Model Checking of a Diabetes-Cancer Model

Haijun Gong Paolo Zuliani Edmund M. Clarke

Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213 USA

Abstract. Accumulating evidence suggests that 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. These diabetes risk factors can activate a number of signaling pathways that are important in the development of different cancers. To systematically understand the signaling components that link diabetes and cancer risk, we have constructed a single-cell, Boolean network model by integrating the signaling pathways that are influenced by these risk factors to study insulin resistance, cancer cell proliferation and apoptosis. Then, we introduce and apply the Symbolic Model Verifier (SMV), a formal verification tool, to qualitatively study some temporal logic properties of our diabetes-cancer model. The verification results show that the diabetes risk factors might not increase cancer risk in normal cells, but they will promote cell proliferation if the cell is in a precancerous or cancerous stage characterized by losses of the tumor-suppressor proteins ARF and INK4a.

Keywords: Model Checking, Diabetes, Cancer, Signaling Pathway, Boolean Network
PACS: 87.17.Aa, 87.18.Mp, 87.18.Vf, 87.19.xj

INTRODUCTION

Cancer and diabetes are two highly malignant diseases which, along with cardiovascular disease, account for two-thirds of all deaths in the United States [1]. Understanding the causes of cancer and diabetes at the molecular level will help us to detect and treat these diseases at an early stage. Diabetes is typically divided into two major subtypes, Type 1, and the most common form Type 2, constituting over 90% of the diabetes population. Type 2 diabetes is characterized by hyperglycemia [2], β-cell dysfunction [3], hyperinsulinaemia caused by insulin resistance [4] or insulin treatment, and activation of the WNT pathway [5].

Tumorigenesis is a multistage process requiring multiple genetic mutations in the normal cell before the malignant transformation. According to the *Hallmarks of Cancer* [6], the progression of cancer requires six essential modifications in cell physiology: "*self-production of growth signals, insensitivity to antigrowth signals, evasion of apoptosis, uncontrolled replicative potential, sustained angiogenesis, and tissue invasion and metastasis*" [6].

Clinical and epidemiological studies indicate that diabetes mellitus might be a potential risk factor for the development of many types of cancer, including those of the pancreas [7], colon [8], and breast [9, 10]. For Type 2 diabetes patients, the risk for pancreatic, colon, and breast cancer increases approximately by 50%, 30%, and 20% respectively [1, 9]. About 80% of patients with cancer of the pancreas were diagnosed to have glucose intolerance or frank diabetes [11]. High blood glucose levels, insulin resistance and its compensatory hyperinsulinaemia might be implicated in the development

of cancer [2, 4]. Studies show that obesity is the most common cause of insulin resistance in humans [3]. Adiposity can increase the free fatty acid (FFA) level, promoting the production of ROS (Reactive Oxygen Species) which can disrupt insulin signaling and reduce insulin receptors' sensitivity [2, 4]. Studies with Type 2 diabetic patients also found that the expression level of multiple WNT pathway components, including TCF and β-Catenin, are elevated with respect to nondiabetic individuals [5]. These diabetes risk factors could provide growth stimulation and energy to activate several signaling pathways that regulate the progression of the cell cycle, leading to cancer cell growth.

We now ask the following question: Could the diabetes risk factors increase cancer risk in diabetes patients? If yes, how do they influence the signaling pathways in cancer cells, at the molecular level? If not, is there any association between diabetes and cancer? Computational modeling and verification techniques can help to better understand the interactions of different signaling pathways influenced by the diabetes risk factors.

Model Checking [12] is an automated verification technique for finite state transition systems which has been applied successfully for the verification of digital circuits. The technique determines whether or not a given model satisfies a desired property/specification written in a propositional temporal logic. Model Checking algorithms exhaustively and efficiently search the state space of the system model to determine the truth of a specification; if the property is not satisfied, it will output a counterexample to the desired specification. Given a biological network modeled as a transition system, Model Checking can help to verify desired properties and signal transduction sequences which will drive the network to a pre-specified state at or before a pre-specified time [13]. Several formal verification studies [14, 15], including our recent works in Statistical Model Checking [16, 17] and Symbolic Model Checking [18], demonstrated that Model Checking is a powerful technique for verifying temporal logic properties of biological networks.

No conclusive evidence has been obtained for the possible association of diabetes mellitus with cancer risk. Current diabetes-cancer association research mostly focuses on cohort analyses of large-scale populations [7, 8, 9, 10]. Some mathematical models have been proposed to study glucose regulation and insulin-signaling pathways in the cell [19, 20]. To the best of the authors' knowledge, this work is the first attempt to apply Model Checking to investigate the association between cancer and diabetes, based on a computational model of the signaling pathway network.

In this paper, we first introduce some well-known signaling pathways activated by the diabetes risk factors which are important in the progression of the cell cycle. The diabetes-cancer model is described by a Boolean network (BN). Then, we introduce Computation Tree Logic (CTL) and the Symbolic Model Checking tool SMV. Finally, a number of temporal logic properties are proposed to study the diabetes-cancer model.

DIABETES-CANCER MODEL

An extensive literature search was performed to construct a signaling network model that includes diabetes risk factors such as hyperglycemia, hyperinsulinaemia, obesity, and activation of the WNT pathway. One of our motivations is to investigate whether there is any association between diabetes and cancer incidence.

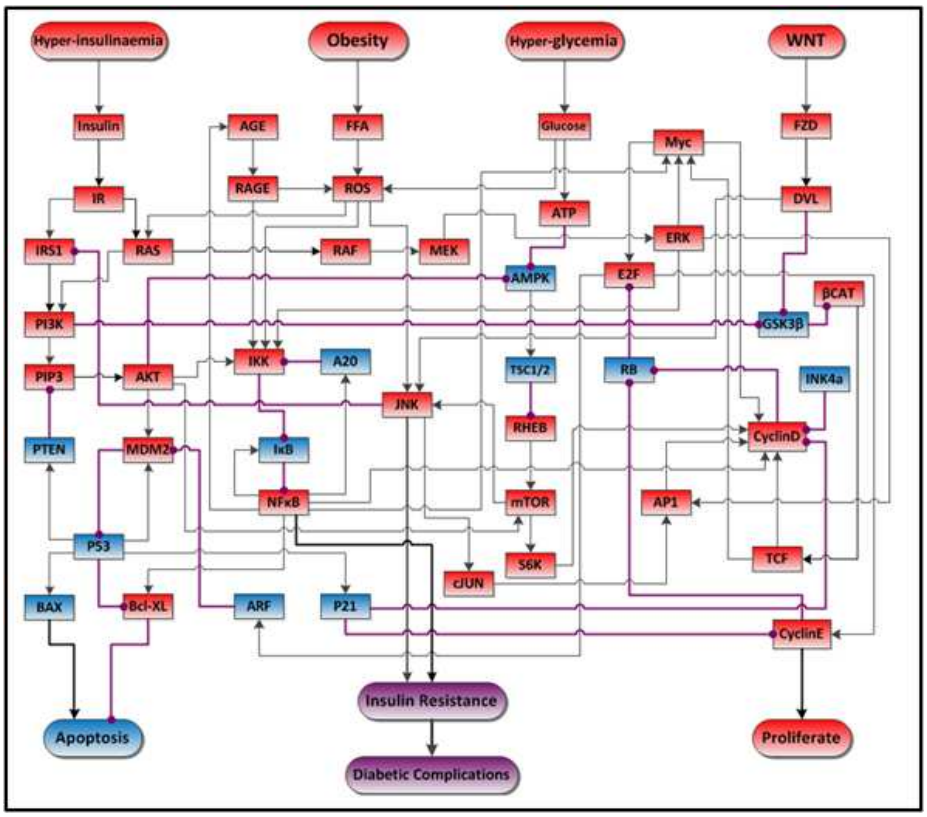

FIGURE 1. Schematic view of signal transduction in the diabetes-cancer model. Blue nodes represent tumor-suppressor proteins, red nodes represent oncoproteins/lipids. Arrow → represents protein activation, circle-headed arrow ⊸ represents deactivation.

In Fig. 1 we illustrate the crosstalk model of different signaling pathways, including the Insulin-PI3K-P53, Insulin-RAS-RB, Obesity-ROS-JNK, Glucose-AMPK-mTOR, RAGE-NFκB and WNT pathways. Some pathways in Fig. 1 have been discussed in our recent works [16, 17, 18]. The crosstalk of these signaling pathways could also influence the cell's fate. We denote activation by →, while inhibition is denoted by ⊸ or ⊣. Here, we will briefly explain the interplay of these signaling pathways, and also their association with apoptosis, tumor proliferation and insulin resistance.

Insulin-PI3K pathway: Insulin secreted from pancreatic β cells in response to elevated glucose levels activates the insulin receptors (IR)[4], leading to the phosphorylation and activation of several intracellular proteins, including the insulin receptor substrate (IRS1). The binding of IRS1 to IR will activate the oncoprotein PI3K, which in turn phosphorylates the lipid PIP2 to PIP3, initiating a cascade of reactions including the phosphorylation and activation of AKT and MDM2, thereby attenuating P53's transcription activity in the nucleus [21]. The tumor-suppressor protein P53 is mutated in more than 50% of cancers [22], and it is also a transcription factor for several tumor-suppressor proteins, including PTEN, BAX, and P21. The main pathway is written as $Insulin \rightarrow IR \rightarrow IRS1 \rightarrow PI3K \rightarrow PIP3 \rightarrow AKT \rightarrow MDM2 \dashv P53 \rightarrow$

$\{PTEN, BAX, P21\}$.

Insulin-RAS-RB pathway: Insulin can also activate the RAS protein [4] to promote cell cycle progression through the activation of the RAF, MEK, and ERK proteins, upregulating the expression level of the regulatory protein Cyclin D, enabling the cell cycle to progress through the $G1$ phase: $RAS \rightarrow RAF \rightarrow MEK \rightarrow ERK \rightarrow CyclinD$. The cell cycle's progression from phase $G1$ to S is induced by the Cyclin E and CDK2 complex. In the normal cell, the unphosphorylated tumor-suppressor protein RB binds to and inhibits $E2F$'s transcription activity. After RB is phosphorylated and inhibited by Cyclin D, $E2F$ will be activated to regulate the transcription of Cyclin E, which promotes the $G1$-S transition [23]. The main pathway is summarized as $CyclinD \dashv RB \dashv E2F \rightarrow CyclinE \rightarrow Proliferate$.

Obesity-ROS-JNK pathway: $Obesity \rightarrow FFA \rightarrow ROS \rightarrow JNK \{\rightarrow cJUN, \dashv IRS1, \rightarrow Insulin\text{-}Resistance\}$. The elevated free fatty acid (FFA) due to obesity can cause oxidative stress [2], leading to the production of ROS which can disrupt insulin signaling. ROS will activate JNK (c-Jun N-terminal Kinases), which in turn decreases the tyrosine phosphorylation of IRS-1 and consequently inhibits insulin signaling and promotes insulin resistance [4]. Besides insulin resistance, ROS can also drive Cyclin D expression through $cJUN$ or activate the NFκB pathway, which plays an important role in tumorigenesis and insulin resistance.

RAGE-NFκB pathway: $\{ERK, RAGE, AKT, ROS\} \rightarrow IKK \dashv I\kappa B \dashv NF\kappa B \rightarrow \{A20, Bcl\text{-}XL, I\kappa B, CyclinD, AGE, Myc\}$. In the resting cell, NFκB is located in the cytoplasm, bound to and inhibited by IκB. Once activated by inflammatory mediators such as, the Receptors for Advanced Glycation End products (RAGEs) [24] or ROS, the IκB Kinase (IKK) will phosphorylate and deactivate IκB, leading to the translocation of NFκB into the nucleus to promote the transcription of a number of genes [24, 25], including Cyclin D, the anti-apoptotic protein Bcl-XL, cytokines AGEs, and its inhibitors A20 and IκB. The secreted AGEs can activate ROS and the NFκB pathway through an autocrine feedback loop [2]. The NFκB pathway is also implicated in the induction of insulin resistance.

Glucose-AMPK-mTOR pathway: $Glucose \rightarrow ATP \dashv AMPK \rightarrow TSC1/2 \dashv RHEB \rightarrow mTOR \rightarrow S6K \rightarrow CyclinD$. The proliferation of cancer cells requires glycolysis for energy, which depends on glucose. AMPK is a cellular metabolic sensor responding to low ATP levels; it is inhibited while ATP increases in hyperglycemia. AMPK regulates the metabolism of lipids, cholesterol and glucose in liver, muscle and adipose tissue, and links cell metabolism to cell growth control [26]. AMPK's activation will promote the transcription of the tumor-suppressor genes TSC1/2, which inhibit the activity of the oncoproteins RHEB and mTOR (mammalian Target Of Rapamycin). The activated mTOR will stimulate protein synthesis by phosphorylating ribosomal protein S6 kinase (S6K). Inappropriate activation of the mTOR-S6K pathway induces insulin resistance and hamartoma syndromes [27]. So, AMPK is a key therapeutic target for both diabetes and cancer.

WNT pathway: It includes the canonical pathway $WNT \rightarrow FZD \rightarrow DVL \dashv GSK3\beta \dashv \beta$-

Catenin $\rightarrow$ *TCF* $\rightarrow$ *CyclinD*, and the non-canonical pathway $WNT \rightarrow FZD \rightarrow DVL \rightarrow JNK$. WNT pathway activation was observed in 65% of pancreatic adenocarcinomas [28], and several pathway components are also overexpressed in Type 2 diabetes. When the WNT protein is absent, β-catenin is localized in the cytoplasm, bound to and inhibited by the complexes composed of Axin, APC, and GSK3β [29]. The canonical WNT pathway is activated by WNT/Frizzled (FZD) interaction, leading to a destabilization of the Axin-APC-GSK3 complex and translocation of β-catenin to the nucleus, where it activates the TCF-LEF transcription factors [30]. The activation of Dishevelled (DVL) [31] by the non-canonical WNT pathway could also promote insulin resistance [5] and cell growth through the JNK pathway.

Next, we will translate the above signaling pathways into to a discrete dynamic Boolean network model. The input signals of the model are the four different diabetes risk factors. Our goal is to investigate how these factors influence the development of cancer and diabetes. Therefore, the outputs of the model are cell proliferation, apoptosis, and insulin resistance. In a Boolean network, the state of each node can be either ON(1) or OFF(0) at any time step, and the state transformation of the node n from time t to $t+1$ is described by a Boolean *transfer* function. As in our recent work [18], the state of most internal nodes at the next time step depends on their current state and that of their immediate incoming neighbors, which can be divided as *activators* and *inhibitors*. The Boolean transfer function for node i can be written using the Boolean logic connectives $\vee$ (or), $\wedge$ (and), $\neg$ (not) as

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

where $A(n)$ and $I(n)$ represent the set of activators and inhibitors of node n, respectively. In Eq. 1 we assume that the activators can change the state of a node only if no inhibitor is acting on that node [18]. This assumption is based on the observation that, in the normal cell, the expression level of most oncoproteins is strictly regulated by the tumor-suppressors. However, in the cancerous stage the oncoproteins will be continuously activated following the loss of the tumor-suppressors inhibitors.

SYMBOLIC MODEL CHECKING

Model Checking [12] is an automatic verification technique for finite state systems modeled by labeled state transition graphs, called Kripke Structures. Let $AP = \{p, q, r, ...\}$ be a set of atomic propositions, where p, q ... are Boolean variables. The Boolean connectives are $\vee$, $\wedge$, $\neg$, and $\rightarrow$ (implication). Let S be a finite set of states, R be a transition relation between states, and $L(s)$ be a labeling function that labels each state s with the atomic propositions true in s. Given the Kripke structure $M = (S, R, L)$, and a temporal logic formula ϕ expressing some desired property/specification, the Model Checking problem [12] is to find the set of all states in S that satisfy ϕ: $\{s \in S \mid M, s \models \phi\}$. In Symbolic Model Checking [32], the transition relation between states is represented implicitly as a Boolean function, which is in turn encoded as an Ordered Binary Decision

Diagram (OBDD) [33], a data structure that can be used to efficiently represent and manipulate Boolean functions.

Temporal logic formulas can be divided into two subtypes, Linear Temporal Logic (LTL) formulas and Computation Tree Logic (CTL) formulas. An LTL formula describes properties of an infinite sequence of states, and is constructed from a set of atomic propositions, Boolean logic connectives, and *temporal* operators describing the properties of a path [12]: $\mathbf{X}p$ – p holds in the neXt state of the path; $\mathbf{F}p$ – p holds at some state in the Future (eventually) on the path; $\mathbf{G}p$ – p holds Globally (always) at every state on the path; $p\mathbf{U}q$ – p holds Until q holds on the path. A CTL formula describes properties of computation trees [12]. The root of a computation tree corresponds to the initial state, and the branches correspond to all possible paths from the root. A CTL formula is constructed from both *path* quantifiers and *temporal* operators. In a CTL formula, the LTL operators **X, F, G, U** must be immediately preceded by a path quantifier **A** – for *all* paths, or **E** – there *exists* a path. There are eight basic CTL operators, $AX, EX, AG, EG, AF, EF, AU, EU$. It has been proved that any CTL formula can be decomposed and expressed in terms of $\neg, \vee, EX, EU$ and EG [12].

CTL formulas can be divided into state formulas ψ and path formulas ϕ, and the syntax of the logic is the following:

$$\psi ::= AP \mid \psi_1 \vee \psi_2 \mid \neg\psi \mid \mathbf{E}\phi \mid \mathbf{A}\phi$$
$$\phi ::= \mathbf{X}\psi \mid \mathbf{F}\psi \mid \mathbf{G}\psi \mid \psi_1\mathbf{U}\psi_2.$$

A path in M is an infinite sequence of states, $\pi = s_0, s_1, ...$, where $s_i \in S$, and for every $i \geq 0$, $(s_i, s_{i+1}) \in R$, *i.e.* the pair (s_i, s_{i+1}) is a valid transition of the system M. We use π^i to denote the suffix of π starting at s_i. The semantics of a CTL formula is [12]:

$M, s \models p$	iff $p \in L(s)$;
$M, s \models \neg\psi$	iff $M, s \models \psi$ does not hold;
$M, s \models \psi_1 \vee \psi_2$	iff $M, s \models \psi_1$ or $M, s \models \psi_2$;
$M, \pi \models \mathbf{X}\psi$	iff $M, \pi^1 \models \psi$;
$M, \pi \models \psi_1\mathbf{U}\psi_2$	iff there exists $k \geq 0$ such that, $M, \pi^k \models \psi_2$ and for all $0 \leq j < k$, $M, \pi^j \models \psi_1$;
$M, s \models \mathbf{E}\phi$	iff there exists a path π from s such that $M, \pi \models \phi$;
$M, s \models \mathbf{A}\phi$	iff for every path π from s, $M, \pi \models \phi$.

The Symbolic Model Verifier (SMV) [32] developed by K. McMillan at Carnegie Mellon University is the first model checker based on OBDDs. The SMV language provides a notation to describe general state transition diagrams, and its output could be either "true" (property is satisfied) or a counterexample trace to the desired property. The Symbolic Model Checking algorithm is shown in Algorithm 1, the interested reader should refer to [12] for details.

Below, a portion of the SMV code for the diabetes-cancer Boolean network model is given as an illustration. In the SMV code, the keyword **VAR** is used to declare variables; **ASSIGN** is used to define the initial states (**init**) and state transitions of the model (**next**). The verification of CTL properties is encoded using the "**SPEC**" statement.

Algorithm 1 Symbolic Model Checking Algorithm

Input: A model $M = (S, R, L)$, and CTL formulas ϕ, ϕ_1, ϕ_2.

Check: takes a CTL formula as its argument and returns the OBDD for the set of states that satisfy the formula.
Output: A set of states of M which satisfy the property ϕ: $M, s \models \phi$.

- if ϕ is an atomic proposition p: return **Check**(p) = p.
- if $\neg\phi$: return **Check**($\neg\phi$) = $\neg$**Check**(ϕ)
- if $\phi_1 \vee \phi_2$: return **Check**(ϕ_1)$\vee$ **Check**(ϕ_2)
- if **EX**ϕ: return **Check**(**EX**(**Check**(ϕ)))
- if **E**[ϕ_1**U**ϕ_2]: return **Check**(**EU**(**Check**(ϕ_1),**Check**(ϕ_2)))
- if **EG**ϕ: return **Check**(**EG**(**Check**(ϕ)))

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

MODEL VERIFICATION

The proposed Diabetes-Cancer Boolean network depicted in Fig. 1 comprises 49 variable nodes and 4 external (control) nodes, leading to 2^{49} possible states in the state-transition diagram. In this work, SMV is applied to answer some questions that were proposed in [1]. For example: "Are there any associations between diabetes and cancer incidence or prognosis", "What components are critical to both diabetes and cancer", and "Do diabetes treatments influence the risk of cancer". The SMV code for our models is available online: http://www.cs.cmu.edu/~haijung/research/DC11.zip. Our results are divided into two groups: group A – *normal* cells in which the initial state of each node in the BN network can be either active (1) or inactive (0); and group B – *precancerous* or *cancerous* cells which are characterized by loss of the tumor-suppressor proteins ARF and INK4a, *i.e.*, the states of both INK4a and ARF are set to be 'OFF' in the model.

Question 1: Do diabetes risk factors influence the risk of cancer or cancer prognosis? The CTL properties for this question are summarized as:

$$\textit{Property}\ 1: \mathbf{AF}(\textit{Proliferate}); \quad \textit{Property}\ 1': \mathbf{EF}(\textit{Proliferate});$$

$$\textit{Property}\ 2: \mathbf{AF}(\textit{Apoptosis}); \quad \textit{Property}\ 2': \mathbf{EF}(\textit{Apoptosis});$$

$$\textit{Property}\ 3: \mathbf{AF}(\textit{Resistance}); \quad \textit{Property}\ 3': \mathbf{EF}(\textit{Resistance});$$

Properties 1-3 (1'-3') mean that, under the influence of some diabetes risk factors, it is **necessary (possible)** for the cell to reach the *Proliferate/Apoptosis/Insulin-Resistance* state at some time in the future. If the property $\mathbf{AF}p$ is true, then $\mathbf{EF}p$ will also be true (of course the converse implication does not hold in general). These properties were verified by SMV with different risk factors including Hyperinsulinaemia, Obesity, Hyperglycemia and WNT:

(a) *Group A – normal cells*: under the influence of these diabetes risk factors, Properties 3 and 2'-3' are true, while the rest are false. That is, diabetes risk factors would augment insulin resistance in diabetes patients, but cell growth is still regulated by the tumor suppressor proteins; apoptosis is possible, while cell proliferation is not permitted. Thus, the diabetes risk factors might not increase the risk of cancer under normal conditions. Property 3 is consistent with the diabetes studies showing that increased adiposity and active WNT pathway are major determinants of insulin resistance [4, 5].

(b) *Group B – precancerous or cancerous cells*: If either of these risk factors is 'ON', our verification results demonstrate that all but Property 2 are true. This means that under the influence of diabetes risk factors, cancer cell proliferation and insulin resistance are unavoidable, while the failure of Property 2 indicates that apoptosis does not necessarily happen. That is, diabetes risk factors promote the growth of precancerous or cancerous cells, augment insulin resistance and worsen the diabetic complications when the loss of the tumor-suppressors INK4a and ARF occurs. These results are consistent with a number of large-scale population-based cohort analyses [8, 9, 10].

Another important question to ask is whether diabetes treatments influence the risk of cancer. Since insulin injections are needed for both Type 1 and 2 diabetes treatment, they could lead to an increase of insulin level in a short time (hyperinsulinemia). There is evidence that insulin therapy could increase the risk of colorectal cancer in Type 2 diabetes patients [34]. Our verification results suggest that insulin injection treatment might not be the cause of cancer incidence in non-cancerous diabetes patients. However, if the diabetes patient is in precancerous or cancerous stage, it could promote tumorigenesis.

Question 2: What signaling components are common and critical to both diabetes and cancer? That is, which proteins' mutation/knockout will promote/inhibit cancer cell growth and insulin resistance in diabetic cancer patients? In temporal logic language, the following CTL formulas in Group B were verified to be true:

$$\textit{Property}\ 4: \mathbf{AG}\{\textit{RAS} \rightarrow \mathbf{AF}(\textit{Resistance}\ \&\ \textit{Proliferate}\ \&\ !\textit{Apoptosis})\}$$

$$\textit{Property}\ 5: \mathbf{AG}\{\textit{AKT} \rightarrow \mathbf{AF}(\textit{Resistance}\ \&\ \textit{Proliferate}\ \&\ !\textit{Apoptosis})\}$$

$$\textit{Property}\ 6: \mathbf{AG}\{\textit{NF}\kappa\textit{B} \rightarrow \mathbf{AF}(\textit{Resistance}\ \&\ \textit{Proliferate}\ \&\ !\textit{Apoptosis})\}$$

$$\textit{Property}\ 7: \mathbf{AG}\{\textit{ROS} \rightarrow \mathbf{AF}(\textit{Resistance}\ \&\ \textit{Proliferate}\ \&\ !\textit{Apoptosis})\}$$

Properties 4-7 mean that regardless of presence of diabetes risk factors, the continuous activation or overexpression of RAS, AKT, NFκB, or ROS will induce the proliferation of precancerous or cancerous cell, inhibit apoptosis and augment insulin resistance. In other words, cell proliferation and insulin resistance are unavoidable. Property 7 is consistent with the experimental results in diabetes and cancer studies showing that the production of ROS can not only disrupt insulin signaling and induce insulin resistance [35], but also cause oxidative damage to DNA, thus promoting tumorigenesis [36].

Question 3: The oscillations of NFκB [37] and the negative feedback of P53-MDM2 have been measured in many *in vitro* experiments, after the cells were stimulated by external signals. Do these phenomena exist in cells subjected to diabetic risk factors?

$$Property\ 8 : \mathbf{AG}\{(!NF\kappa B \rightarrow \mathbf{AF}(NF\kappa B)) \ \&\ (NF\kappa B \rightarrow \mathbf{AF}(!NF\kappa B))\}$$

$$Property\ 9 : \mathbf{AG}\{(P53 \rightarrow \mathbf{AF}(MDM2)) \ \&\ (MDM2 \rightarrow \mathbf{AF}(!P53))\}$$

Properties 8 and 9 were determined to be true when all the risk factors are 'ON', meaning that the oscillations of NFκB and P53-MDM2 also exist in cells influenced by different diabetes risk factors. These two properties could be tested by future experiments.

CONCLUSIONS

Our diabetes-cancer signaling network model is the first attempt to investigate the association of diabetes and cancer, based on the known signaling pathways at the molecular level. We have applied Model Checking to formally verify some temporal logic properties of our model. Though it is still debatable whether diabetes increases the risk of cancer or not, our work provides some insights into the association of diabetes with cancer. For example, our verification results indicate that the diabetes risk factors might not increase the risk of cancer in normal cells. However, they could promote cell proliferation, inhibit apoptosis and augment insulin resistance, after losses of the tumor-suppressor proteins ARF and INK4a occur in the precancerous stage. We have also studied some signaling components, including RAS, AKT, NFκB, and ROS, which are common and critical to the development of both diabetes and cancer. The CTL properties proposed and verified in this work might provide new insights into cancer and diabetes therapies. In addition, the diabetes-cancer model predicts oscillation phenomena which might be observed in diabetes studies. Our work demonstrates that Model Checking and the diabetes-cancer signaling network model can capture the most important characteristics of the pathways implicated in the interaction between diabetes and cancer.

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. E. Giovannucci, D. M. Harlan, M. C. Archer, et al., *CA Cancer J Clin* **60**, 207–221 (2010).
2. J. L. Evans, I. D. Goldfine, B. A. Maddux, et al., *Endocrine Reviews* **23**, 599–622 (2002).
3. G. Bell, and K. Polonsky, *Nature* **414**, 788–791 (2001).
4. I. F. Godsland, *Clinical Science* **118**, 315–332 (2010).
5. S. H. Lee, C. Demeterco, I. Geron, et al., *Experimental Diabetes Research* **2008**, 728763 (2008).
6. D. Hanahan, and R. A. Weinberg, *Cell* **100**, 57–70 (2000).
7. W. Fisher, *World J. Surg.* **25**, 503–508 (2001).
8. C. Tseng, *Diabetes Care* **34**, 616–21 (2011).
9. U. Smith, and E. Gale, *Diabetologia* **52**, 1699–1708 (2009).
10. I. Wolf, S. Sadetzki, R. Catane, A. Karasik, and B. Kaufman, *Lancet Oncol* **6**, 103–111 (2005).
11. F. Wang, M. Herrington, J. Larsson, and J. Permert, *Molecular Cancer* **2**, 4 (2003).
12. E. M. Clarke, O. Grumberg, and D. A. Peled, *Model Checking*, MIT Press, 1999.
13. C. Langmead, and S. Jha, *J. of Bioinformatics and Computational Biology* **7**, 323–338 (2009).
14. F. Fages, S. Soliman, and N. Chabrier-Rivier, *J. of Biol. Phys. and Chem.* **4**, 64–73 (2004).
15. X. Shen, J. Collier, D. Dill, et al., *PNAS* **105**, 11340–11345 (2008).
16. H. Gong, P. Zuliani, A. Komuravelli, J. R. Faeder, and E. M. Clarke, *Proceedings of Algebraic and Numeric Biology, LNCS* **6479** (2010).
17. H. Gong, P. Zuliani, A. Komuravelli, J. Faeder, and E. M. Clarke, *BMC Bioinformatics* **11** (2010).
18. H. Gong, Q. Wang, P. Zuliani, M. Lotze, J. Faeder, and E. M. Clarke, *Proceedings of the International Conference on Bioinformatics and Computational Biology (BICoB)* (2011).
19. A. R. Sedaghat, A. Sherman, and M. J. Quon, *American Journal of Physiology-Endocrinology and Metabolism* **283**, E1084 (2002).
20. Y. H. Chew, Y. L. Shia, C. T. Lee, et al., *Molecular and Cellular Endocrinology* **303**, 13–24 (2009).
21. Y. Haupt, R. Maya, A. Kasaz, and M. Oren, *Nature* **387**, 296–299 (1997).
22. N. Bardeesy, and R. A. DePinho, *Nature Reviews Cancer* **2**, 897–909 (2002).
23. G. Yao, T. J. Lee, S. Mori, J. Nevins, and L. You, *Nature Cell Biology* **10**, 476–482 (2008).
24. J. R. van Beijnum, W. A. Buurman, and A. W. Griffioen, *Angiogenesis* **11**, 91–99 (2008).
25. A. Hoffmann, A. Levchenko, M. Scott, and D. Baltimore, *Science* **298**, 1241–1245 (2002).
26. D. G. Hardie, *International Journal of Obesity* **32**, S7–S12 (2008).
27. K. Inoki, M. Corradetti, and K. Guan, *Nature Genetics* **37**, 19–24 (2005).
28. G. Zeng, M. Germinaro, A. Micsenyi, et al., *Neoplasia* **8**, 279–289 (2006).
29. A. Wodarz, and R. Nusse, *Annu Rev Cell Dev Biol* **14**, 59–88 (1998).
30. B. Vogelstein, and K. Kinzler, *Nature Medicine* **10**, 789–799 (2004).
31. J. Wallingford, and R. Habas, *Development* **132**, 4421–36 (2005).
32. K. L. McMillan, *PhD thesis: Symbolic model checking - an approach to the state explosion problem*, Carnegie Mellon University, 1992.
33. R. Bryant, *IEEE Transactions on Computers* **35**, 677–691 (1986).
34. Y. Yang, S. Hennessy, and J. Lewis, *Gastroenterology* **127**, 1044–1050 (2004).
35. G. King, and M. Loeken, *Histochem. Cell. Biol.* **122**, 333–338 (2004).
36. S. Martien, and C. Abbadie, *Ann. N.Y. Acad. Sci.* **1119**, 51–63 (2007).
37. S. Kauffman, *J. Theoret. Biol* **22**, 437–467 (1969).