Cell*, **122**, 565–578.

Samoilov,M. *et al*. (2005) Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations. *Proc. Natl Acad. Sci. USA*, **102**, 2310–2315.

Sanchez,L. and Thieffry,D. (2003) Segmenting the fly embryo: a logical analysis of the pair-rule cross-regulatory module. *J. Theor. Biol*., **224**, 517–537.

Shmulevich,I. *et al*. (2002) Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, **18**, 261–274.

Thomas,R. and Kaufman,M. (2001) Multistationarity, the basis of cell differentiation and memory. II. Logical analysis of regulatory networks in terms of feedback circuits. *Chaos*, **11**, 180–195.

Tyers,M. (2004) Cell cycle goes global. *Curr. Opin. Cell. Biol*., **16**, 602–613.

van Kampen,N.G. (1992) *Stochastic Processes in Physics and Chemistry*. Amsterdam, North Holland.