Modeling the Mammalian Cell Cycle as Boolean GRN

Theresa Mason Fisher

SES 598

**I. Introduction and System Description**

Understanding the cell cycle is critical due to its role in both cancer and normal developmental processes. Consequently, the gene regulatory network (GRN) that governs mammalian cell division has been subjected to considerable study.

The mammalian cell cycle is made up of five phases: G0, G1, S, G2, and M. G0 is the resting stage between bouts of cell division, when the cell is quiescent. G1, S, and G2 are collectively referred to as interphase, and are characterized by the cell making internal changes to enable it to divide. During G1 the cell increases in size to prepare for DNA replication. In S, DNA replication occurs, leading to the formation of sister chromatids. Following S, the cell continues to grow in size during G2 phase. Finally, the cycle enters M phase, where mitosis, the actual physical division of the cell into two new daughter cell, occurs.

The cell cycle is strictly regulated using a series of activators and inhibitors, in order to balance the demands of growth and reproduction against the risk of tumor formation. Inhibitors are proteins that bind to and inactivate transcription factors, and act as control mechanisms to prevent the cell from proceeding through to the next step in cell division; this allows time for DNA proof-reading and repair. As one might expect, inhibitors play a crucial role in tumor suppression, and mutations leading to their deactivation or increased degradation are strongly correlated in the development of cancer (Chu et al, 2008; Murphree and Benedict 1984). Growth is induced by the presence of an external signal, usually in the form of a cyclin-dependent kinase (CDK) protein, which, by activating its target kinase proteins, can then trigger a cascade that inactivates inhibitors and activates promoters (Nigg 1995).

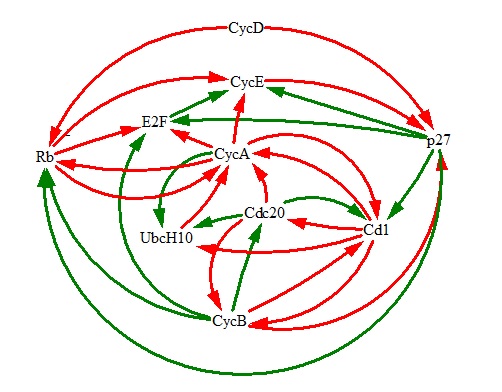
This combination of activation and inactivation forms a series of “checkpoints” which regulate when the cell enters a new phase of the cell cycle. Major checkpoints include the so-called “restriction point” at G1 (where the cell “commits” to cell division, and no longer requires an external signal) (Pardee 1989), and between G2 and M. Once the criteria for cell division is met (as expressed by the right sequence of promoter activation and inhibitor inactivation), the cell can then proceed through the checkpoint onto the next phase of the cell cycle.

Due to the dichotomy of activation/inactivation, the mammalian cell cycle GRN can be modeled as a Boolean system, where different proteins can be considered “on” or “off”.The ability to model this network can then potentially be used to study network dynamics, such as the presence of stable states or limit cycles (Faure et al 2006); this, in turn, inform experiments *in vivo*. For example, the identification of control kernels can guide the development of treatments targeting the corresponding protein in cancer cells (Sahin et al 2009).

**Model description and prior work**

The cell cycle has been a target of previous Boolean GRN modeling efforts. Initial efforts mostly focused on the cell cycle of simpler eukaryotes- particularly both budding and fission yeast (Li et al 2004, Davidich and Bornholdt 2008). The mammalian cell cycle has since been analyzed in this fashion as well- in addition to the model created by Faure et al (2006) that forms the foundation of this work, Sahin et al (2009) developed a Boolean GRN specifically examining the network existing between Rb and the ERBB receptor tyrosine kinase, with the goal of developing better targets for anticancer treatments.

As conceptualized by Faure et al (2006), the cell cycle is initiated by the presence of CycD (used as a stand-in for the cdk4/6-Cyclin D complex). Within the model, CycD is assumed to be constitutively active, and is treated as a constant. CycD acts to inactivate the tumor suppressors retinoblastoma (Rb) and p27, preventing them from blocking DNA transcription. This in turns releases the E2F transcription factors from being suppressed by Rb. E2F, a promoter, in turn activates Cyclin E (CycE) at the G1/S transition. E2F also inactivates Rb, forming a positive feedback loop that pushes the cell towards further stages of the cycle. E2F also activates Cyclin A (CycA), which maintains the inhibition of Rb and p27. CycA also inactivates E2F and CycE, as well as the Anaphase Promoting Complex, represented by its activators, Cdh1 and Cdc20. Upon inactivation of the APC, Cyclin B (CycB) production is released from inhibition; *in vivo*, CycB promotes the transition to mitosis and physical cell division during telophase. Cdc20 is activated by CycB in a negative feedback loop, and then goes on to degrade CycA and indirectly activate Cdh1. Finally, UbcH10 helps regulate the deactivation rate of CycA in combination with Cdh1 (see fig. 1).



Faure et al (2006)'s model was implemented in GINsim as Boolean GRN, where each of the above proteins is presented as a node, and assigned a value of 1 if active, 0 if inactive. Interactions between proteins (inhibition or promotion) makes up the edges of the network in turn. The model was run with three different rules for updating- synchronous, asynchronous, and mixed rules. The synchronous form was simplest, yielding two attractors (one fixed, one a seven-state limit cycle), but was considered an unrealistic model of *in vivo* processes. At the other extreme, the asynchronous model also generated

Fig. 1. The causal graph for the network of the cell cycle.

the same fixed point attractor, and a much more complex limit cycle composed of 112 states. As a middle-ground solution, the mixed rules variant modeled some processes as synchronous, but others asynchronously, via the use of priority classes. The priority classes were selected to represent *in vivo* fast and slow biochemical processes. The mixed rules version yielded a simpler 18-state limit cycle, and the same fixed attractor. Over all, the overall dynamics of the system was preserved regardless of updating rules.

Due to its simplicity, the synchronous model developed by Faure et al (2006) was selected for replication and analysis. The model and its analysis was coded in Python, making heavy use of the Network X library. Much of the code concerning the network dynamics was derived from code developed to analyze the Boolean GRN of fission yeast (Walker et al 2015), avoiding the need to “reinvent the wheel.” The resulting modeled network is composed of ten nodes connected via 34 edges.

Each protein in the GRN is assigned a value- either a 1, indicating an active state, or a 0, indicating inactivation. The protein and its corresponding value are stored in a dictionary structure.

The update rules for the system is based on those devised for the model developed by Faure et al (2006). A protein's state is determined by use of simple Boolean statements that determine whether or not the node's inhibitors and activators are in the right state of activation and inactivation e.g.

if prevState['E2F'] == 1 and prevState['Rb'] == 0:

currState['CycE']= 1('Cdh1', 'CycB') 0.331732

else:

currState['CycE'] = 0

In the example above, CycE is switched an active state if E2F is active in the previous timestep and Rb is inactive. If these conditions are untrue, CycE is switched to inactive. The model program iterates through the dictionary, verifies the state of each node's edges, and then, following the update rules, changes the value accordingly.

For the purposes of analyzing the network dynamics of the Boolean GRN, the initial values were set at as CycD and Rb=1, and all others set to zero.

**Results**

The model was able to replicate the attractors reported by Faure et al (2006) (see table 1, figs. 2 and 3). The fixed point represents the quiescent state where CycD is not present and Rb is active. The limit cycle approximately corresponds to the *in vivo* cell cycle of division activated by the presence of CycD.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Protein** | CycD | Rb | E2F | CycE | CycA | p27 | Cdc20 | Ubch10 | Cdh1 | CycB |
| Fixed Point | | | | | | | | | | |
| **Value** | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| Limit Cycle | | | | | | | | | | |
| **Value** | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
|  | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 |
|  | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
|  | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
|  | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
|  | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
|  | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |

**Table 1**. The attractor states for the Boolean GRN model in terms of node values. The attractor states are made up of a fixed point and a limit cycle.

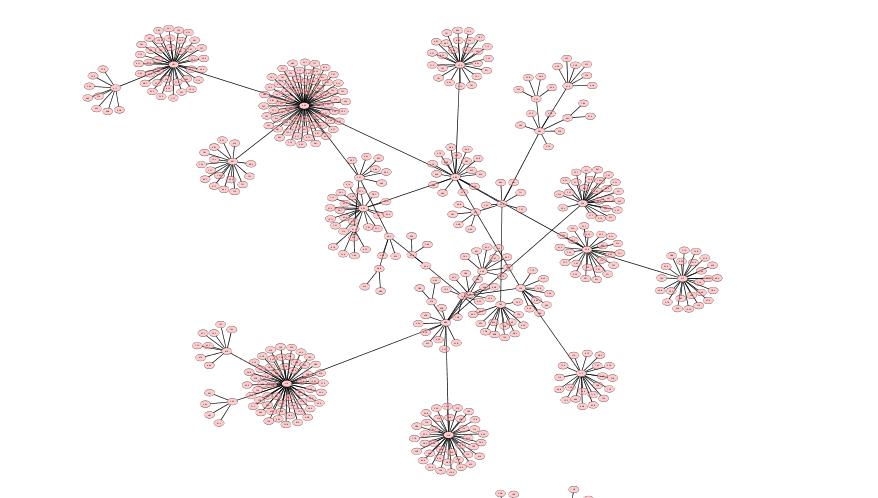


Fig.2 The attractor basin of the limit cycle.

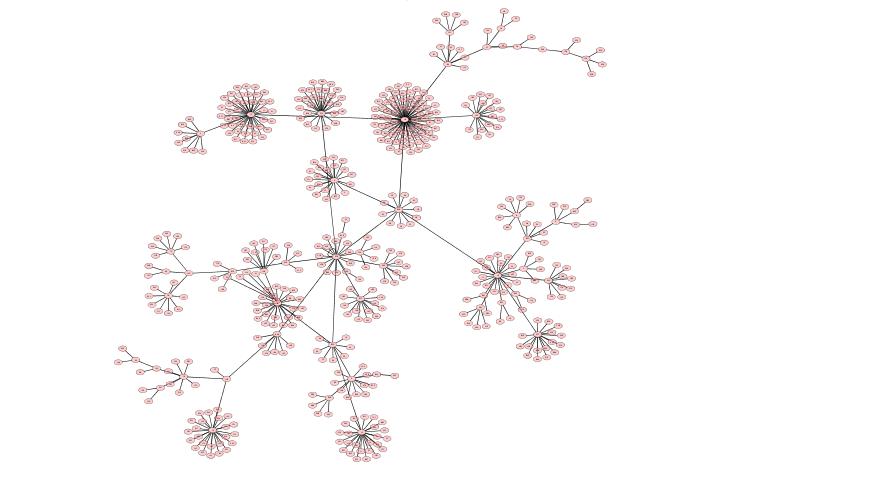


Fig. 3 The attractor basin for the fixed point attractor.

The control kernels of the network appear to be Rb, and CycD. Knocking out CycD eliminates the limit cycle attractor, as does making Rb constitutively active. Knocking out Rb eliminates the fixed point, and also generates an additional limit cycle attractor.

The second cyclical attractor is an 8 state limit cycle made that proceeds 0001001011, 0011010010, 00011110010, 0011100010, 00001100000, 00000100101, 00000101101, 00000001110. Notably, this cycle does not require CycD to proceed, and presumably could operate without the presence of an external signal to proliferate. Thus, *in* vivo it could potentially be thought of as carcinogenic mutation due loss of function in a tumor suppressor.

P27 follows a similar pattern to Rb- constitutive activity suppresses the limit cycle, though a fixed point attractor still exists if p27 is knocked out (unlike Rb). This fixed point is almost identical to that of the wild type's, the only difference being p27 inactivity. As such, it may not necessarily be considered a control node in a strict sense, despite the important role it plays in the regulation of the system dynamics.

Transfer entropy per node appears to be approximately scale-free (see fig. 3), despite the network itself not having a scale-free structure (though this may be due to its small size). As many biological networks exhibit scale-free behavior (Barabási and Oltvai 2004), this was unsurprising. The nodes CycD, Rb have the highest TE on a per node basis, of 1.064 and 1.063, respectively- this is likely due to their status as control nodes. P27 also has a high TE value, of 1.047, reflecting its important role in regulation.

Fig. 3. Transfer entropy of node-node interactions of the wild type and Rb knockout models, in rank order.

Interesting, in terms of node-node interactions, the TE highest individual values (those above 0.3) were

found where a node was interacting with CycB or Cdc20. This may be due to the negative feedback loop that exists between CycB and Cdc20, and the role CycB has in internally regulating the cell cycle as a deactivator of both Rb and p27 (inhibitors) and E2F (a promoter).

The transfer entropy values in the Rb knockout variant are substantially different compared to the wildtype, and over all, the system appears to be processing less information (see fig. 3). This may reflect loss of information from extracellular signals that normally function to control and regulate the cell cycle. In this sense, the reduction in TE can be thought of a warning signal towards carcinogenesis- the cell is reaching towards a state where it no longer requires external information to reproduce, having become metaphorically deaf to the body's cries.

Active information is highest for the control nodes Rb and CycD, as well as p27 (see fig. 4),likely due to the fact that these nodes remain essentially in the same state throughout the simulation due to the initial conditions. More importantly, they all play an important role in regulating the cycle. Correspondingly, in the Rb knockout mutant, p27's AI is reduced, whereas the AI of proteins that promote cell division- CycD, E2F, CycA, and CycE- increased. As with transfer entropy, this may reflect a loss of control in the dynamics of the cell cycle, and a shift towards a state cell division is prioritized at the loss of other cell functions.

Fig. 4. Active information of each node, for wild type and Rb knockout models.

Overall, the onset of carcinogenesis appears to be associated with major changes in the information dynamics of the cell, with a degradation of processing abilities, and increased activity in nodes associated with proliferation.

**Summary and Conclusion**

By building off of the Boolean GRN model developed by Faure et al (2006), we were able to shine a light on the processes that contribute to loss of health due to overproliferation and carcinogenesis, and analyze them from a perspective of information processing.

In normal, healthy cells, a quiescent, non-proliferative state exists as an attractor along with the limit cycle associated with cell division. The shift between the two can be regulated by extracellular signaling that allows the host organism fine control over cell growth. This extracellular regulation is aided by the presence of internal negative feedback loops, keeping the cell in the right critical state. We have demonstrated that two proteins, CycD and Rb, act as control nodes, and are central to determining which attractor basin the cell is placed in (the former associated with the limit cycle, the latter with the fixed point quiescent state).

In the absence of regulation by the control node of Rb, carcinogenesis becomes possible, as the cell gains the ability to reproduce in the absence of an external signal, leading to a second, CycD independent limit cycle and the loss of the fixed point attractor. This is accompanied by an overall loss in the information processing of the cell, as its ability to self-regulate is degraded. In turn, the activity of proliferative factors are elevated, further increasing the drive to divide endlessly.

The dynamics of the cell cycle are crucial to physiological health and development (or lack thereof). Boolean GRN models offer a useful tool in understanding how this system functions, and how it can go awry.

**Works Cited**

Chu, I. M., Hengst, L., & Slingerland, J. M. (2008). The Cdk inhibitor p27 in human cancer: Prognostic potential and relevance to anticancer therapy. *Nature Reviews Cancer Nat Rev Cancer,* *8*(4), 253-267. doi:10.1038/nrc2347

Davidich, M. I., & Bornholdt, S. (2008). Boolean Network Model Predicts Cell Cycle Sequence of Fission Yeast. *PLoS ONE,* *3*(2). doi:10.1371/journal.pone.0001672

Faure, 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*(14). doi:10.1093/bioinformatics/btl210

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,* *101*(14), 4781-4786. doi:10.1073/pnas.0305937101

Murphree, A., & Benedict, W. (1984). Retinoblastoma: Clues to human oncogenesis. *Science,* *223*(4640), 1028-1033. doi:10.1126/science.6320372

Pardee, A. (1989). G1 events and regulation of cell proliferation. *Science,* *246*(4930), 603-608. doi:10.1126/science.2683075

Sahin, Ö, Fröhlich, H., Löbke, C., Korf, U., Burmester, S., Majety, M., . . . Arlt, D. (2009). Modeling ERBB receptor-regulated G1/S transition to find novel targets for de novo trastuzumab resistance. *BMC Systems Biology BMC Syst Biol,* *3*(1), 1. doi:10.1186/1752-0509-3-1