Carnegie Mellon

Carnegie Mellon
SCHOOL OF COMPUTER SCIENCE

Model Checking and Pancreatic Cancer

Haijun Gong, Paolo Zuliani, Anvesh Komuravelli, Qinsi Wang, Michael Lotze, James Faeder, Edmund Clarke

CMACS

Abstract

Pancreatic ductal adenocarcinoma (PDAC), a 4th leading cause of cancer-related mortality in the United States, is characterized by a number of genetically altered cellular signaling pathways and overexpressed growth factors. Model Checking is a formal verification technique widely used for the automated verification and analysis of hardware systems and digital circuits.

Recent studies on pancreatic cancer cells have found that the overexpression of HMGB1, a DNA-binding protein, can decrease apoptosis and increase cancer cell's survival time. To systematically understand the signaling components that link HMGB1 and cancer risk, we constructed a rule-based model [1, 2] of the HMGB1 network which was implemented using the BioNetGen language. In [1,2], we applied Statistical Model Checking method to verify some linear temporal logic (LTL) properties in the rule-based stochastic models of HMGB1.

Accumulating evidence suggests that pancreatic cancer incidence might be associated with diabetes mellitus, especially Type II diabetes which is characterized by hyperinsulinaemia, hyperglycaemia, obesity, and overexpression of multiple WNT pathway components. In [3], we constructed a single-cell Boolean network model, and applied Symbolic Model Checking method to verify some computation tree logic (CTL) properties related to insulin resistance, cancer cell proliferation and apoptosis.

Rule-based Model of HMGBl [1,2]

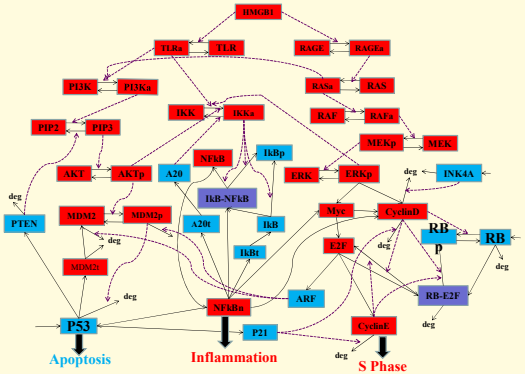


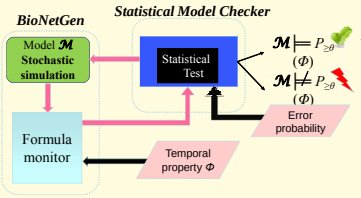
Figure 1. Schematic view of HMGB1 signal transduction. Blue nodes represent tumor suppressor proteins, red nodes represents oncoproteins/lipids. Solid lines with arrows denote protein transcription, degradation or changes of molecular species; dashed line with arrows denote activation processes.

We formulated a reaction model corresponding to the reactions illustrated in Fig.1 in the form of rules specified in the BioNetGen language, The ordinary differential equation (ODE) method and stochastic simulation algorithm (SSA) are used to simulate the model. Example ODE and BioNetGen rules:

$$\frac{d}{dt}MDM2_p(t) = k_1 AKT_p(t) MDM2(t) - d_1 MDM_p(t) - d_2 MDM2_p(t) - d_3 ARF(t) MDM2_p(t)$$

- MDM2 phosphorylation: $MDM2(a\sim U) + AKTp \rightarrow MDM2(a\sim p) + AKTp$ k1
- MDM2p dephosphorylation: $MDM2(a\sim p) \rightarrow MDM2(a\sim U)$ d1
- MDM2p degradation: $MDM2(a\sim p) \rightarrow Trash()$ d2
- MDM2p degradation : $MDM2(a\sim p) + ARF \rightarrow ARF$ d3

Statistical Model Checking: $M \models P_{\geq\theta}(\Phi)$



Require: *Property* $P_{\geq\theta}(\Phi)$, *Threshold* $T \geq 1$, *Prior* g
$n := 0$ {number of traces drawn so far}
$x := 0$ {number of traces satisfying Φ so far}
repeat
  $\sigma :=$ draw a sample trace from BioNetGen (iid)
  $n := n + 1$
  if $\sigma \models \Phi$ **then**
    $x := x + 1$
  endif
  $\mathcal{B} := BayesFactor(n, x, \theta, g) = \frac{P(X|H_0)}{P(X|H_1)}$
until ($\mathcal{B} > T \vee \mathcal{B} < 1/T$)
if ($\mathcal{B} > T$) **then**
  return "H_0 *accepted*"
else
  return "H_0 *rejected*"
endif

Figure 2. Flow chart and Pseudocode for Statistical Model Checking

HMGB1 Simulation Results [1,2]

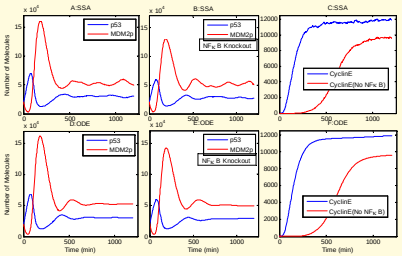


Figure 3. Baseline simulation results for the SSA (A-C) and ODE (D-F) models.

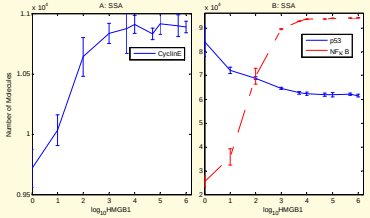


Figure 4. Overexpression of HMGB1 leads to the increase of DNA replication proteins Cyclin E and NFkB, decrease of p53.

Verification of HMGB1 Stochastic Model [1,2]

Property 1: Overexpression of HMGB1 will induce the expression of the cell cycle regulatory protein CyclinE

$$P_{\geq 0.9}\, F^{600}\,(\,CyclinE > 900\,)$$

HMGB1	# samples	# Success	Result
10^2	9	0	False
10^3	55	16	False
10^6	22	22	True

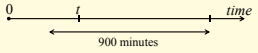
"*within 600 minutes, the number of CyclinE molecules will be greater than 900*"

Property 2: p53 is expressed at low levels in normal human cells:

$$P_{\geq 0.9}\, F^{t}(\, G^{900}\,(\,p53 < 3.3 \times 10^4\,)\,)$$

t (min)	# Samples	# Success	Result	Time (s)
400	53	49	True	597.59
500	23	22	True	271.76
600	22	22	True	263.79

"*within t minutes, p53 will stay low for 900 minutes*"



Property 3: Expression level of HMGB1 influences the 1st peak of p53's level

$$P_{\geq 0.9}\, F^{100}(\,p53 \geq a\ \&\ F^{100}(\,p53 \leq 4 \times 10^4\,)\,)$$

HMGB1	a (x 10^4)	# Samples	# Success	Result	Time (s)
10^3	5.5	20	3	False	29.02
10^2	5.5	22	22	True	19.65
10^2	6.0	45	12	False	56.27
10	6.0	38	37	True	41.50

"*within 100 minutes, p53 will pass a, and in the next 100 minutes it will eventually be below 4x10^4*"

Diabetes-Cancer Boolean Network Model [3]

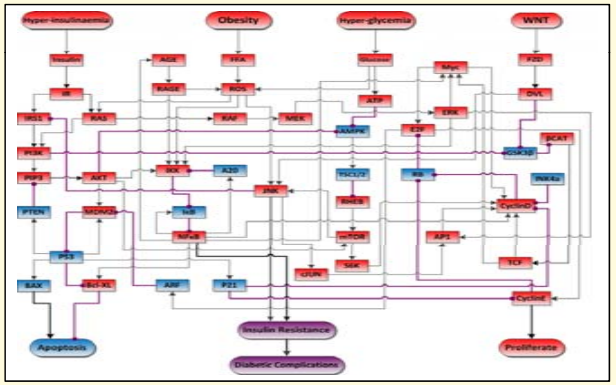


Figure 5 Schematic view of signal transduction in the diabetes-cancer model.

1. Graph G
2. Boolean transfer function
3. State of each node could be either ON (1) or OFF (0) at any time step.

The Boolean transfer function describes the transformation of the state of node from time t to t + 1 [3,4,5],

$$n(t+1) = \left(n(t) \vee \bigvee_{a \in A(n)} a(t)\right) \wedge \neg\left(\bigvee_{i \in I(n)} i(t)\right)$$

Symbolic Model Checking (SMV)

```
MODULE MAIN
VAR
    Obesity, Hyperglycemia, Hyperinsulinaemia: boolean;
    AMPK, TSC, AKT, ... : boolean;
ASSIGN
    init(TSC):={0,1}; init(AMPK):={0,1}; ...
    next(TSC):=      //update rule for TSC
      case
        AMPK: 1;
            1: TSC;
      esac;
    ... ...          //update rules for other variables
SPEC AG(AKT → AF(Resistance & Proliferate)); // property verification
```

SMV code can be divided into three parts:

Variable declarations ("boolean");

Initialization of the states for each variable with **init**;

Implementation – updating the state of each node in the state transition diagram with **next**

The verification of CTL properties is encoded using the **SPEC** statement.

Model Verification[3]

Question 1: Could the diabetes risk factors induce the oscillations of NFkB's expression level in the nucleus; and negative feedback of P53-MDM2?

AG{ (!NFkB → **AF** (NFkB)) & (NFkB → **AF** (!NFkB))}

AG{ (**P53** → **AF** (MDM2)) & (MDM2 → **AF** (!P53))}

Question 2: Do diabetes risk factors influence the risk of cancer or cancer prognosis?

a: AF(Proliferate)	**a': EF**(Proliferate)
b: AF(Apoptosis)	**b': EF**(Apoptosis)
c: AF(Resistance)	**c': EF**(Resistance)

Normal Cells: Properties c and b'- c' are true, while the rest are false. Diabetes risk factors might not increase the risk of cancer under normal conditions.

Precancerous or cancerous cells: All but property b are true. Under the influence of diabetes risk factors, cancer cell proliferation and insulin resistance are unavoidable.

Question 3: What signaling components are common and critical to both diabetes and cancer?

AG{ **RAS** → **AF** (Resistance & Proliferate & !Apoptosis) }
AG{ **AKT** → **AF** (Resistance & Proliferate & !Apoptosis) }
AG{ **NFkB** → **AF** (Resistance & Proliferate & !Apoptosis) }
AG{ **ROS** → **AF** (Resistance & Proliferate & !Apoptosis) }

Continuous activation or overexpression of RAS, AKT, NFkB or ROS will induce the proliferation of precancerous or cancerous cells, inhibit apoptosis and augment insulin resistance regardless of presence of diabetes risk factors.

Conclusions

1. Overexpression of HMGB1 will promote the expression of cell cycle regulatory protein Cyclin E and NFkB, inhibit the pro-apoptotic protein p53.
2. Diabetes risk factors could increase the risk of cancer after the proteins ARF and INK4a lose their functions.
3. Statistical Model Checking and Symbolic Model Checking techniques can be effectively combined in the signaling pathway and verify some important temporal properties.

Acknowledgement

This work was supported by a grant from the U.S. national Science Foundation's Expeditions in Computing Program (award ID 0926181).

References

[1] H. Gong, P. Zuliani, A. Komuravelli, J. Faeder, E. Clarke, Analysis and Verification of the HMGB1 Signaling Pathway, *BMC Bioinformatics*, 2010

[2] H. Gong, P. Zuliani, A. Komuravelli, J. Faeder, E. Clarke, Computational Modeling and Verification of Signaling Pathways in Cancer, *Algebraic and Numeric Biology* , 2010

[3] H. Gong, P. Zuliani, E. Clarke, Model Checking of a Diabetes-Cancer Model, *International Symposium on Computational Models for Life Sciences, AIP Conf. Proc. 1371*, 2011

[4] H. Gong, Q. Wang, P. Zuliani, J. Faeder, E. Clarke, Symbolic Model Checking of Signaling Pathways in Pancreatic Cancer, *3rd International Conference on Bioinformatics and Computational Biology*, 2011

[5] H. Gong, P. Zuliani, Q. Wang, E. Clarke, Formal Analysis for Logical Models of Pancreatic Cancer, *50th IEEE Conference on Decision and Control and European Control Conference*, 2011