MATHEMATICAL MODELING AND INFERENCE OF CANCER NETWORKS

A Dissertation

by

HASWANTH VUNDAVILLI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Aniruddha Datta
Committee Members,	Shankar P. Bhattacharyya
	Yang Shen
	Catherine Yan
Head of Department,	Miroslav M. Begovic

May 2021

Major Subject: Electrical Engineering

ABSTRACT

Cancer is a group of diseases characterized by abnormal cell growth. Old cells do not die and grow uncontrollably, forming a mass of tissue, called a tumor. In order to understand this abnormal cell growth, there have been various efforts to model the interactions between different molecules and pathways that initiate and drive cell proliferation. In this work, we analyze Bayesian and Boolean techniques that can aid in modeling different cancer networks and infer the drug combinations that can effectively kill tumor cells.

Signaling pathways supervise cellular processes such as growth, differentiation, and death. In healthy cells, these processes are tightly regulated, however, in cancerous cells, mutations in crucial genes often lead to irregularities in these processes and eventually cancer. In this work, we study pathways and genes characterizing Breast cancer, Pancreatic cancer, and Lung cancer. We make use of biological literature to construct the pathways and then use mathematical modeling techniques to analyze and rank different therapeutic interventions. We first develop a Bayesian network of Breast cancer and using a messaging passing algorithm, we infer the network and rank drugs according to their ability to induce apoptosis. We then model the signaling network and mutations of Pancreatic cancer using a multi-fault Boolean framework and simulate the network to theoretically assess the efficacy of drug combinations. Finally, we use a modified Boolean approach to mathematically model feedback loops in Lung cancer and determine the drug combinations that produce cell death for the majority of mutations.

Our theoretical analyses point out that drug combinations containing Cryptotanshinone, a compound found in traditional Chinese herbs, result in significantly increased cell death in each of Breast, Pancreatic, and Lung cancer pathways. We corroborated our theoretical results with experiments on MCF-7 breast cancer cell lines, Human Pancreatic Cancer (HPAC) cell lines, H2073 and SW900 lung cancer cell lines.

DEDICATION

To my mother, my father, and my twin brother.

ACKNOWLEDGMENTS

I count myself very fortunate to have been advised by Professor Aniruddha Datta. His unwavering support, invaluable mentoring, and practical advice have made my doctoral adventure truly a smooth journey. I am forever grateful to him for consistently supporting me on this journey. I am also thankful to Professor Shankar Bhattacharyya, Professor Yang Shen, and Professor Catherine Yan for their feedback, discussions, and serving on my dissertation committee.

I would also like to thank Professor Sampsa Hautaniemi and his team and Professor Kenji Mizuguchi and his team for hosting me at the University of Helsinki, Finland and at the National Institutes of Biomedical Innovation, Health and Nutrition, Japan respectively. I am very thankful for the memorable and life-changing experiences I picked up there.

Most importantly, without my awe-inspiring mother, ever-encouraging father, and a mix of cool and nerdy twin-brother, this PhD dissertation would not have taken any shape or form whatsoever.

I would also like to thank all my friends and colleagues for their continued support and unforgettable memories. Finally, this dissertation is incomplete without acknowledging my best friend Hari who has always helped me and been my well-wisher.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Aniruddha Datta (advisor), Professor Shankar Bhattacharyya and Professor Yang Shen of the Department of Electrical & Computer Engineering and Professor Catherine Yan of the Department of Mathematics.

The experiments on cell lines in Sections 2, 3, and 4 were carried out by Dr. Chao Sima, Dr. Jianping Hua, Dr. Rosana Lopes, and Dr. Michael Bittner of TEES-AgriLife Center for Bioinformatics and Genomic Systems Engineering (CBGSE) at Texas A&M University.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

This work was funded in part by the National Science Foundation under grants ECCS-1404314, ECCS-1609236 and ECCS-1917166, and in part by TEES-AgriLife Center for Bioinformatics and Genomic Systems Engineering (CBGSE) startup funds.

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION *

Breast Cancer is the second leading cause of cancer death among US women, hence identifying potential drug targets is an ever increasing need. In Section 2, we integrate existing biological information with graphical models to deduce the significant nodes in the Breast Cancer signaling pathway. We make use of biological information from the literature to develop a Bayesian network. Using the relevant gene expression data we estimate the parameters of this network. Then, using a messaging passing algorithm, we infer the network. The inferred network is used to quantitatively rank different interventions for achieving a desired phenotypic outcome. The particular phenotype considered here is the induction of apoptosis. The theoretical analysis pinpoints to the role of Cryptotanshinone, a compound found in traditional Chinese herbs, as a potent modulator for bringing about cell death in the treatment of cancer. Using a mathematical framework, we showed that the combination therapy of mTOR and STAT3 genes yields the best apoptosis in Breast Cancer. The computational results we arrived at are consistent with the experimental results that we obtained using Cryptotanshinone on MCF-7 breast cancer cell lines and also by the past results of others from the literature, thereby demonstrating the effectiveness of our model.

The number of deaths associated with Pancreatic Cancer has been on the rise in the United States making it an especially dreaded disease. The overall prognosis for pancreatic cancer patients continues to be grim because of the complexity of the disease at the molecular level involving the potential activation/inactivation of several diverse signaling pathways. In Section 3, we first model the aberrant signaling in pancreatic cancer using a multi-fault Boolean Network. Thereafter, we theoretically evaluate the efficacy of different drug combinations by simulating this

*Parts of this section are reprinted with permission from H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "Bayesian Inference Identifies Combination Therapeutic Targets in Breast Cancer," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 9, pp. 2684-2692, 2019, doi: 10.1109/TBME.2019.2894980 © 2019 IEEE; and H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "In Silico Design and Experimental Validation of Combination Therapy for Pancreatic Cancer," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 3, pp. 1010-1018, 2020, doi: 10.1109/TCBB.2018.2872573 © 2020 IEEE; and H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "Cryptotanshinone Induces Cell Death in Lung Cancer by Targeting Aberrant Feedback Loops," *in IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 8, pp. 2430-2438, 2020, doi: 10.1109/JBHI.2019.2958042 © 2020 IEEE.

boolean network with drugs at the relevant intervention points and arrive at the most effective drug(s) to achieve cell death. The simulation results indicate that drug combinations containing Cryptotanshinone, a traditional Chinese herb derivative, result in considerably enhanced cell death. These in silico results are validated using wet lab experiments we carried out on Human Pancreatic Cancer (HPAC) cell lines.

Signaling pathways oversee highly efficient cellular mechanisms such as growth, division, and death. These processes are controlled by robust negative feedback loops that inhibit receptor-mediated growth factor pathways. Specifically, the ERK, the AKT, and the S6K feedback loops in Lung Cancer attenuate signaling via growth factor receptors and other kinase receptors to regulate cell growth. Irregularity in any of these supervised processes can lead to uncontrolled cell proliferation and possibly cancer. These irregularities primarily occur as mutated genes, and an exhaustive search of the perfect drug combination by performing experiments can be both costly and complex. Hence, in Section 4, we model the Lung Cancer pathway as a Modified Boolean Network that incorporates feedback. By simulating this network, we theoretically predict the drug combinations that achieve the desired goal for the majority of mutations. Our theoretical analysis identifies Cryptotanshinone, a traditional Chinese herb derivative, as a potent drug component in the fight against cancer. We validated these theoretical results using multiple wet lab experiments carried out on H2073 and SW900 lung cancer cell lines.

2. BAYESIAN INFERENCE IDENTIFIES COMBINATION THERAPEUTIC TARGETS IN BREAST CANCER *

2.1 Introduction

Cancer is a collection of diseases characterized by out-of-control cell growth. Cancer develops when the body's normal control system stops working. Old cells do not die and grow out of control, forming new, aberrant cells. These extra cells may form a mass of tissue, called a tumor. Despite major progress over the last 40 years, it is estimated that in 2017 alone over 250,000 women will be diagnosed with breast cancer and more than 40,000 will die just in the United States [4]. This calls for safe, contemporary and effective tools in the battle against breast cancer. One way to approach several problems in Medicine and other life sciences is as a control problem, with the objective being to find methods to drive an undesirable state of a Gene Regulatory Network (GRN) into another, more desirable one, by means of an intervention, such as a therapeutic treatment. The rationale for this is that if we can build a good model and find the targets with the most "significance", we may be able to design drugs for diseases, such as cancer, resulting from gene misregulations. In this vein, there have been numerous attempts at modeling genetic regulatory networks, such as Boolean networks [5, 6], Differential equations [7], Probabilistic Boolean networks [8] and Bayesian networks [9, 10, 11].

In this work, we describe a methodology that utilizes current biological knowledge from the literature to build a GRN model and integrates this knowledge with experimental Genomic data using a Bayesian Network based approach. Bayesian Networks are a class of Directed Acyclic Graphs(DAGs) that encode independencies in a given network. They are suitable to the problem since they can be used to represent causal relationships, analogous to the interactions in biological signaling pathways. In the given network, our interest is in selecting genes which can be used as modulators in order to drive the pathway dynamics to a desirable state. As the network grows in

*Parts of this section are reprinted with permission from H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "Bayesian Inference Identifies Combination Therapeutic Targets in Breast Cancer," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 9, pp. 2684-2692, 2019, doi: 10.1109/TBME.2019.2894980

size, this modulator selection problem becomes intractable. Fortunately, we have a polynomial-time algorithm called Pearl's Message-Passing Algorithm for performing modulator selection in Bayesian networks. We apply this algorithm to the Breast Cancer pathway in the human body to derive effective drug targets related to breast cancer.

2.2 Methodology

2.2.1 Bayesian Networks

Bayesian networks (BNs), also known as belief networks, belong to the family of probabilistic graphical models (PGMs). In the last decade, BNs have become extremely popular and have been used for applications in various areas, such as machine learning [12], speech recognition [13], bioinformatics [9], plant genomics [14], etc.

Interactions in gene regulatory networks are usually sparse. i.e., each gene interacts with a very small number of genes compared to the total number of genes in the network. Due to this sparseness, the generated BN models are easy to categorize, which promises well for both reconstruction and visualization scalability, and this makes BN models a promising tool for the analysis of gene regulatory networks [15, 16]. Also, biological systems are naturally stochastic, and the probabilistic nature of Bayesian networks is well suited for capturing the uncertainties involved in gene networks.

We represent a BN by $\langle \mathcal{G}, \theta \rangle$, where $\mathcal{G}$ is the DAG over the set of random variables $\boldsymbol{X}$. The nodes $X_1, X_2, ..., X_n$ in $\mathcal{G}$ represent random variables, and the edges in $\mathcal{G}$ represent the direct dependencies between them. In a BN, each variable is independent of its nondescendents in the graph given the state of its parents, which is consistent with a Markovian property, where the CPD at each node depends only on its parents. The second component θ denotes the set of parameters of the network. Accordingly, $\mathcal{G}$ defines a unique joint probability distribution(JPD) over $\boldsymbol{X}$ given by:

$$P(X_1, X_2, ..., X_n) = \prod P(X_i | Pa(X_i)) \tag{2.1}$$

where $Pa(X_i)$ are the parents of X_i.

In a biological application, we can use the signaling pathway knowledge available in the biological literature to construct the graph skeleton according to which the JPD of our model gets factorized.

2.2.2 Integrating Gene Expression Data

Once we obtain the graph structure $\mathcal{G}$ of the Bayesian network, we proceed to update the model parameters using gene expression data. Gene expression data obtained from public repositories such as NCBI usually consists of expressions of key genes measured across different experiments, and results in a matrix with raw data as shown in Table 2.1.

Table 2.1: Example gene expression table. (Reprinted from [1])

	Exp_1	Exp_2	...	Exp_m
$Gene_1$	0.34	0.49	...	0.05
$Gene_2$	0.86	0.62	...	0.35
$\vdots$	$\vdots$		$\ddots$	
$Gene_n$	0.44	0.29	...	0.87

We use a binary quantization of gene expression data. There are several advantages to doing so such as enhanced noise robustness and reduction of computational complexity. Gene expression levels are heavily skewed in linear scale. Hence, using a standard threshold for discretizing multiple genes would lead to biasing error. Therefore, for each gene, assuming the gene expression data is normally distributed, we used a maximum likelihood estimator for the mean μ, and used it as the threshold to discretize the expression data for that specific gene. Expression values above the threshold are assigned a 1 value and those below are assigned a 0 value. The choice of this threshold is also justified using the Central Limit Theorem and the Law of large numbers. We next integrate the gene expression data with the Bayesian model.

Hoff [17] showed that whenever we have a conjugate prior, the posterior distribution belongs to the same family of distributions as the prior. The beta distribution is known to be a conjugate

to the binomial likelihood, hence, if we choose the prior to be a Beta distribution, the conditional posterior probability distributions of nodes are again given by a beta distribution, i.e.

$$P(p|X) \sim Beta(\bar{\alpha}_X, \bar{\beta}_X) \tag{2.2}$$

where $\bar{\alpha}_X = (\alpha_X + m)$, $\bar{\beta}_X = (\beta_X + n - m)$, n is the number of observations, m is the number of 1's in the data of X and α_X, β_X are the shape parameters of the Beta distribution.
The expected value for this Beta distribution is given by,

$$E(p|X) = \frac{\bar{\alpha}_X}{\bar{\alpha}_X + \bar{\beta}_X} \tag{2.3}$$

Initially, we assign a prior for all nodes as Beta(1,1), which is a uniform distribution over the finite interval [0,1]. Using Equation (2.2), we update the prior using the gene expression data to obtain the posterior distribution. With the posterior distribution in hand, we can calculate the posterior mean which is the expected value of each state of a target node conditioned on each possible combination of states of its parent nodes. As more data are observed, we can update the values of $\bar{\alpha}_X$ and $\bar{\beta}_X$ so that the posterior probabilities approach the true underlying distribution. When this occurs, the expected values converge to the CPDs of our Bayesian network. A simple algorithm implementing the above procedure of integrating gene expression data to calculate the CPDs is given in Algorithm 1 below.

With the Bayesian network graph $\mathcal{G}$ and its CPDs calculated, we want to identify the most significant nodes, which when used as modulators will drive our network to the desired output.

2.2.3 Significant Genes

Biologically, a gene with a stronger effect on a target (gene or output) has a higher chance of desirably affecting the latter as compared to other genes that have only minimal effects. Not all genes have the same influence in a network and hence identifying the gene(s) which have high "significance" is of utmost importance while trying to carry out drug design. For example, p53 is a

1 Calculating the CPDs of a Bayesian Network

procedure PROBABILITIES($\mathcal{G}, S$) ▷ Input graph matrix $\mathcal{G}$ and dataset S
  for each column in S **do**
   find the mean
    for each row in the column **do**
      if value $<$ mean **then**
       value $\leftarrow 0$
      else
       value $\leftarrow 1$
      end if
    end for
  end for

  for each node in $\mathcal{G}$ **do**
   find its parents and store them
   the number of conditional probabilities $= 2^{(\text{number of parents})}$
    for each data-point of the node **do**
     compare with the data-points of its parents
      if match **then** count $\leftarrow$ count + 1
      using count calculate $\bar{\alpha}_X$ and $\bar{\beta}_X$ from (2.2)
      end if
     $P(X_i|Pa(X_i)) = \frac{\bar{\alpha}_X}{\bar{\alpha}_X + \bar{\beta}_X}$
    end for
  end for
end procedure

well known tumor suppressor gene, whereas p73 another tumor suppressor gene is less productive in activating apoptosis compared to p53 [18].

Given the bayesian network and its model parameters, we are interested in finding the nodes that contribute the most to our desired output. In other words, we want to maximize the conditional probability of our output when other genes are intervened. We illustrate this further by modeling drug intervention in an example bayesian network.

2.2.4 Drug Intervention

Drugs generally work by interacting with receptors on the surface of cells or enzymes within cells. Binding to the target receptor site, they can either block the function of the protein (inhibitory drug), or imitate it's effect (enhancing drug).

Suppose we have two drugs Drug-1 and Drug-2 that bind to the receptors A and B respectively as shown in Figure 2.1, thereby disabling their functions further downstream. Mathematically, we model this by setting the probability of that node having the value one equal to zero. That is, in the case of Drug-1, we have, $P(A = 1) = 0$ and similarly for Drug-2, we have, $P(B = 1) = 0$.

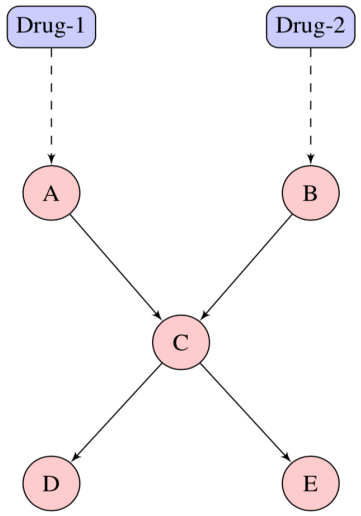


Figure 2.1: Modeling the drug intervention in an example bayesian network. (Reprinted from [1])

Let us assume that our desired objective is to minimize the gene expression of E. Given the two drug choices, we would like to find the more efficient drug for achieving this. In other words, we are interested in determining the smaller of $P(E = 1|A = 0)$ and $P(E = 1|B = 0)$. If our computations result in $P(E = 1|A = 0)$ to be the smaller of the two, we conclude Drug-1 to be the more effective one.

To calculate $P(E|A)$ we compute:

$$P(E|A) = \frac{P(E, A)}{P(A)} = \frac{\sum_C P(E, A, C)}{P(A)} = \frac{\sum_C P(E|A, C)P(A, C)}{P(A)}$$

In a Bayesian network each variable is conditionally independent of all its non-descendants in the

graph given the value of all its parents. Hence, $P(E|A,C) = P(E|C)$. Substituting this into the last expression and proceeding, we obtain:

$$\frac{\sum_C P(E|A,C)P(A,C)}{P(A)} = \frac{\sum_C P(E|C)P(A,C)}{P(A)} = \frac{\sum_C P(E|C)\big(\sum_B P(A,B,C)\big)}{P(A)}$$

$$= \frac{\sum_C P(E|C)\big(\sum_B P(C|A,B)P(A,B)\big)}{P(A)} = \frac{\sum_C P(E|C)\big(\sum_B P(C|A,B)P(A)P(B)\big)\big)}{P(A)}$$

Therefore, finally we obtain

$$P(E|A) = \sum_C P(E|C)\Big(\sum_B P(C|A,B)P(B)\Big) \tag{2.4}$$

Similarly $P(E|B)$ is given by:

$$P(E|B) = \sum_C P(E|C)\Big(\sum_A P(C|A,B)P(A)\Big) \tag{2.5}$$

Using the CPDs calculated earlier, we can compute (2.4), (2.5) and compare the two numbers, based on which we can make a decision about whether to intervene with A or with B. Clearly, as the network grows in size, computing these probabilities manually becomes tedious and intractable. Shimony [19] showed that the probabilistic reasoning problems for general Bayesian networks are NP-hard. Fortunately, there are algorithms for special cases that solve reasoning problems in time that is a polynomial function of the number of variables. Message passing algorithms such as Pearl's message-passing propagation algorithm efficiently solve the inference problem in singly-connected networks [20, 21]. We now briefly discuss the working of this algorithm.

2.2.5 Message-Passing Algorithm

Shimony [19] showed that the probabilistic reasoning problems for general Bayesian networks are NP-hard. However, there are algorithms for special cases that solve reasoning problems in time that is a polynomial function of the number of variables. In this section we review a polynomial-

time algorithm called Pearl's Message-Passing Algorithm, which is also known as Pearl's belief propagation algorithm. Pearl's algorithm provides exact solutions to inference problems in singly-connected graphs [22]. The algorithm exploits the loop-less feature in a singly-connected graph that allows us to partition the graph in an efficient manner, greatly simplifying the reasoning problem.

Consider a Bayesian Network over a set of nodes $\mathcal{X}$ with a DAG $\mathcal{G}$ that is a singly-connected directed graph. Let $\mathbf{E} \in \mathcal{X}$ be a set of evidence variables such that $\mathbf{E}$ = e. For all random variables $X \in \mathcal{X}$, we define λ message, λ value, π message and π value as the following [21, 23]:

- λ message: The message a child Y passes to its parent X. It is denoted by $\lambda_y(x)$.
- π message: The message a parent Z passes to its child X. It is denoted by $\pi_z(x)$.
- λ and π values: Each node has values of λ and π for each state that it may pass on.

An illustration of λ and π messages being exchanged in a graph is shown in Figure 2.2 below. These messages and values are defined as follows.

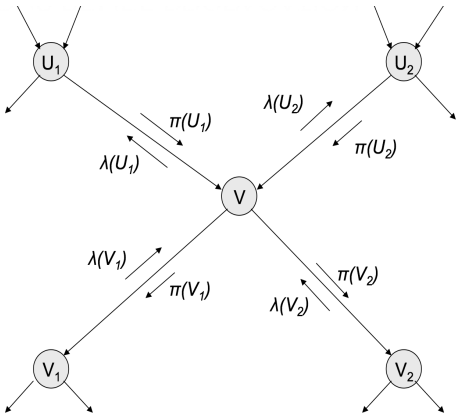


Figure 2.2: λ messages and π messages in a singly-connected directed graph. (Reprinted from [1])

1. **λ messages (child $\rightarrow$ parent) :** For every child $Y \in Ch(X)$ and all $x \in Val(X)$, define

$$\lambda_Y(x) \equiv \sum_y \left[\sum_{w_1,...,w_k} \left(P(y|x, w_1, ..., w_k) \prod_{i=1}^{k} \pi_Y(w_i) \right) \right] \lambda(y) \tag{2.6}$$

where $W_1, ..., W_k$ are other parents of **Y**.

2. **π messages (parent $\rightarrow$ child) :** For a parent $Z \in Pa(X)$ and all $z \in Val(Z)$, define

$$\pi_X(z) \equiv \pi(z) \prod_{U \in Ch(Z) - \{X\}} \lambda_U(z) \tag{2.7}$$

3. **λ values :**

(i) If $X \in \mathbf{E}$ and the observed values is $X = \hat{x}$, for all $x \in Val(X)$, define

$$\lambda(x) \equiv \begin{cases} 1 & \text{if } x = \hat{x} \\ 0 & \text{else} \end{cases} \tag{2.8}$$

(ii) If $X \notin \mathbf{E}$ and X is a leaf, for all $x \in Val(X)$, define $\lambda(x) \equiv 1$

(iii) If $X \notin \mathbf{X}$ and X is not a leaf, for all $x \in Val(X)$, define

$$\pi_X(z) \equiv \pi(z) \prod_{U \in Ch(Z) - \{X\}} \lambda_U(z) \tag{2.9}$$

4. **π values :**

(i) If $X \in \mathbf{E}$ and the observed value is $X = \hat{x}$, for all $x \in Val(X)$, define

$$\pi(x) \equiv \begin{cases} 1 & \text{if } x = \hat{x} \\ 0 & \text{else} \end{cases}$$

(ii) If $X \notin \mathbf{E}$ and X is a root, for all $x \in Val(X)$, define $\pi(x) \equiv P(x)$

(iii) If $X \notin \mathbf{E}$ and X is not a root, for all $x \in Val(X)$, define

$$\pi(x) \equiv \sum_{z_1,...z_k} \left(P(x|z_1, ...z_k) \prod_{i=1}^{k} \pi_X(z_i) \right) \tag{2.10}$$

where $Z_1, ...Z_k \in Pa(X)$ are the parents of **X**.

5. If we define the messages and values as above, for all $X \in \mathcal{X}$, $x \in Val(X)$ we can calculate

our required conditional probability given evidence **e** by :

$$P(x|\mathbf{e}) = \alpha\lambda(x)\pi(x) \qquad \text{(where } \alpha \text{ is a normalizing constant)}$$

We can compute $\lambda(x)$, $\pi(x)$ and $P(x|\mathbf{e})$ for every random variable $X \in \mathcal{X}$ and for all values $x \in Val(X)$ using Algorithm 2 given below.

2 Message-Passing algorithm

function INITIALIZE NETWORK
  for $X \in \mathcal{X}$ set $\lambda(x) = 1$
  $\forall Z \in Pa(X)$ set $\lambda_X(z) = 1$, $\forall Y \in Ch(X)$ set $\pi_Y(x) = 1$
  for every Root $R \in \mathcal{X}$ set $\pi(x) = P(x)$, $P(r|\mathbf{e}) = P(r)$
  for $W \in Ch(R)$ send_π_message(R, W)
end function

function UPDATE NETWORK ▷ New evidence $V = \hat{v}$
  $\mathbf{E} = \mathbf{E} \cup \{V\}$
  for $v \in Val(V)$, if $v = \hat{v}$ set $\lambda(v) = \pi(v) = P(v|\mathbf{e}) = 1$, else set $= 0$
  $\forall\big(Z \in Pa(V)$ && $Z \notin \mathbf{E}\big)$ send_λ_message(V, Z)
  for $Y \in Ch(V)$ send_π_message(V, Y)
end function

function SEND_λ_MESSAGE(Y, X) ▷ λ message Y(child) $\rightarrow$ X(parent)
  Compute $\lambda_Y(x)$ using equation (2.6)
  $\lambda(x) = \prod_{U \in Ch(X)} \lambda_U(x)$, $\tilde{P}(x) = \lambda(x)\pi(x)$
  $\alpha = \sum_x \tilde{P}(x) \Rightarrow P(x|\mathbf{e}) = \frac{1}{\alpha}\tilde{P}(x)$
  $\forall\big(Z \in Pa(X)$ && $Z \notin \mathbf{E}\big)$ send_λ_message(X, Z)
  $\forall\big(U \in Ch(X) - \{Y\}\big)$ send_π_message(X, U)
end function

function SEND_π_MESSAGE(Z, X) ▷ π message Z(parent) $\rightarrow$ X(child)
  Compute $\pi_X(z)$ using equation (2.7)
  if $X \notin \mathbf{E}$ **then**
    compute $\pi(x)$ using equation (2.10)
    $\tilde{P}(x) = \lambda(x)\pi(x) \Rightarrow P(x|\mathbf{e}) = \frac{1}{\alpha}\tilde{P}(x)$ where $\alpha = \sum_x \tilde{P}(x)$
    for $Y \in Ch(X)$ send_π_message(X, Y)
  end if
  if $\exists\ x$ such that $\lambda(x) \neq 1$ **then**
  for $\big(W \in Pa(X) - \{Z\}$ && $W \notin \mathbf{E}\big)$ send_λ_message(X, W)
  end if
end function

2.3 Breast Cancer Pathways

The Breast cancer pathway mainly consists of three important sub-pathways, the JAK/STAT, the MAPK, and the PI3K/mTOR which all interact with each other.

The Janus kinase (JAK)/signal transducer and activator of transcription (STAT) cell signaling pathway functions as the primary component of gene transcription and immune control. Abnormal activation of the JAK/STAT pathway has been reported in various disease states [24], and in solid tumors, constant phosphorylation of STAT3 has been demonstrated in breast cancer. JAKs employ Cytokine receptors (CRLF2) and mediate tyrosine phosphorylation of STAT3 [25].

The mammalian target of rapamycin (mTOR), lies downstream of receptor phosphoinositide 3 kinase (PIK3CA) [26], and PI3K/Akt/mTOR is a well known pathway, which causes cell growth and tumor proliferation in breast cancer [27]. Upregulated mTOR activates downstream ribosomal p70S6 kinase (RPS6KB1) and hyperphosphorylates Eukaryotic Translation Initiation Factor 4E (eIF4E) [28]. Hussain [29] showed a compelling biological connection between NF-κB and PI3K/Akt pathway, where Akt activates the IKB kinase, a positive regulator of the survival factor NF-κB. Studies have shown that Akt can directly inactivate pro-apoptotic factors such as BAD [30].

Mitogen-Activated Protein Kinase (MAPK) families are well known to play an important role in cellular functions such as proliferation, development and apoptosis. Raf activates the MAPK/ERK kinase (MAP2K1), which then activates MAPK [31].

Wilson [32] displayed that SRC kinases are prime second messengers of HER2 (ERBB2), and the connection between SRC activation and overexpression of ERBB2 has been reported in breast carcinoma [33]. Ceramide (CERK) has been linked in diverse cellular processes, including proliferation and apoptosis [34], and Phosphoglycolate Phosphatase (PGP) is a Protein Coding gene that could increase the level of CERK in cells [35].

BCL-2 an apoptosis regulator is over-expressed in about 85% of ER-positive tumors [36]. Studies have shown that anamolous STAT3 signaling advances breast tumor growth through deregulation of BCL-2, BIRC5 and MCL1 [37]. Targeting the mTOR pathway also resulted in decreased

MCL1 expression postulating a link between them [38].

On the basis of the above biological information, we formulated the Breast Cancer signaling pathway as shown in Figure 2.3. The gene interactions are represented by arrows, where a normal arrow denotes activation and a red hammer arrow denotes inhibition.

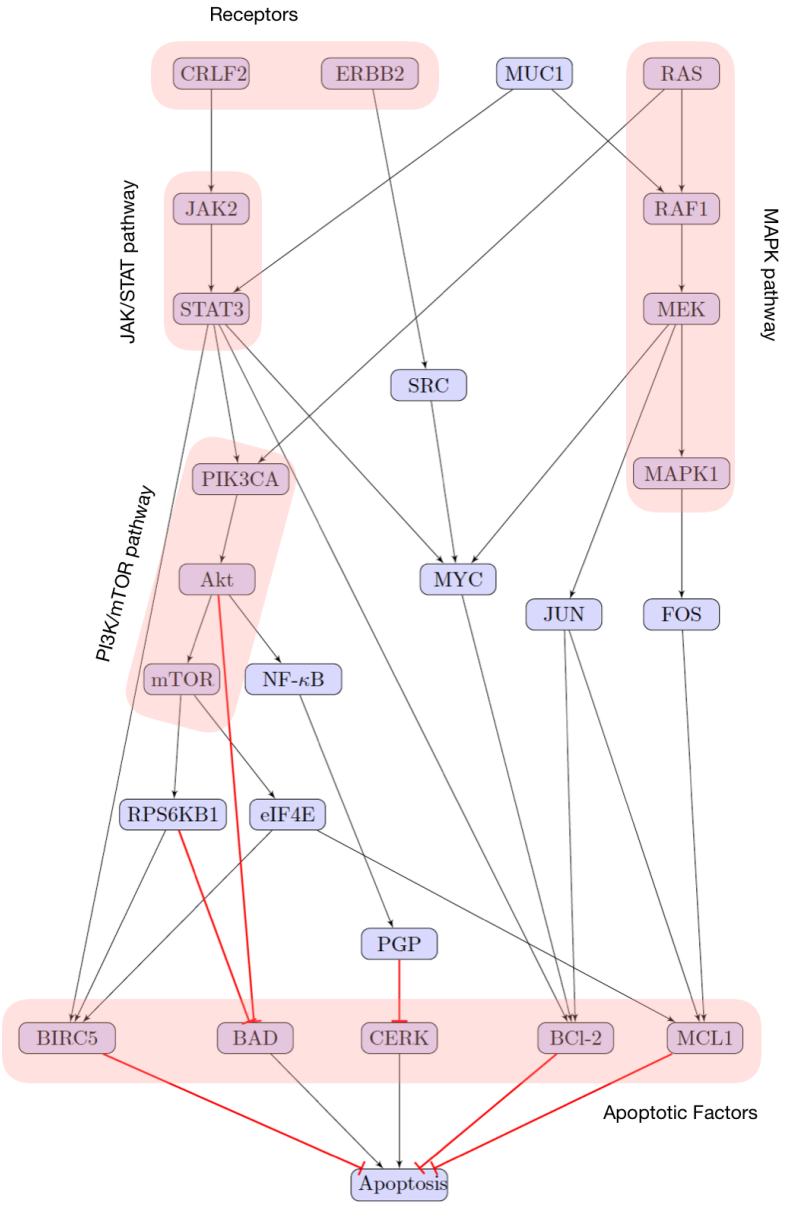


Figure 2.3: **Breast Cancer Signaling Pathway.** An arrow indicates activation and a red hammer arrow represents inhibition. (Reprinted from [1])

2.4 Results

2.4.1 Simulations

In this section, we discuss the application of our model to select critical points of intervention in the Breast Cancer pathway. We use real experimental data deposited in the NCBI database to test the efficacy of our model. Analysis of expression data has three key uses: classifying diseases, identifying decisive genes, and decoding biological pathways. Our choice of datasets meets these pivotal requirements. Two datasets GSE2990 [39] and GSE6532 [40] were discretized and pooled together. Each of these datasets contains the gene expression data obtained from the microarray experiments carried out on primary breast tumors. Using this gene expression data we calculated the CPDs of our Bayesian model using the method discussed in section 2.2.

Cancer is a disease in multicellular organisms that results from an imbalance between cell proliferation and programmed cell death leading to the formation of tumors that become malignant. Thus, a possible approach for treating cancer could be to enhance Apoptosis(cell death). It is well known that the BCL-2 protein family, consisting of anti-apoptotic and pro-apoptotic members, is involved in the regulation of apoptotic cell death. The anti-apoptotic members prevent apoptosis by preventing the release of Cytochrome-c into the cytoplasm. On other hand, enhanced expression of pro-apoptotic molecules can result in increased mitochondrial outer membrane permeability(MOMP), which leads to the release of Cytochrome-c. The hemeprotein Cyotchrome-c then recruits Apaf-1 and pro-caspase-9 to form the Apoptosome, which triggers the Caspase 9/3 cascade, resulting in apoptosis [41]. In our signaling pathway, CERK and BAD are pro-apoptotic genes, whereas, MCL1, BCL-2 and BIRC5 are anti-apoptotic genes.

We define a mathematical expression that incorporates the effect of both pro-apoptotic and anti-apoptotic factors. Define the Apoptosis Ratio by:

$$\text{Apoptosis Ratio} = \frac{\text{Gene Expression (Pro-Apoptotic genes)}}{\text{Gene Expression (Anti-Apoptotic genes)}} \tag{2.11}$$

Clearly, the higher the ratio, the greater the chances of Cytochrome-c release and consequently

greater cell death.

Now, using 's messaging passing algorithm discussed earlier, we computed this ratio conditioned on the evidence set **E**, where **E** is the set of the genes being directly targeted by drugs. In the next section, we discuss the results obtained.

Using the approach described above, we calculated the Apoptosis Ratio for two scenarios. First, we assumed a single gene intervention and then we looked at a combination therapy involving the simultaneous modulation of two genes.

2.4.2 Single Gene Intervention

We ran the algorithm for modulation using only one gene at a time and computed the corresponding Apoptosis Ratios as shown in Table 2.2. We then overlaid these ratios on the breast cancer pathway as shown in Figure 2.4, and for elegance, we plotted only the most significant genes in Figure 2.5. From the figure and the graph, it is evident that mTOR and Akt1 are the most preferred nodes for single gene intervention. From the figure, it is also evident that both the location of the gene and the number of crucial genes it intervenes with influence its apoptosis ratio.

2.4.3 Combination Therapy using Two Genes

Assuming that we can simultaneously intervene with a pair of genes, we ran the algorithm for every set of two genes. Due to the large number of combinations, we only tabulated the most significant ones in Table 2.3 and plotted them in Figure 2.6. From the graph, we infer that a combination of mTOR and STAT3 is the most preferred pair for intervention purposes.

2.4.4 Drug Intervention

We now look at the drugs that are well known to intervene and bind specific genes in the pathway, and compare computational results with the experimentally obtained ones.

APTSTAT3-9R is a specific STAT3-binding peptide that blocks STAT3 phosphorylation and reduces expression of STAT3 targets in various types of cancer cells [42]. Another STAT3 inhibitor HO-3867 has been shown to disrupt the JAK/STAT3 signaling pathway thereby reducing the expression levels of both JAK and STAT3 [43]. Lapatinib, a reversible inhibitor of HER-2/ERBB2,

Table 2.2: Ratio of Apoptosis: Single gene intervention. (Reprinted from [1])

Number	Gene Name	State	Value
1	CRLF2	OFF	0.484234
		ON	0.325834
2	JAK2	OFF	0.492859
		ON	0.300862
3	MUC1	OFF	0.387579
		ON	0.343696
4	STAT3	OFF	0.570441
		ON	0.262752
5	RAS	OFF	0.484234
		ON	0.333091
6	PIK3CA	OFF	0.742564
		ON	0.235739
7	RAF1	OFF	0.395569
		ON	0.350117
8	mTOR	OFF	1.249280
		ON	0.215840
9	eIF4E	OFF	0.697592
		ON	0.200470
11	Akt	OFF	1.047970
		ON	0.182339
12	MEK	OFF	0.43825
		ON	0.336071
14	NF-κB	OFF	0.716033
		ON	0.169933
15	MAPK1	OFF	0.422930
		ON	0.342773
16	PGP	OFF	0.813258
		ON	0.190818
17	FOS	OFF	0.380059
		ON	0.346885
18	CERK	OFF	0.133953
		ON	0.780397
22	JUN	OFF	0.423528
		ON	0.332256
23	RPS6KB1	OFF	0.899283
		ON	0.220879
24	ERBB2	OFF	0.363305
		ON	0.362379
25	SRC	OFF	0.364518
		ON	0.362207

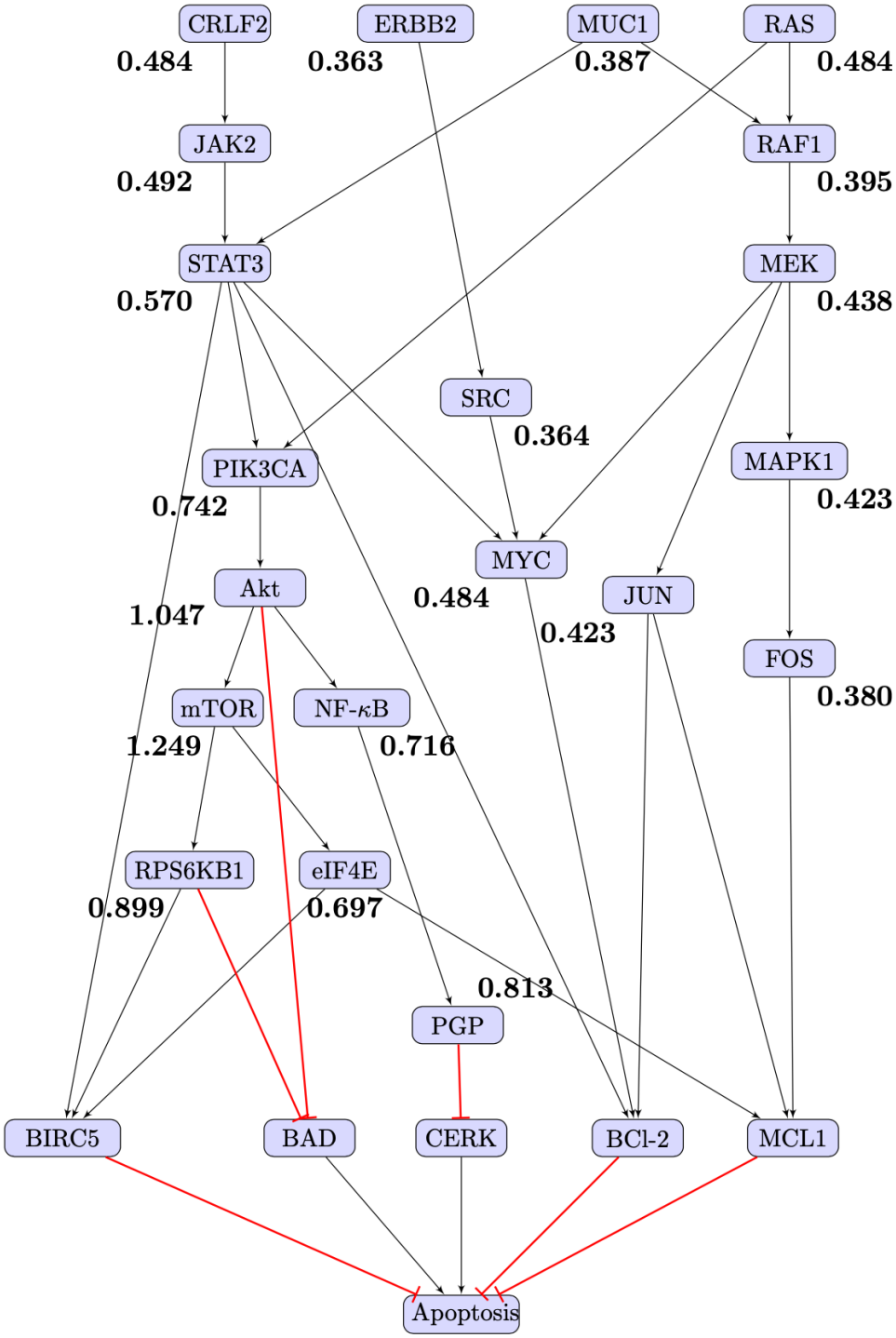


Figure 2.4: Apoptosis Ratios calculated by intervening different nodes independently. (Reprinted from [1])

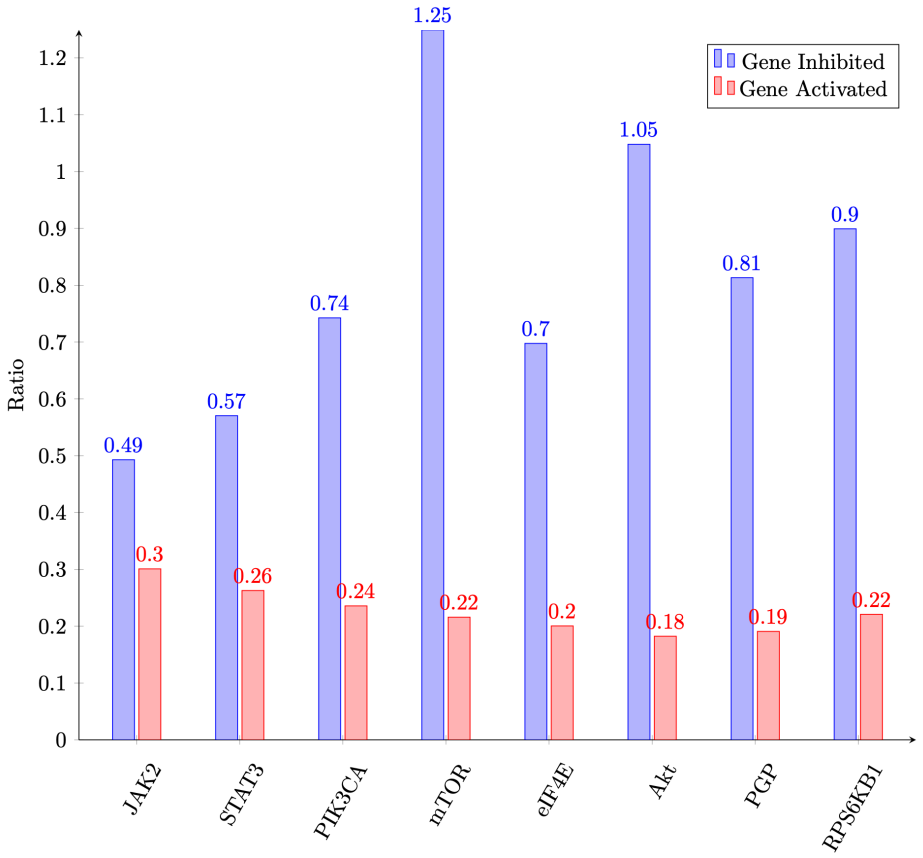


Figure 2.5: Bar graph illustrating the Apoptosis Ratio calculated by intervening different nodes. (Reprinted from [1])

has shown some success in different cancers [44]. LY294002, a selective PI3K inhibitor significantly induced cell apoptosis in MCF-7 cells [45]. Temsirolimus is a potent mTOR inhibitor that showed significant activity in vitro against a variety of cancer cells including MCF-7 breast cancer cells [46]. An et al. [47] showed the suppression of human breast cancer cell line using U0126 which is a specific MEK inhibitor. Cryptotanshinone, a naturally occurring drug has been shown to suppress the mTOR signaling pathway [48], and STAT3 signaling through blocking its dimerization [49]. Using the above biological information, we tabulated the drugs and the gene(s) they intervene in Table 2.4, and plotted the apoptosis ratio for different drugs.

Table 2.3: Ratio of Apoptosis: Combination therapy involving two genes. (Reprinted from [1])

Number	Gene Name	State	Value
1,2	CRLF2+JAK2	OFF	0.492859
		ON	0.300862
2,4	JAK2+STAT3	OFF	0.582264
		ON	0.262945
4,6	STAT3+PIK3CA	OFF	0.883097
		ON	0.222168
6,8	PIK3CA+mTOR	OFF	1.477788
		ON	0.188257
8,9	mTOR+eIF4E	OFF	1.275501
		ON	0.196231
6,9	PIK3CA+eIF4E	OFF	1.477788
		ON	0.188257
4,8	STAT3+mTOR	OFF	2.08768
		ON	0.186954
8,11	mTOR+Akt	OFF	1.281037
		ON	0.181387
9,11	eIF4E+Akt	OFF	1.264280
		ON	0.165171
1,8	CRLF2+mTOR	OFF	1.808874
		ON	0.205043
11,14	Akt+NF-κB	OFF	1.105998
		ON	0.137457
14,16	NF-κB+PGP	OFF	0.868523
		ON	0.148477
11,23	Akt+RPS6KB1	OFF	1.325971
		ON	0.174062

Table 2.4: Drugs used and their gene intervention points. (Reprinted from [1])

Drug	Gene(s) targeted
APTSTAT3-9R	STAT3
Lapatinib	HER-2/ERBB2
LY294002	PIK3CA
HO-3867	STAT3
Temsirolimus	mTOR
U0126	MEK
Cryptotanshinone	STAT3 + mTOR

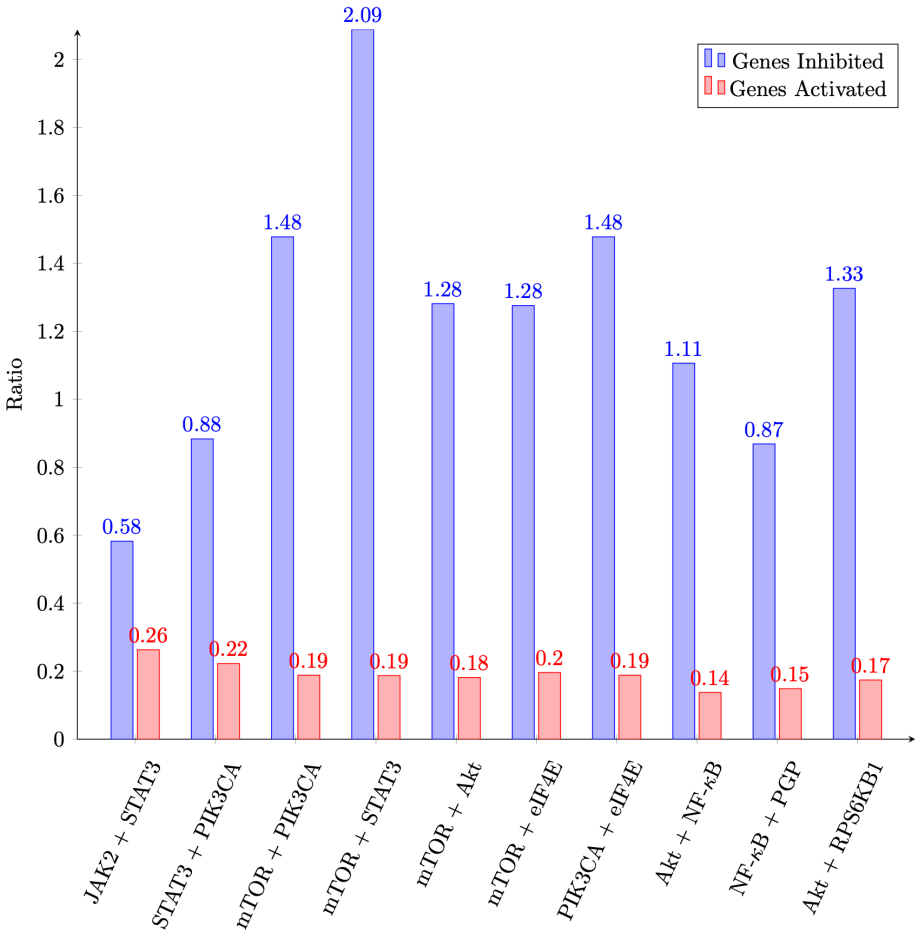


Figure 2.6: Bar graph illustrating the Apoptosis Ratio calculated by intervening a set of two nodes. (Reprinted from [1])

2.4.5 Single Drug Intervention

We calculated the apoptosis ratios when a single drug is used and plotted them in Figure 2.7.

2.4.6 Combination of Two Drugs

Considering the harmful side-effects of drugs, we restricted the maximum number of drugs per combination to two in our theoretical analysis and experiments. Hence, here we are interested in finding the drug combinations that maximize cell death with minimum side-effects. We calculated the apoptosis ratios when a combination of two drugs is used and plotted them in Figure 2.8.

From the plots in Figure 2.7 and Figure 2.8, it is clear that Cryptotanshinone either by itself, or

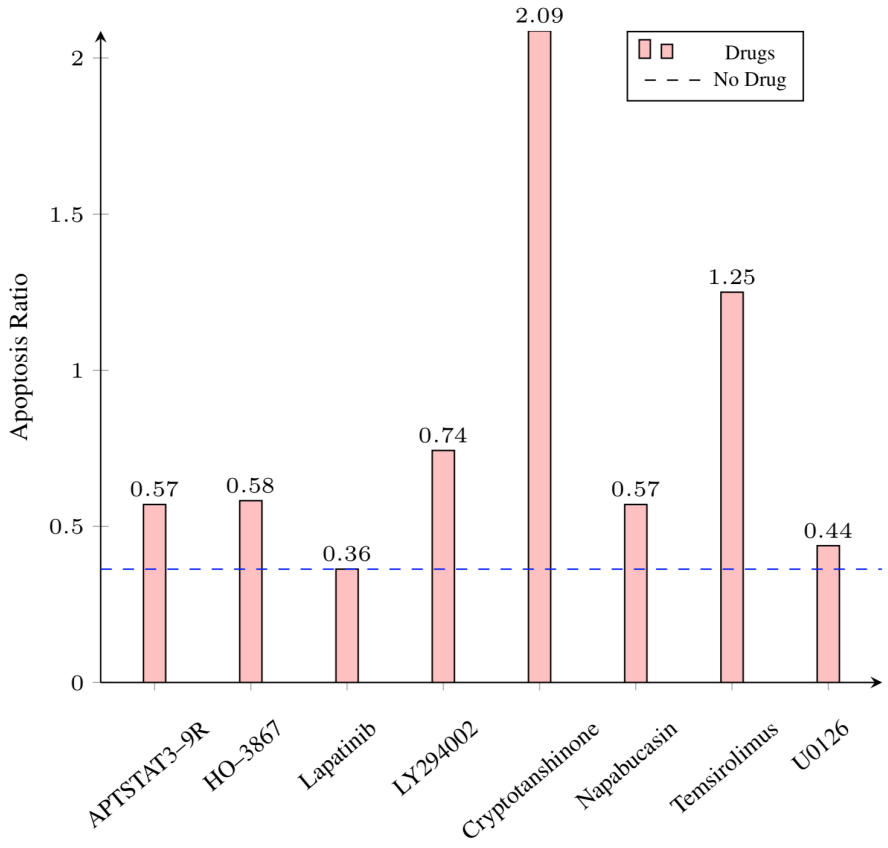


Figure 2.7: Bar graph illustrating the Apoptosis Ratio calculated when different drugs are used independently. (Reprinted from [1])

in combination with other drugs results in greatly enhanced cell death. To verify this, we treated MCF-7 breast cancer cell lines with these drug combinations and the experimental results agree closely with the computational predictions. We now present the experimental results.

2.4.7 Experimental Validation

The theoretical results obtained above were validated using MCF–7 breast cancer cell lines subjected to the various alternative drug treatments. The cell line was supplied to Dr. Jeffrey Trent's laboratory at the NIH from ATCC in 1996. This line is currently being used in the current collaboration between Dr. Trent's Institute, TGen, and the Texas A&M Center for Bioinformatics and Genomic Systems Engineering. We used high–content fluorescent protein reporter imaging

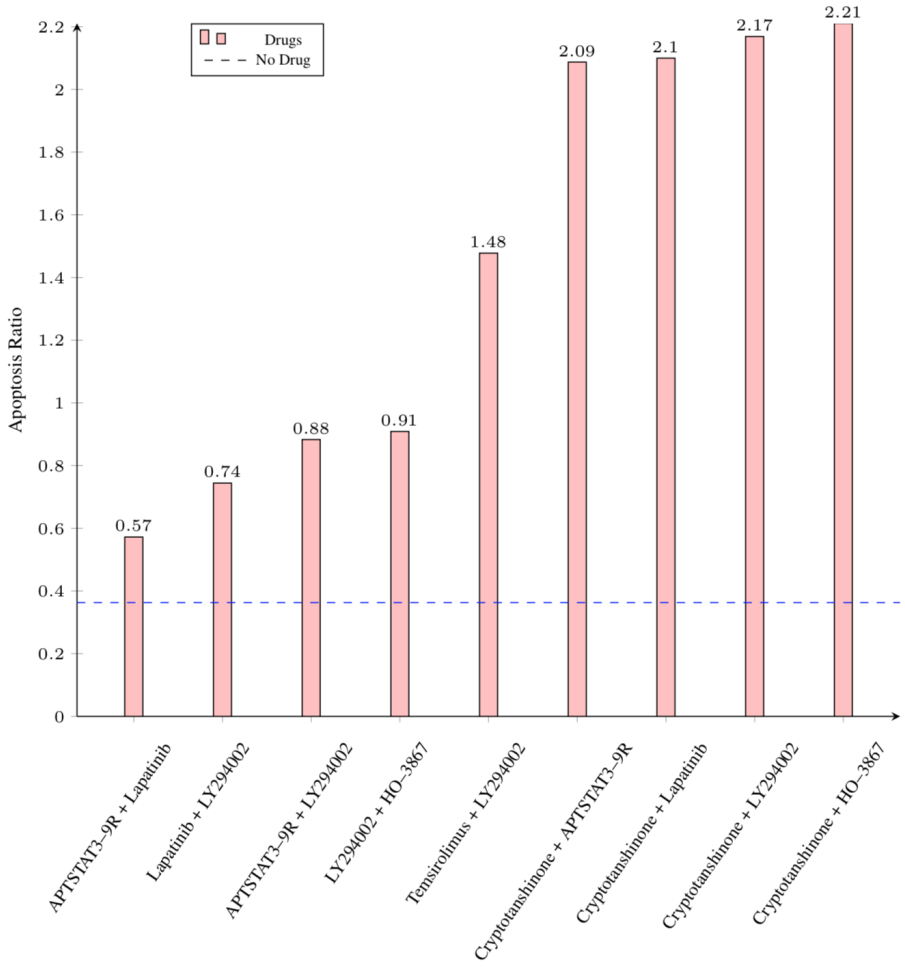


Figure 2.8: Bar graph illustrating the Apoptosis Ratio calculated when a set of two drugs are used. (Reprinted from [1])

method to track cell death in MCF–7 cells. We extracted cell processing dynamics using a two-part data processing procedure introduced in Hua et. al. [50]. We condensed this data obtained into expression profiles and represented as plots to interpret further.

The plots in Figure 2.9 show the cell killing produced in MCF7 breast cancer cell lines under the effect of different drug combinations compared against the untreated cell line. Cryptotanshinone has been used in each of the drug combinations and from the plots, it is clear that in each case, impressive cell death occurs within 6-8 hours. Furthermore, all the drug combinations have more than 95% apoptosis in 10 hours.

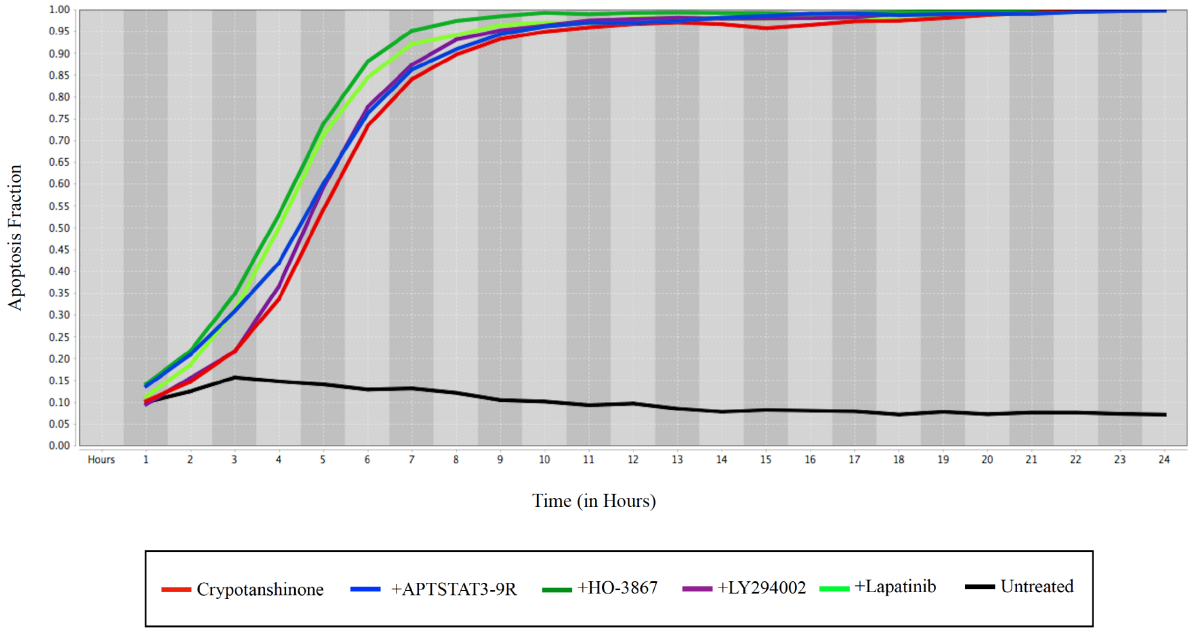


Figure 2.9: Apoptosis fraction versus time(in hours) for different drug combinations. The drug combinations in the legend from left to right are Cryptotanshinone, Cryptotanshinone + APTSTAT3-9R, Cryptotanshinone + HO-3867, Cryptotanshinone + LY294002, Cryptotanshinone + Lapatinib and Untreated cell line. (Reprinted from [1])

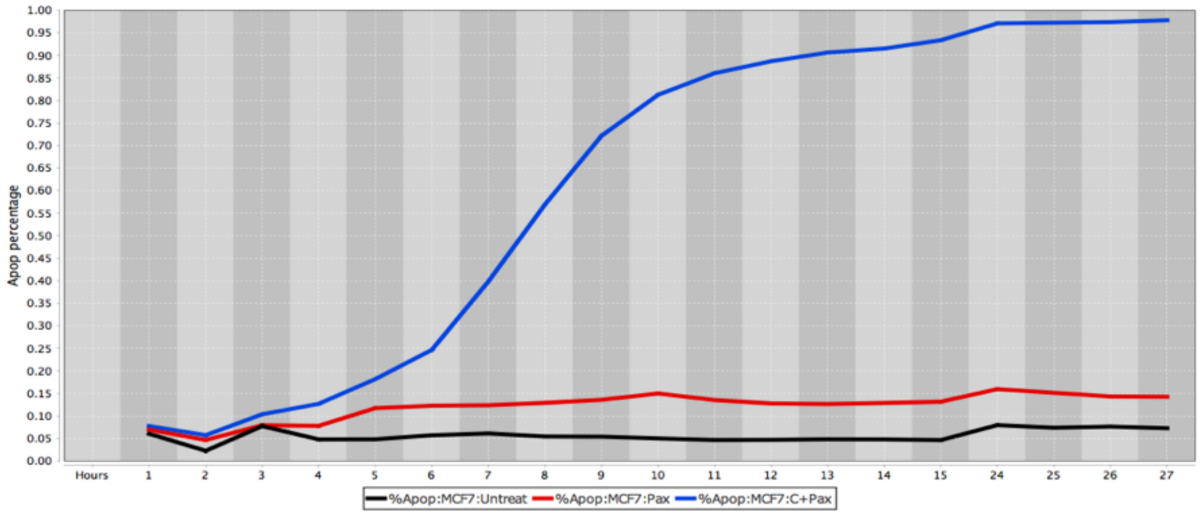


Figure 2.10: Apoptosis fraction versus time (in hours) for different drug combinations. The drug combinations in the legend from left to right are Untreated cell line, Paxlitaxel, Cryptotanshinone + Paxlitaxel. (Reprinted from [1])

In order to compare our results with the standard therapy for breast cancer, we treated the MCF7 cell line with Paxlitaxel [51] and its combination with Cryptotanshinone and compared against untreated cell line. We plotted these results as shown in Figure 2.10. From the figure, the drug is completely ineffective by itself, but upon the introduction of Cryptotanshinone, the performance of cell death improves significantly over time. These results further strengthen our argument that Cryptotanshinone substantially enhances cell death.

The ordering of the efficacy of cell death induction by the different combinations appears to be in line with what is predicted by our computational approach. Hence, our computational results agree closely with the experimentally obtained ones.

3. IN SILICO DESIGN AND EXPERIMENTAL VALIDATION OF COMBINATION THERAPY FOR PANCREATIC CANCER*

3.1 Introduction

Pancreatic ductal adenocarcinoma (commonly referred to as pancreatic cancer) is the most common malignancy of the pancreas. The American Cancer Society recently reported that the number of deaths associated with pancreatic cancer has been increasing at an alarming rate making it the third leading cause of cancer related deaths in the United States [4]. In 2018, it is estimated that there will be 55,440 new diagnosed cases and 44,330 deaths linked to pancreatic cancer in the United States alone. It is also estimated that by 2030, annual deaths due to pancreatic cancer will exceed that of breast, prostate, and colorectal cancers [52].

The annual number of deaths for most cancers has decreased in recent decades, but the death rate for pancreatic cancer has remained significantly flat. The absence of clinical progress in pancreatic cancer in comparison with other cancers is ascribed to a lack of success in developing novel and effective therapies. This grim outlook for pancreatic cancer is linked to various reasons. Pancreatic cancer is usually diagnosed at advanced stages, which is usually due to an absence of early symptoms and a lack of detecting/imaging techniques for early-stage tumors [53]. Pancreatic cancer also sets itself apart from other cancers because of its exceptional resistance to most traditional medications, including radiotherapy and chemotherapy. This resistance stems from the complexity pancreatic cancers carry at the genomic level, with diverse activated pathways and apparent cross-talk [54].

Historically, biologists have captured cause-effect interactions between different biological molecules using signaling pathways. Although marginal in nature, such information can provide useful therapeutic pointers for diseases that result from a simple breakdown of such signaling. In

*Parts of this section are reprinted with permission from H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "In Silico Design and Experimental Validation of Combination Therapy for Pancreatic Cancer," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 3, pp. 1010-1018, 2020, doi: 10.1109/TCBB.2018.2872573

the case of cancer, however, the success with this approach has been very limited mainly because of the complexity of the possible breakdowns resulting in the manifestation of the disease [55]. Consequently, in recent years there have been several attempts to holistically model the interactions between different biological molecules of interest. Specific approaches used include Bayesian networks [14], Differential equations [56], Boolean Networks [57] and Probabilistic Boolean networks [8].

In this work, we describe a methodology that utilizes current biological knowledge from the literature to build a Boolean Network model of the pancreatic cancer pathway. We model the gene interactions in the pathway using appropriate logic gates. We simulate this boolean network with drugs at appropriate intervention points to calculate a "measure", which is defined to capture the extent of activation/deactivation of cell death. Such a measure can aid us in theoretically assessing the effectiveness of drugs, provided the primary goal of therapy is to facilitate the death of cancer cells.

3.2 Methodology

Gene regulatory networks, which describe the interactions between genes and other molecules, play a pivotal role in orchestrating most biological processes such as cell proliferation, differentiation, metabolism, and apoptosis. Understanding the mechanics of these networks can assist us in dissecting the mechanisms of the diseases that result when these cellular processes deviate from the norm [58]. Mathematical and computational methods have been developed for modeling these gene interactions. These mathematical models have had some success in capturing some of the complexities of biological networks [59].

One such widely used model is the Boolean Network (BN) model. Modeling of biological interactions using BN has been successfully used for studying the Growth Factor signaling pathway [57] and the Prostate Cancer pathway [6]. We now briefly review the key concepts needed for modeling biological pathways with Boolean networks.

3.2.1 Boolean Networks

In a Boolean network, each node can be in one of two binary states, inactive or active. It is customary to assign a '0' value to the inactive state and a '1' value to the active state. In a gene regulatory network (GRN), genes can be binarily quantized [60] and classified as either up-regulated or down-regulated based on their gene expression levels. This switch-like behavior of genes clearly can be modeled within a binary framework, thereby making boolean networks an apparent choice to model GRNs. Furthermore, in a GRN, a gene is influenced by one or more genes. This interaction among different genes can be modeled as a boolean logic function, where the nodes represent the genes and the edges represent the interactions among the genes.

3.2.2 Modeling Abnormalities

Cancer is a collection of diseases in which abnormal cells are busy producing more cells, the death of many produced cells, and the movement of cancer cells to other places in the body. This abnormality can be caused by a malfunction in the normal signaling pathways leading to the loss of cell cycle control and uncontrolled cell proliferation. One of the most common and well known malfunctions is the mutation of a gene, which leads to its over- or under-expression. This aberration can be modeled as a stuck-at fault in the BN. When a stuck-at fault occurs at a gene, its value gets fixed at (0/1) and is no longer influenced by the activity status of other genes.

3.2.3 Modeling Drug Intervention

Drugs usually work by interacting with receptors on the surface of cells or on enzymes within cells. They can either block the function of the protein (inhibitory drug), or induce it's effect (enhancing drug) by binding to the target receptor site. This interaction of the drug with a gene can be mapped to a BN by either forcibly suppressing or enhancing the value of the gene at the appropriate location.

We illustrate the type of modeling just discussed with the help of an example. Let us assume a scenario where genes A and B independently activate gene C, genes C and D form a heterodimer and activate gene E which then activates gene F. Furthermore, let us suppose that gene D is mutated

and a drug precisely inhibits gene F. These interactions and the associated equivalent BN could be represented as shown in Figure 3.1.

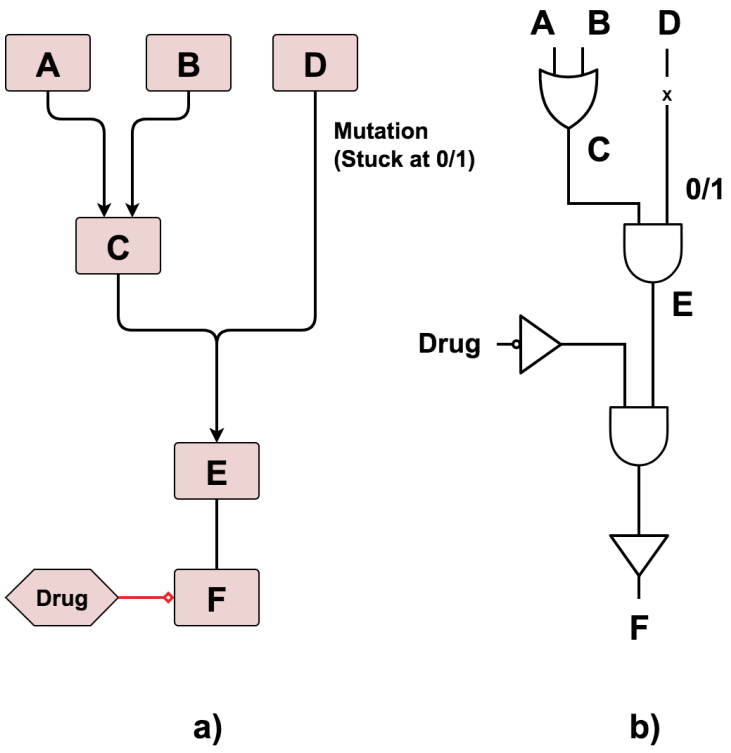


Figure 3.1: a) Example gene regulatory network. b) Boolean equivalent of the example gene regulatory network. (Reprinted from [2])

We next analyze the signaling pathways and the specific drug intervention points associated with pancreatic cancer.

3.3 Pancreatic Cancer Pathways

Signaling pathways supervise cellular processes such as growth, division, and death. Abnormality in any of these controlled processes can lead to cancer. Abnormal paracrine and autocrine signalling cascades in pancreatic cancer advance cell proliferation, invasion and metastasis.

Signalling molecules such as epidermal Growth Factor (EGF), insulin-like growth factor 1 (IGF1), heparin binding EGF like growth factor (HBEGF), and their respective tyrosine kinase

receptors such as epidermal growth factor receptor (EGFR), IGF1 receptor (IGF1R), receptor tyrosine-protein kinase (ERBB2), activate different pathways that strengthen pancreatic cancer cells' self-reliance and boost migration and invasion [61]. EGFR can form both homodimers and heterodimers with ERBB2, and in the case of mutated Kirsten ras (KRAS), downstream signalling has been shown to be excessively activated. In the presence of growth factor receptor-bound protein 2 (GRB2), these tyrosine kinase receptors were shown to be more effectively activated [62].

Liver kinase B1 (LKB1) is a well known serine/threonine kinase which senses changes in cellular energy and adjusts metabolic processes by triggering its downstream kinase, AMP-activated protein kinase (AMPK) [63]. These binding events make way for the activation of RAS, RAF and mitogen-activated protein kinase (MAPK) signalling. In addition to these signaling cascades, anti-apoptotic and pro-survival pathways such as signal transducer and activator of transcription 3 (STAT3), phosphatidylinositol 3-kinase (PI3K) and AKT are simultaneously activated [64]. Additionally, extracellular-regulated kinase (ERK1/2) and c-jun N-terminal kinase (JNK), members of the MAPK family, have been demonstrated to regulate cell survival.

These pathways are supervised by a series of phosphatases, kinases and multiple exchange proteins, and mutations in these pathways can lead to uncontrolled cell proliferation and eventually cancer.

3.3.1 Drug Interventions

In the literature, scientists have established the specific receptors/enzymes where different drugs intervene in a signaling pathway. We superimpose that data on our BN model to specify the genes that will be inhibited by a particular drug. Before doing so, we first list our drugs of interest and their specific targets.

HO-3867, a specific STAT3-binding peptide has been shown to block STAT3 phosphorylation which results in the disruption of the JAK/STAT3 signaling pathway [65]. Temsirolimus is a selective mTOR inhibitor that has demonstrated significant activity in vitro against a variety of cancer cells [66, 67]. Lapatinib, a reversible inhibitor of HER-2/ERBB2 and growth factor receptors has shown some success in multiple cancers [68, 69]. LY294002 is a potent PI3K inhibitor that has

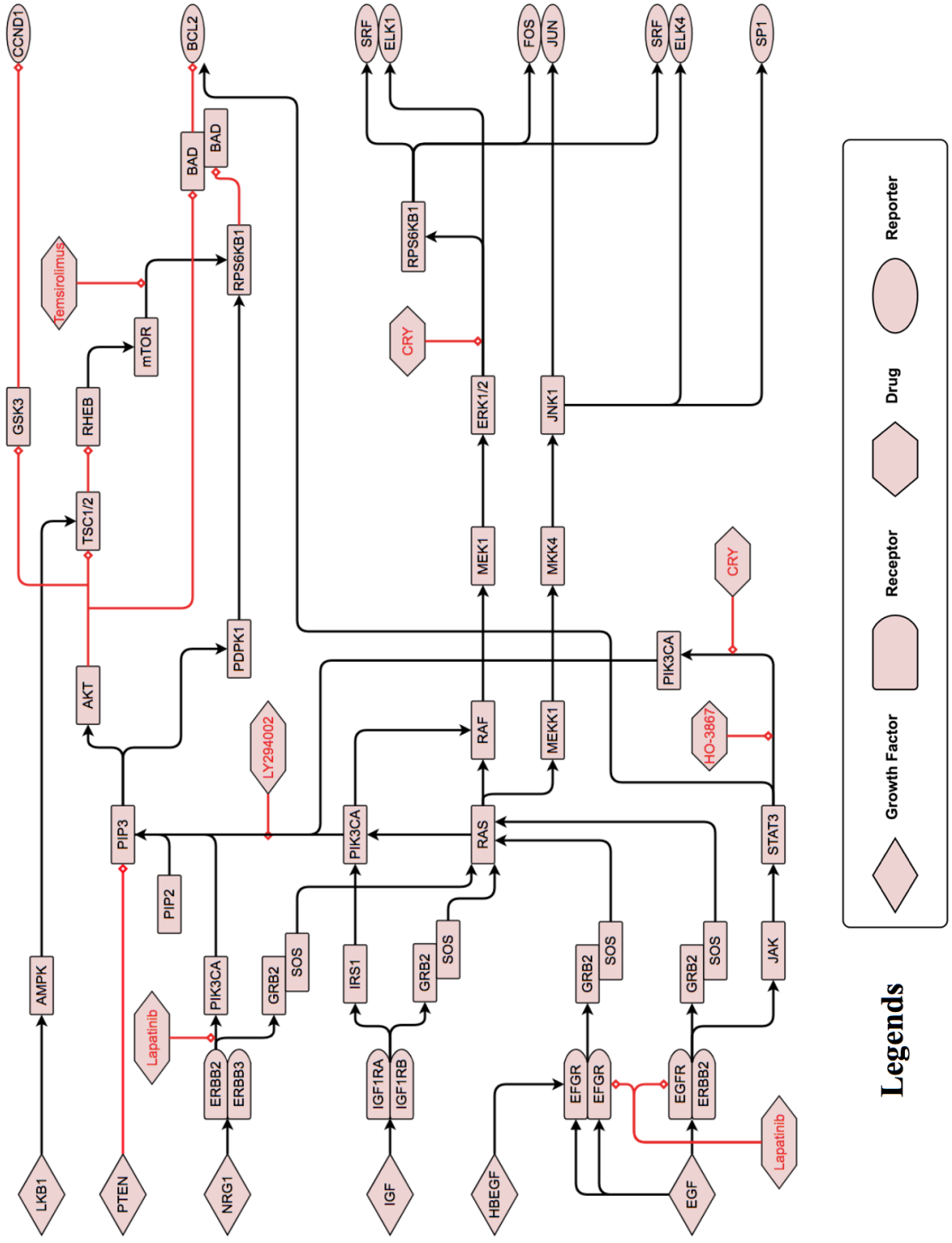


Figure 3.2: **Pancreatic cancer pathway.** A black arrow denotes activation and a red arrow denotes inhibition. The legends explain the role of different bounding boxes. (Reprinted from [2])

notably induced cell death in cancer cells by impeding the PI3K/AKT pathway [70]. Cryptotanshinone (CPT), a naturally occurring drug has been shown to suppress STAT3 signaling in pancreatic cancer cells by arresting the cell cycle in the G1-G0 phase, indicating that CPT is an effective STAT3 inhibitor [71]. In the same study, the expression level of ERK1/2 was also significantly inhibited by CPT. Cryptotanshinone has shown exceptional success in not only pancreatic cancer cells but also in prostate cancer [48], human glioma [72], and chronic myeloid leukemia cells [73]. Consequently, we decided to include CPT in our list of drugs so that we could theoretically analyze its beneficial effects, both while acting solo or in combination with other drugs.

Using the information presented above and the available literature, we constructed the gene regulatory network of pancreatic cancer as shown in Figure 3.2. The gene interactions are represented by arrows, where a black arrow denotes activation and a red arrow denotes inhibition.

The boolean equivalent of this regulatory network is shown in Figure 3.3. The fault locations are denoted by numbers in parentheses, where black numbers denote stuck-at-1 faults and blue numbers denote stuck-at-0 faults. We next discuss the theoretical results obtained from the developed model and then present experimental results to support our theoretical conclusions.

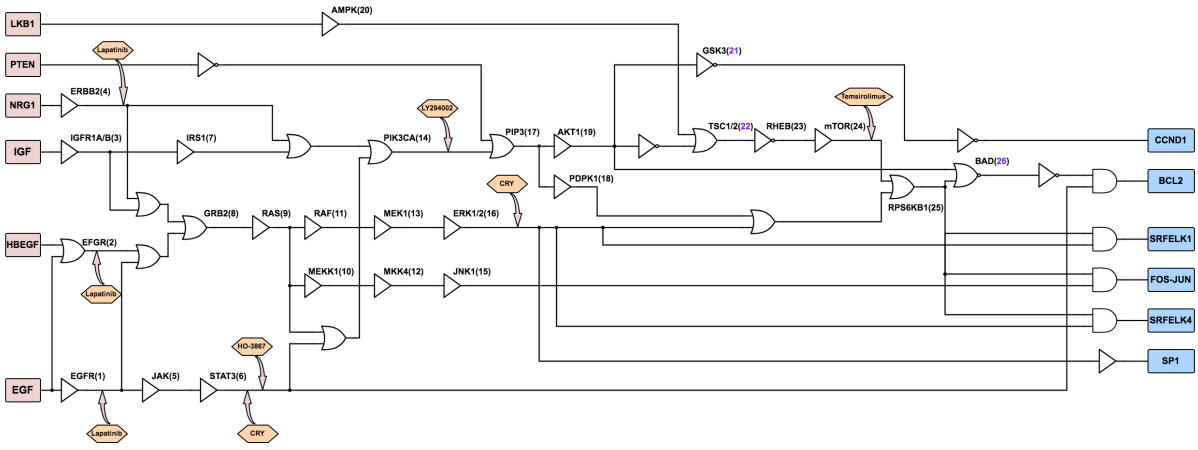


Figure 3.3: **Boolean equivalent of pancreatic cancer pathway.** The numbers in parentheses represent the identifying number assigned to a fault at that location. Here, black numbers denote stuck-at-1 faults and blue numbers denote stuck-at-0 faults. (Reprinted from [2])

3.4 Results

3.4.1 Simulations

Using the Boolean model constructed above, we can now compare the different drug combinations and their efficacies. For each possible mutation/fault, the objective is to find the best combination therapy that can alleviate the deleterious impact of that fault.

Referring to the Boolean model in figure 3, we have six inputs and six outputs. The inputs include two tumor suppressors (PTEN, LKB1) and four growth factors (EGF, HBEGF, IGF, NRG1). The outputs consist of pivotal genes (CCND1, BCL2, SRF-ELK1, FOS-JUN, SRF-ELK4, SP1) associated with and indicative of cell proliferation and apoptosis.

For mathematical analysis, we can represent these inputs and outputs as row vectors. A zero corresponds to an inactive gene in the matching location, and on the other hand, a one corresponds to an active gene. So the binary input and output vectors will be given by:

$$\text{Input} = [\text{PTEN, LKB1, EGF, HBEGF, IGF, NRG1}] \text{ and}$$

$$\text{Output} = [\text{CCND1, BCL2, SRF-ELK1, FOS-JUN, SRF-ELK4, SP1}].$$

For the input [110000], the tumor suppressors are active and the growth factors are inactive, and this input corresponds to an absence of proliferation and a non-reduction in apoptosis. In the Boolean network with no faults, this input produces the output [000000], which also corresponds to a lack of cell proliferation and a non-suppression of apoptosis. However, for the same input, the network with faults will produce a different non-zero output vector. Our goal here is to guide this non-zero output vector closer to the zero vector with the assistance of drugs. Biologically, this is equivalent to driving a mutated pathway towards non-proliferation and unsuppressed apoptosis using therapy.

The drugs we used in our simulations and experiments are, Cryptotanshinone (20 μM), LY294002 (10 μM), Temsirolimus (10 μM), Lapatinib (5 μM), and HO-3867 (10 μM). The drug dosage levels of Lapatinib and Temsirolimus are those of human medical use, and the dosages for HO-3867,

LY294002, and Cryptotanshinone are at levels similar to the tests of their utilities on human and canine cell lines. Once again, we can represent the activity status of the different drugs using a row vector:

Drug = [Cryptotanshinone, LY294002,Temsirolimus, Lapatinib, HO-3867].

Each component of the drug vector will be one or zero according to whether that particular drug is applied or not.

As stated above, we are interested in steering the output vector of a network with faults in the direction of a desirable output vector. In order to quantify the dissimilarity between two output vectors, we introduce a measure called Size Difference (SD). This quantity measures how different two vectors are, and its value is proportional to the dissimilarity between the two vectors.

Let $\bar{a} = (a_1, \ldots, a_n)$ and $\bar{b} = (b_1, \ldots, b_n)$ be two binary vectors. Then, we can count the number of matches and mismatches at each bit location and construct a confusion matrix as shown in Figure 3.4. Here the entries B and C keep count of the two types of possible mismatches summed across all locations while the entries A and D keep count of the two types of possible matches summed across all locations.

	$a_i = 1$	$a_i = 0$
$b_i = 1$	A	B
$b_i = 0$	C	D

Figure 3.4: Example confusion matrix. (Reprinted from [2])

Using these quantities, the Size Difference (SD) is then defined by:

$$d_s(\bar{a}, \bar{b}) = \left(\frac{B + C}{A + B + C + D} \right)^2 \tag{3.1}$$

Since a higher size difference correlates with a higher deviation from the ideal output, the outputs with greater SD would correspond to greater cell proliferation and/or reduced apoptosis and possibly a higher risk of cancer.

The Boolean network when simulated across all possible faults and drug combinations yields a matrix of SDs. An example matrix is shown below in Table 3.1. For each fault, we compare the row entries, and the drug vector that corresponds to the row with the smallest value yields the most desirable combination (for that fault). Similarly, in order to ascertain the most advantageous drug combination across all faults, we add all the columns and choose the row corresponding to smallest sum. In our example matrix, the second combination is favored across all faults and specifically for faults 1 and 3.

Table 3.1
Example size difference matrix. (Reprinted from [2])

	Fault 1	Fault 2	Fault 3
Drug combination 1	0.3	0.4	0.3
Drug combination 2	0.1	0.5	0.2
Drug combination 3	0.5	0.3	0.6

In this work, we additionally examined the existence of two faults and three faults simultaneously. Considering the harmful side-effects of drugs, we restricted the maximum number of drugs per combination to two in our experiments. A simple algorithm to summarize the complete method discussed is presented in Algorithm 3 below. The outputs 1 and 2 correspond to the best drug combinations for each fault and across all faults respectively.

Using the method discussed above, we implemented the boolean network and simulated the model in Matlab. We now present the theoretical results obtained followed by the experimental results.

3 Algorithm to find the best drug combinations

function
  construct pathway from literature
  design boolean network from pathway
  simulate the boolean network
    for each fault j **do**
      for each drug combination i **do**
        calculate SD(i, j)
      end for
    end for
  output$_1(j)$ = **arg min**$_i$ SD(i, j)
  output$_2$ = **arg min**$_i$ $\sum_j$ SD(i, j)
end function

3.4.2 Theoretical Results

We calculated the Size Difference for each combination of drug and fault. This arithmetic was performed for one, two and three faults occurring simultaneously, and since there are 26 possible fault locations, a total of ${}^{26}C_1 + {}^{26}C_2 + {}^{26}C_3 = 2951$ combinations were considered. We plotted these values as a box plot as shown in Figure 3.5. Clearly, the drug combinations with Cryptotanshinone yield small SD for a higher fraction of networks.

Furthermore, for each therapy, we find the average of size differences across all faults, and call it an “overall measure". Since we are interested in finding the best drug combination, the smallest overall measure corresponds to the most favorable combination. We first present the overall measure values for at most one and at most two faults occurring simultaneously for each drug combination in Table 3.2. We further present the overall measure for at most three faults occurring simultaneously for each drug combination in Table 3.3. From the table, it is evident that the bottom rows involving Cryptotanshinone resulted in very small measures, and the combination of Cryptotanshinone with LY294002 had the lowest value. This mathematical output promises low cell-proliferation and/or enhanced apoptosis in cells when Cryptotanshinone is used.

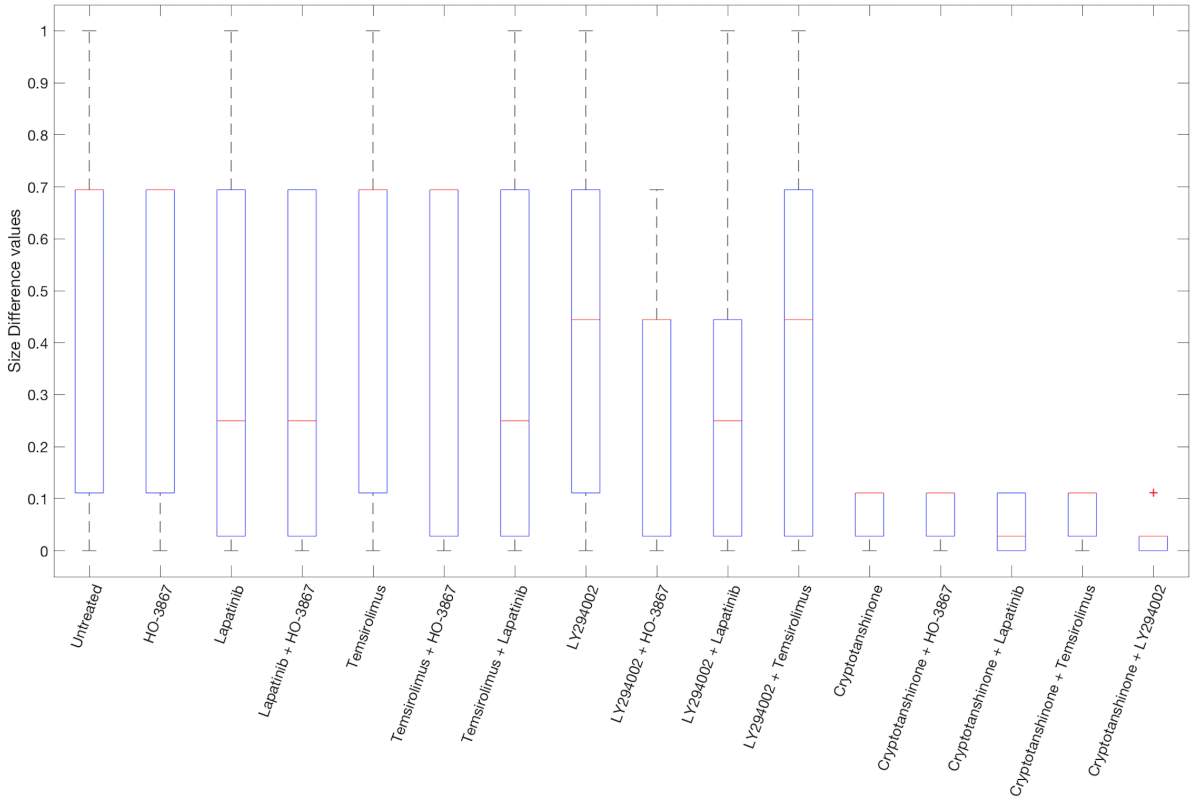


Figure 3.5: Box plot of the size differences calculated across all faults for different drug combinations. (Reprinted from [2])

3.4.3 Experimental Results

The theoretical results we obtained above were corroborated using experiments conducted on HPAC pancreatic cells subjected to different drug treatments.

Using a high-content fluorescent protein reporter imaging method, we detected cell death in HPAC cells. We extracted cell processing dynamics using a two-step data processing methodology introduced in 2012 [50]. Subsequently, we condensed this acquired data into expression profiles and plotted them to elucidate further.

The plots in Figure 3.6 display the cell killing produced in HPAC pancreatic cancer cell lines under the effect of different drug combinations. The black line denotes the untreated cell line which serves as a reference. Cryptotanshinone has been used in each of the drug combinations and from the plots, it is apparent that in each instance, impressive cell death occurs within 6-8 hours.

Table 3.2
The overall measure obtained for each of the drug combinations for at most one fault & two faults occurring simultaneously. (Reprinted from [2])

Drug combinations	Overall Measure (one fault)	Overall Measure (two faults)
Untreated	1.0000	1.0000
HO-3867	0.9055	0.8710
Lapatinib	0.5721	0.6245
Lapatinib + HO-3867	0.5323	0.5570
Temsirolimus	1.000	0.9982
Temsirolimus + HO-3867	0.9055	0.8693
Temsirolimus + Lapatinib	0.5721	0.6227
LY294002	0.6716	0.7089
LY294002 + HO-3867	0.6269	0.6299
LY294002 + Lapatinib	0.3881	0.4422
LY294002 + Temsirolimus	0.6716	0.7060
Cryptotanshinone	0.1443	0.1402
Cryptotanshinone + HO-3867	0.1443	0.1402
Cryptotanshinone + Lapatinib	0.0846	0.0899
Cryptotanshinone + Temsirolimus	0.1443	0.1384
Cryptotanshinone + LY294002	0.0149	0.0293

Moreover, all the drug combinations have more than 90% apoptosis in 15 hours. The fluorescent images capturing cell death over time for Cryptotanshinone + LY294002 are shown in Figure 3.7. From the figure, the tumor cells show increase in fluorescence and a lack of membrane integrity over time and this demonstrates the cell killing carried out by Cryptotanshinone over time.

In order to compare our results with the standard therapy for pancreatic cancer, we treated the PANC1 cell line with Gemcitabine [74], Gefitinib [75, 76] and their combinations with Cryptotanshinone and plotted these results in Figure 3.8. From the figure, the drugs are completely ineffective by themselves, but upon the introduction of Cryptotanshinone, the performance of cell

Table 3.3
The overall measure obtained for each of the drug combinations for at most three faults occurring simultaneously. (Reprinted from [2])

Drug combinations	Overall Measure
Untreated	1.0000
HO-3867	0.8493
Lapatinib	0.6727
Lapatinib + HO-3867	0.5852
Temsirolimus	0.9983
Temsirolimus + HO-3867	0.8473
Temsirolimus + Lapatinib	0.6705
LY294002	0.7460
LY294002 + HO-3867	0.6386
LY294002 + Lapatinib	0.4984
LY294002 + Temsirolimus	0.7419
Cryptotanshinone	0.1377
Cryptotanshinone + HO-3867	0.1377
Cryptotanshinone + Lapatinib	0.0957
Cryptotanshinone + Temsirolimus	0.1353
Cryptotanshinone + LY294002	0.0419

death improves significantly over time. These results further strengthen our argument that Cryptotanshinone substantially enhances cell death.

As a final observation, we identify that different drugs have different kinetics in Figure 3.6, and we believe that there are primarily two reasons for this. Firstly, since the binding of a drug to a specific receptor is a chemical process, the docking and activation probably takes different times for different structures. Secondly, since different drugs act at different locations in the pathway, the locations of these receptors might play an important role in the time course of the experiment.

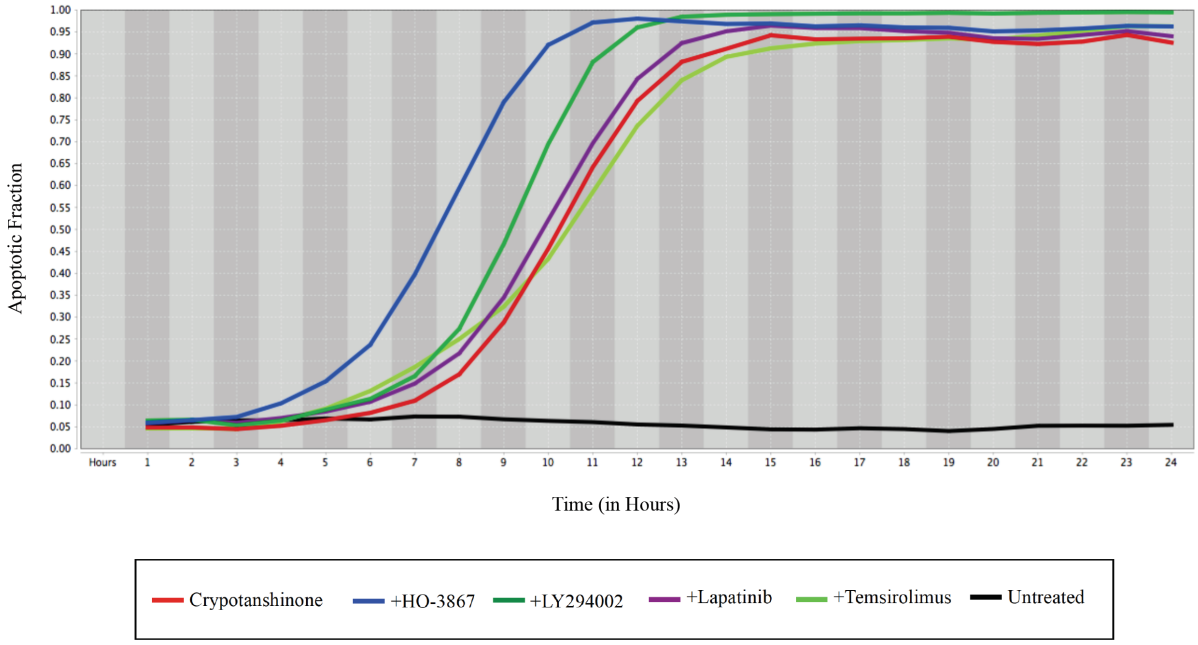


Figure 3.6: Apoptosis fraction versus time (in hours) for different drug combinations. The drug combinations in the legend from left to right are Cryptotanshinone, Cryptotanshinone + HO-3867, Cryptotanshinone + LY294002, Cryptotanshinone + Lapatinib, Cryptotanshinone + Temsirolimus, Untreated cell line. (Reprinted from [2])

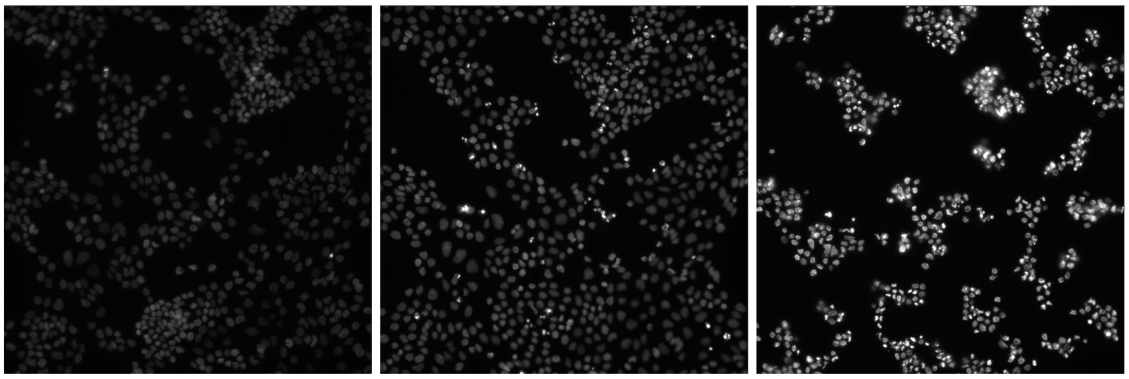


Figure 3.7: **Fluorescent images capturing cell death over time.** (a) Without the presence of any therapy, the tumor cells are intact and there is no fluorescence. (b,c) After adding Cryptotanshinone + LY294002, the tumor cells show a lack of membrane integrity with the presence of fluorescence over time. The increase in fluorescence over time demonstrates the cell killing carried out by Cryptotanshinone over time. (Reprinted from [2])

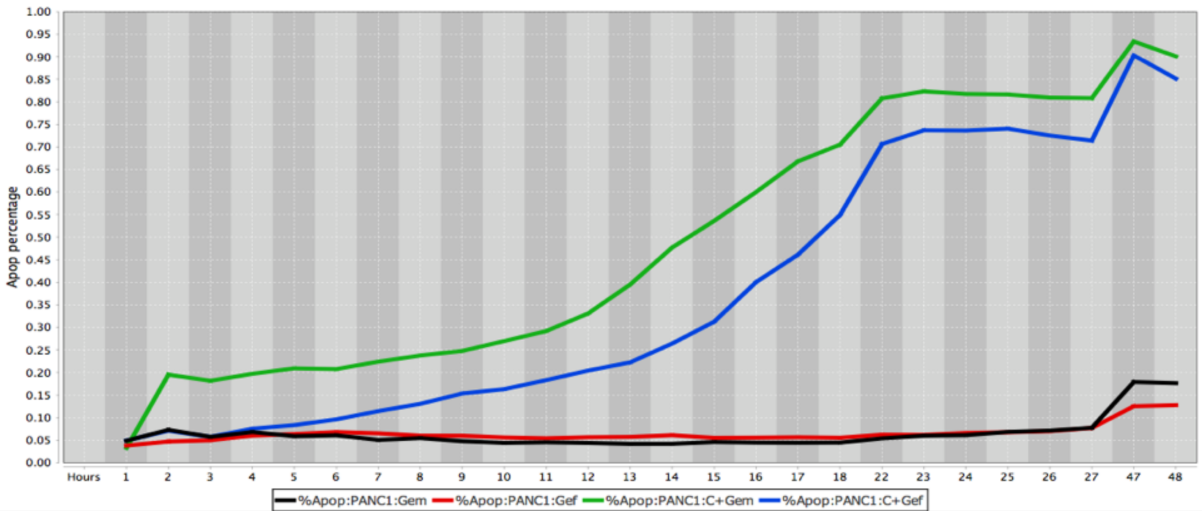


Figure 3.8: Apoptosis fraction versus time (in hours) for different drug combinations. The drug combinations in the legend from left to right are Gemcitabine, Gefitinib, Cryptotanshinone + Gemcitabine, Cryptotanshinone + Gefitinib. (Reprinted from [2])

The ordering of the potency of cell death by the different combinations seems to be in line with what is expected by our mathematical approach. Hence, our computational results concur with the experimentally obtained ones.

4. CRYPTOTANSHINONE INDUCES CELL DEATH IN LUNG CANCER BY TARGETING ABERRANT FEEDBACK LOOPS*

4.1 Introduction

Lung Cancer is globally the dominant cancer killer for both sexes. It is estimated that in the United States alone, there will be 228,150 new diagnosed cases and 142,670 deaths linked to lung cancer in 2019 [4]. In the last 40 years, the 5-year survival rate in the US has only increased to 20% from 12% showing a large room for improvement. This meager progress in the treatment of lung cancer is mainly linked to its complex and heterogeneous molecular basis. Since lung cancers advance through a multistage process comprising the evolution of multiple mutations, a deeper understanding of the mutations at multiple levels and their significance has the potential to help develop treatment strategies that can impact the diagnosis and treatment of the disease [77].

Multicellular organisms have developed highly sophisticated communication networks to integrate and coordinate various biological processes. Potent negative feedback loops regulate these processes in a controlled fashion and hence the elucidation of these feedback loops has surfaced as an important research area for designing effective cancer therapies [78]. In recent times, scientists have approached this drug-design problem as a control theoretic one and have used signaling pathways to examine the cause-effect interactions between biological molecules and therapeutic drugs [79]. The major approaches used to date for modeling gene regulatory network (GRN) interactions include Differential equations [80], Bayesian networks [9, 81, 1], Boolean Networks [57], and Probabilistic Boolean networks [82, 83]. Specifically, Boolean networks have lately shown considerable success in modeling various cancers when modeling of biological feedbacks is not all that crucial. On the other hand, they are not well-suited for capturing the typical feedback loops in GRNs that administer many cellular processes. Therefore, we propose here a Modified Boolean

*Parts of this section are reprinted with permission from H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "Cryptotanshinone Induces Cell Death in Lung Cancer by Targeting Aberrant Feedback Loops," *in IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 8, pp. 2430-2438, 2020, doi: 10.1109/JBHI.2019.2958042

Network that can address this crucial aspect. We make first use of the literature to construct the lung cancer pathway. We then design the appropriate boolean network using the modified rules. Lastly, we simulate this boolean network with drugs at appropriate intervention points to theoretically assess their effectiveness for killing cancer cells and validate our theoretical results with experiments on cell lines.

4.2 Methodology

It is intuitively obvious that a better comprehension of the workings of gene regulatory networks could aid us in dissecting the mechanisms of diseases such as cancer that arise when cellular processes behave in an aberrant fashion. In order to achieve this, several mathematical frameworks have been developed to model these networks [59].

4.2.1 Boolean Network

Boolean Network (BN) modeling is one such framework that has recently proven useful for studying multiple cancers [6, 2]. In a nutshell, for a Boolean network, we assign binary values (‘0’ for an inactive state and ‘1’ for an active state) to each gene in the network and model the interactions between them using boolean logic gates. This quantization of genes in binary space is justified because genes are either down-regulated or up-regulated in the majority of cellular processes [84]. When aberrations, such as those due to mutations, develop in controlled and well regulated biological processes such as apoptosis, cells can multiply uncontrollably and possibly form a tumor. We model these anomalies as faults in the network, where the mutated gene’s activity status is stuck at some value and is non-responsive to the inputs from its regulator genes.

Although this traditional approach to BN modeling has provided some degree of success with respect to biological relevance [85, 86], it is not well-suited for incorporating the feedback loops that often arise in a biological context. Hence, in order to accommodate this, we propose a modified boolean network. We now discuss the proposed modifications and then explain their benefits with the help of an example.

4.2.2 Modified Boolean Network

To date, we have used BNs to study the genes in a regulatory network that are abnormally up-regulated or down-regulated and have used this knowledge to establish the decisive targets that merit intervention. However, there are two major drawbacks with this classical approach.

First, this technique is incapable of distinguishing between the severity of two different gene mutations. To circumvent this drawback, we introduce the following rules:

- Rule 1: Each node in the network can take values in the positive integer set $\mathbb{Z}^+ \in \{0, 1, 2, \dots\}$ where '0' corresponds to the gene being down-regulated and the value $n > 0$ corresponds to n units of the gene product.
- Rule 2: The output of an OR gate is the sum of its inputs and the output of an AND gate is the minimum of its inputs, as shown in Figure 4.1.

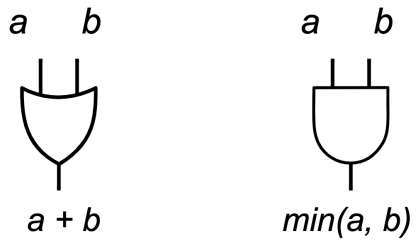


Figure 4.1: Modified rules of 'OR' and 'AND' logic gates. The output of an OR gate is the sum of its inputs ($\in \mathbb{Z}^+$) and the output of an AND gate is the minimum of its inputs. (Reprinted from [3])

The central idea of these rules is to not only qualitatively capture the up-regulation and down-regulation of genes occurring in the network but also to quantify their activity status. Let us elucidate this with the help of an example.

Consider a simple boolean network as shown in Figure 4.2a with possible faults occurring at F and G. With the conventional approach, the scenarios of a fault occurring at either of F or G and

that of faults at both F and G will produce the same output $J = 1$, thereby making the two scenarios indistinguishable from the output J. On the other hand, with the new rules incorporated, a fault occurring at either of F or G will return an output $J = 1$ whereas the simultaneous occurrence of faults at both F and G will produce the output $J = 2$. This increased output can possibly demonstrate enhanced proliferation and a faster-growing cancer.

The **second** drawback of classical Boolean network modeling stems from the fact that pivotal genes in pathways oversee and control cellular processes by constraining the upstream activators. This feedback necessitates a comparative approach where a gene applies brakes based on the difference between its abundance and the need for the particular gene product [87]. Clearly, the traditional approach of BN modeling fails to incorporate this. Once again, we shall illustrate this with the help of an example.

Consider a simple gene regulatory network (GRN) with 8 genes as shown in Figure 4.2b. Suppose gene A activates gene B, and genes A, B dimerize and stimulate gene E, genes B and E independently regulate gene C, gene E activates gene F which stimulates gene H and further dimerizes with C to form G. Additionally, let us assume that genes C and H negatively regulate genes A and D respectively through a feedback loop, gene E is mutated, and a drug inhibits gene F. Using the conventional approach, we can construct the boolean equivalent of this GRN as shown in

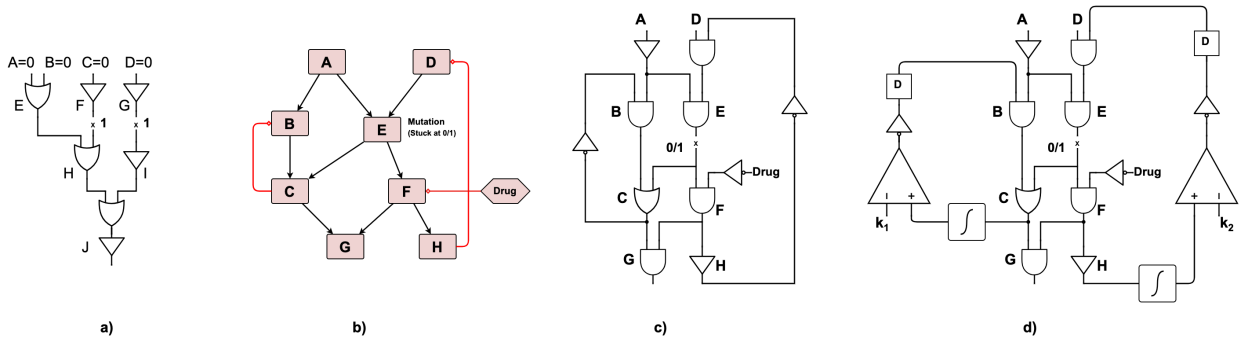


Figure 4.2: a) Example boolean network with possible faults occurring at F and G. b) Example gene regulatory network. c) The conventional boolean network of the example GRN. d) The Modified boolean network of the example GRN. (Reprinted from [3])

Figure 4.2c, but this network is missing the controlled feedbacks. Hence, in order to incorporate the controlled feedback discussed above, we modeled the feedback using an integrator, a comparator, and a delay block as shown in Figure 4.2d. Over time, the amount of gene products of genes C and H will accumulate and an integrator computes this and feeds it to a comparator that determines whether the brakes need to be applied. The delay block models the feedback delay that might occur during inhibition. Here, k_1, k_2 and the amount of delay are design parameters.

Now having understood the methodology and examined its benefits, we apply it in the context of lung cancer. First, we build the gene interaction network of lung cancer from the literature and then simulate it using the framework discussed.

4.3 Lung Cancer Pathways

Lung cancer develops through a multistage process involving the progression of multiple genetic aberrations. These abnormalities mainly occur in the three important sub-pathways, the PI3K/AKT/mTOR, the JAK/STAT, and the RAS/RAF/ERK which all connect and interact with each other [88].

The PI3K/AKT/mTOR pathway is a critical signal transduction pathway that is a key player in the regulation of proliferation, differentiation, and survival of cells [89]. Mutations in this pathway have been reported in various lung cancers. This pathway is activated downstream through tyrosine kinase receptors including epidermal growth factor receptor (EGFR), insulin-like growth factor 1 (IGF1), and receptor tyrosine-protein kinase (ERBB2) [90]. Activated receptor tyrosine kinases engage PI3K to phosphorylate PIP2 to PIP3 which in turn recruits the serine-threonine kinase AKT. AKT controls the expression of EGFR through a negative feedback. AKT also inhibits the tuberous sclerosis complex 1/2 (TSC1/2) which indirectly activates mTOR, a key manager of cell growth and metabolism. Adenosine monophosphate-activated protein kinase (AMPK) is an energy sensor in the cell which when activated by Metformin, a well known anti-diabetic drug, phosphorylates TSC1/2 which in turn inhibits mTOR [91]. Upregulated mTOR activates downstream ribosomal p70S6 kinase (RPS6KB1) which promotes growth signaling and regulates Insulin Receptor Substrate 1 (IRS1) through a negative feedback loop [92].

The Janus kinase (JAK)/signal transducer and activator of transcription (STAT) pathway plays the role of a fundamental block in immune control and gene transcription. Abnormal activation of the JAK/STAT pathway has been reported in multiple cancers. JAKs employ receptors and mediate phosphorylation of STAT3 [93].

Finally, the RAS/RAF/ERK pathway (MAPK pathway) is an intracellular pathway that is integral in the cellular proliferation, differentiation, survival, and apoptosis. When stimulated aberrantly, this pathway can induce tumorigenesis and has been linked with multiple malignancies [94]. EGFR is an important tyrosine kinase receptor involved in the induction of the MAPK pathway. RAS is a protein that is crucial for EGFR signaling whose mutations can activate downstream cascade despite the regulation of EGFR. RAF is a downstream protein of RAS which upon activation phosphorylates MEK and subsequently ERK [95]. The gene ERK promotes growth signaling and also regulates GRB2/SOS activation through a negative feedback loop [96].

In the literature, there is generally broad agreement among scientists about the specific locations of receptors/genes where different drugs intervene in a signaling pathway. We tabulate the list of drugs of interest to us and their targets in Table 4.1, along with the relevant references. The arrows in the parentheses represent whether a drug inhibits (↓) or activates (↑) its target(s).

Table 4.1: Drugs used and their gene intervention points. (Reprinted from [3])

Drug	Gene(s) targeted	References
HO-3867	STAT3 (↓)	[65]
LY294002	PIK3CA (↓)	[70]
Temsirolimus	mTOR (↓)	[67]
Metformin	AMPK (↑)	[97]
Cryptotanshinone	STAT3 (↓) + ERK (↓)	[71, 48]
Lapatinib	EGFR (↓) + ERBB2 (↓)	[68, 69]

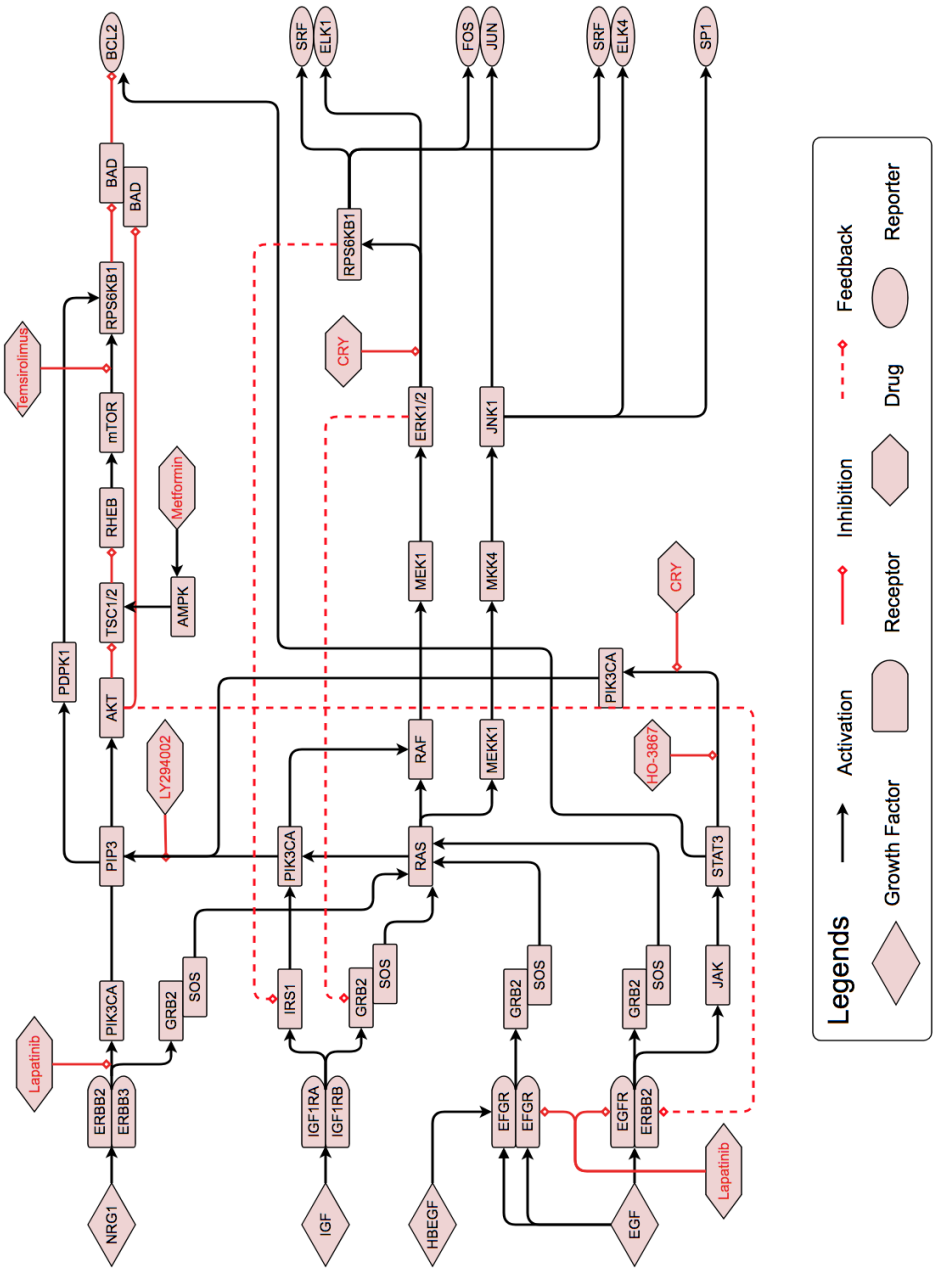


Figure 4.3: **Lung cancer signaling pathway.** A black arrow denotes activation, a red arrow denotes inhibition, and a dashed-red arrow denotes negative feedback. The legends explain the role of different bounding boxes. Growth factors are signaling proteins that promote cell-growth, survival, and differentiation. Receptors are proteins which bind to ligands such as growth receptors and cause responses in the immune system. They also play an important role in signal transduction and immunetherapy. Reporter genes are genes that help us in reporting expression levels and activity of important processes such as cell growth and apoptosis. (Reprinted from [3])

Using the above-discussed information available from the literature, we constructed the gene regulatory pathway of lung cancer as shown in Figure 4.3. The gene interactions are represented by arrows, where a black arrow denotes activation and a red arrow denotes inhibition. Using the methodology discussed earlier, we constructed the boolean equivalent of this gene regulatory network as shown in Figure 4.4. We now discuss the simulations followed by the theoretical and experimental results.

4.4 Results

4.4.1 Simulations

Utilizing the Boolean model constructed in Figure 4.4 and the methodology discussed, when the growth factors (EGF, HBEGF, IGF, NRG1) are present, the proliferation of cells measured using the genes SRF-ELK4, FOS-JUN, SP1, SRF-ELK1, and BCL2 is controlled with the help of the negative feedback loops present at AKT, ERK1/2, and RPS6KB1 genes. As discussed earlier, we modeled each of these negative feedback loops as a cascade of an integrator and a comparator. In Figure 4.4, k_1, k_2, and k_3 are model parameters which decide whether to apply the brakes or not. However, if a gene is mutated (over-expressed or under-expressed), the feedback loops can no longer keep the proliferation in check and this can possibly cause cancer. Hence, our goal here is to find the best drug combination that can mitigate the damaging effect of the majority of the abberations/faults. For our simulations, we chose $k_1 = k_2 = k_3 = 50$.

As discussed in the methodology section, each gene can assume a value in $\mathbb{Z}^+$ where '0' corresponds to the gene being down-regulated and the value $n > 0$ corresponds to n units of the gene product.

In case of inactive growth factors, all of EGF, HBEGF, IGF, and NRG1 are equal to 0, and in the network with no faults, this corresponds to all the output genes, SRF-ELK4, FOS-JUN, SP1, SRF-ELK1, and BCL2 equal to 0. However, in a network with faults present, the output genes will yield non-zero values.

Now, to assess the extent of abnormality in the network, we plot the sum of output genes'

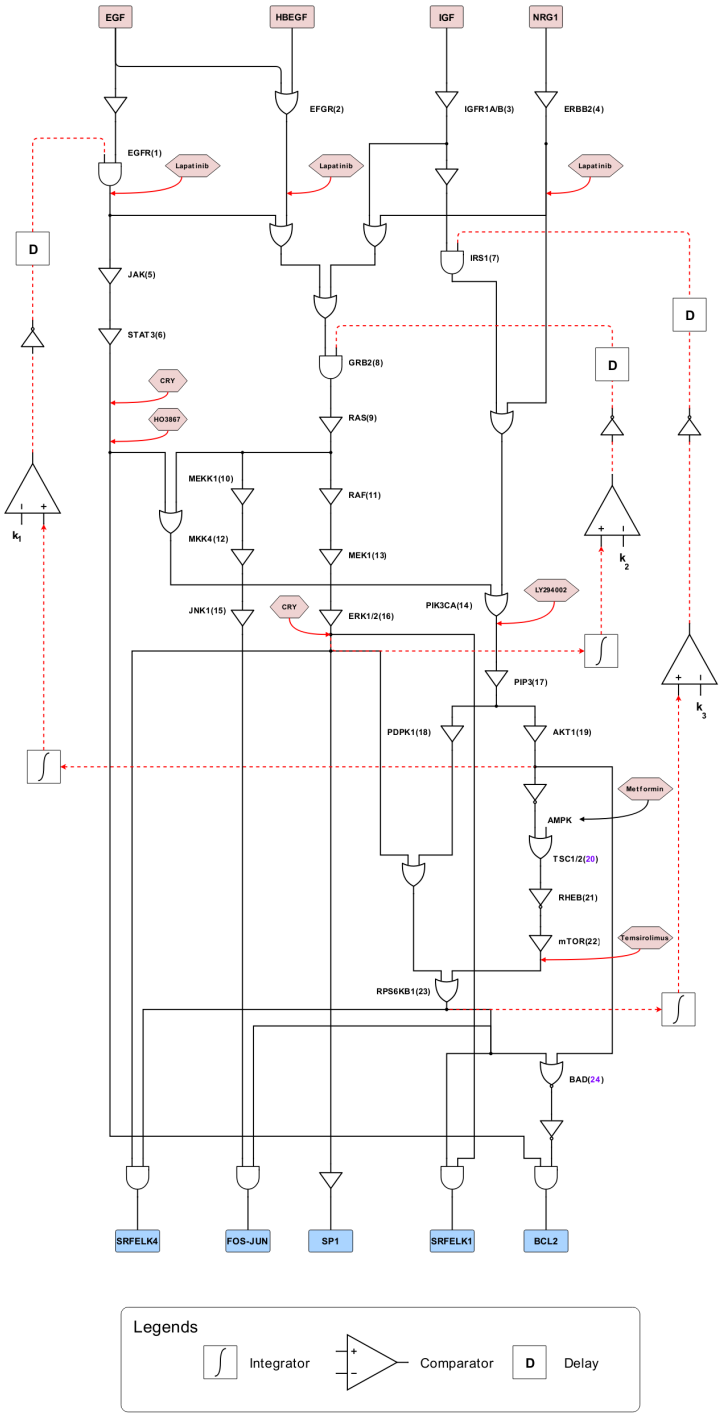


Figure 4.4: **The Modified Boolean equivalent of lung cancer pathway.** The numbers in parentheses represent the identifying number assigned to a fault at that location. Here, black numbers denote stuck-at-1 faults and blue numbers denote stuck-at-0 faults. (Reprinted from [3])

values over time and compute its Area Under Curve (AUC). Biologically, the AUC is comparable to the total number of cells produced in that time. Clearly, if the output genes' values are equal to 0, then the AUC in that scenario is equal to 0 and this corresponds to inactive cell proliferation. Since non-zero output genes' values correlate with a cancerous network, a higher AUC associates with greater cell proliferation and/or reduced apoptosis and possibly a higher risk of cancer.

We now simulate our lung cancer network across all possible faults and drug combinations and this will return a matrix (rows = faults, columns = drug combinations) of AUCs. For each fault, we compare the entries in the corresponding row, and the drug combination that matches the column with the smallest AUC yields the most desirable combination (for that fault). Similarly, in order to determine the most potent drug combination across all possible faults, we sum all the rows and select the column with the smallest value.

In this work, we also examined the existence of two faults, three faults, and four faults occurring simultaneously. Considering the harmful side-effects of drugs, in our experiments, we restricted ourselves to a maximum of three drugs per combination. Here, we provide the simulation results for at most three drugs per combination.

Using the method discussed above, we implemented the boolean network and simulated the model. We now present the theoretical results obtained followed by the experimental ones.

4.4.2 Theoretical Results

For our lung cancer model, we calculated the Area Under Curve for each combination of drug and fault. This arithmetic was performed for one, two, three, and four faults occurring simultaneously, and since there are 24 possible fault locations, we examined a total of ${}^{24}C_1 + {}^{24}C_2 + {}^{24}C_3 + {}^{24}C_4 = 12950$ combinations of faults. Furthermore, as explained above, to find the most dominant drug combination, we find an average AUC across all faults, and the smallest average AUC corresponds to the most favorable combination. In Table 4.2, we present the normalized (with no therapy as the reference) average AUC for each drug combination. Here, we present the values for at most four faults occurring simultaneously. From the table, it is evident that the bottom rows (27-42) involving Cryptotanshinone result in remarkable therapeutic success.

Table 4.2
The normalized average AUC obtained for each of the drug combinations. (Reprinted from [3])

	Drug combinations	Average AUC
1	Untreated	1.0000
2	Metformin	0.9988
3	HO-3867	0.8922
4	HO-3867 + Metformin	0.8904
5	Lapatinib	0.7849
6	Lapatinib + Metformin	0.7834
7	Lapatinib + HO-3867	0.7067
8	Lapatinib + HO-3867 + Metformin	0.7045
9	Temsirolimus	0.9951
10	Temsirolimus + Metformin	0.9951
11	Temsirolimus + HO-3867	0.8846
12	Temsirolimus + HO-3867 + Metformin	0.8846
13	Temsirolimus + Lapatinib	0.7767
14	Temsirolimus + Lapatinib + Metformin	0.7767
15	Temsirolimus + Lapatinib + HO-3867	0.6949
16	LY294002	0.9857
17	LY294002 + Metformin	0.9829
18	LY294002 + HO-3867	0.8834
19	LY294002 + HO-3867 + Metformin	0.8807
20	LY294002 + Lapatinib	0.7698
21	LY294002 + Lapatinib + Metformin	0.7666
22	LY294002 + Lapatinib + HO-3867	0.7018
23	LY294002 + Temsirolimus	0.9672
24	LY294002 + Temsirolimus + Metformin	0.9672
25	LY294002 + Temsirolimus + HO-3867	0.8696
26	LY294002 + Temsirolimus + Lapatinib	0.7468
27	Cryptotanshinone	0.2253
28	Cryptotanshinone + Metformin	0.2157
29	Cryptotanshinone + HO-3867	0.2253
30	Cryptotanshinone + HO-3867 + Metformin	0.2157
31	Cryptotanshinone + Lapatinib	0.1634
32	Cryptotanshinone + Lapatinib + Metformin	0.1597
33	Cryptotanshinone + Lapatinib + HO-3867	0.1634
34	Cryptotanshinone + Temsirolimus	0.2030
35	Cryptotanshinone + Temsirolimus + Metformin	0.2030
36	Cryptotanshinone + Temsirolimus + HO-3867	0.2030
37	Cryptotanshinone + Temsirolimus + Lapatinib	0.1450
38	Cryptotanshinone + LY294002	0.1498
39	Cryptotanshinone + LY294002 + Metformin	0.1340
40	Cryptotanshinone + LY294002 + HO-3867	0.1498
41	Cryptotanshinone + LY294002 + Lapatinib	0.1158
42	Cryptotanshinone + LY294002 + Temsirolimus	0.0771

For a better visual depiction, we plotted a heat map of AUCs for two faults occurring simultaneously for different drug combinations. We ran the simulations using the same parameters as provided in the codes online. In Figure 4.5, we have heat maps which are 24×24 matrices (for each of 24 faults) for three scenarios: Untreated, Temsirolimus + Lapatinib, and Cryptotanshinone + LY294002. The color in each cell represents the magnitude of AUC for that combination of two faults. Here, a color closer to red in the spectrum represents a higher AUC value and a color closer to green in the spectrum represents a lower AUC value. From the figure, the mutated pathway when treated with Temsirolimus + Lapatinib has a minimal effect, whereas, Cryptotanshinone + LY294002 shows promising therapeutic outcome. Further, we also provide the heat maps for all two-drug combinations in Figure 4.6.

We also plotted the sum of output genes' values for the fault-free network with active growth factors and the network with fault at ERK1/2 before and after it is treated with Cryptotanshinone in Figure 4.7. From the figure, the network without mutations is stabilized when growth factors are present. However, with a fault (at ERK1/2), the network is driven to an abnormally active state, and upon introduction of Cryptotanshinone, the growth is controlled. This mathematical output promises low cell-proliferation and/or enhanced apoptosis in cells when Cryptotanshinone is used.

4.4.3 Experimental Results

The theoretical results we obtained above were corroborated using experiments conducted on H2073 and SW900 lung cancer cell lines subjected to different drug treatments. We used a high-content fluorescent protein reporter imaging method and detected cell death in these cells. Then, using a well-known two-step data processing methodology, we extracted cell processing dynamics [50]. To demonstrate further, we condensed this collected data into expression profiles and plotted them.

The plots in Figure 4.8 demonstrate the cell killing produced in the H2073 lung cancer cell line using the intervention of different drug combinations. The black line denotes the untreated cell line which serves as a reference. Cryptotanshinone (CRY) has been used in each of the drug combinations and from the plots, it is apparent that in each instance, impressive cell death occurs

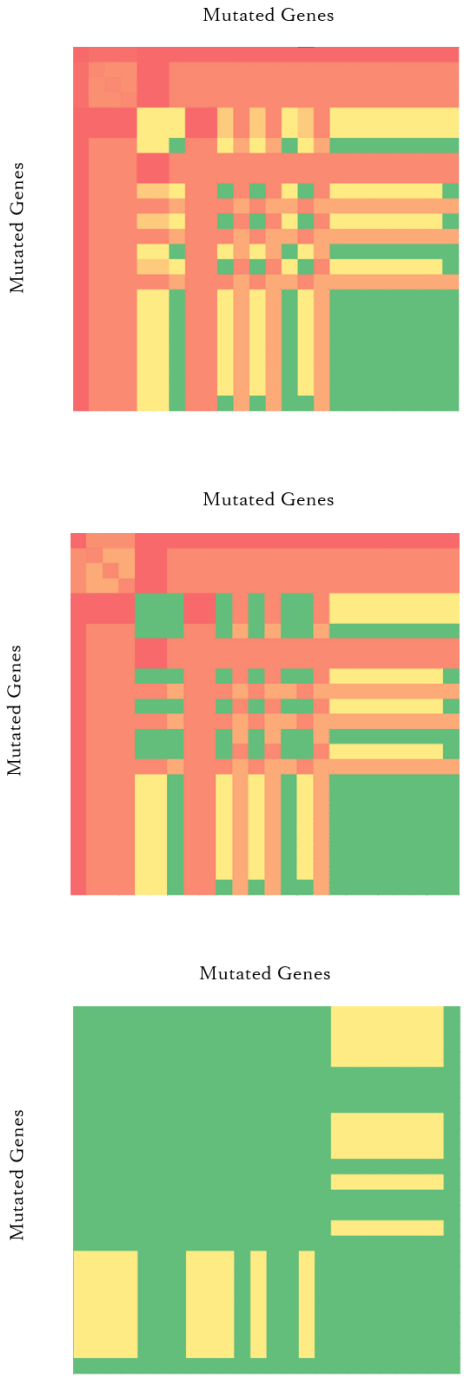


Figure 4.5: Heat map of AUC values for two faults occurring simultaneously for different drug combinations. The drug combinations (from top to bottom) are Untreated, Temsirolimus + Lapatinib, and Cryptotanshinone + LY294002. Here, a color closer to red in the spectrum represents a higher AUC value and a color closer to green in the spectrum represents a lower AUC value. (Reprinted from [3])

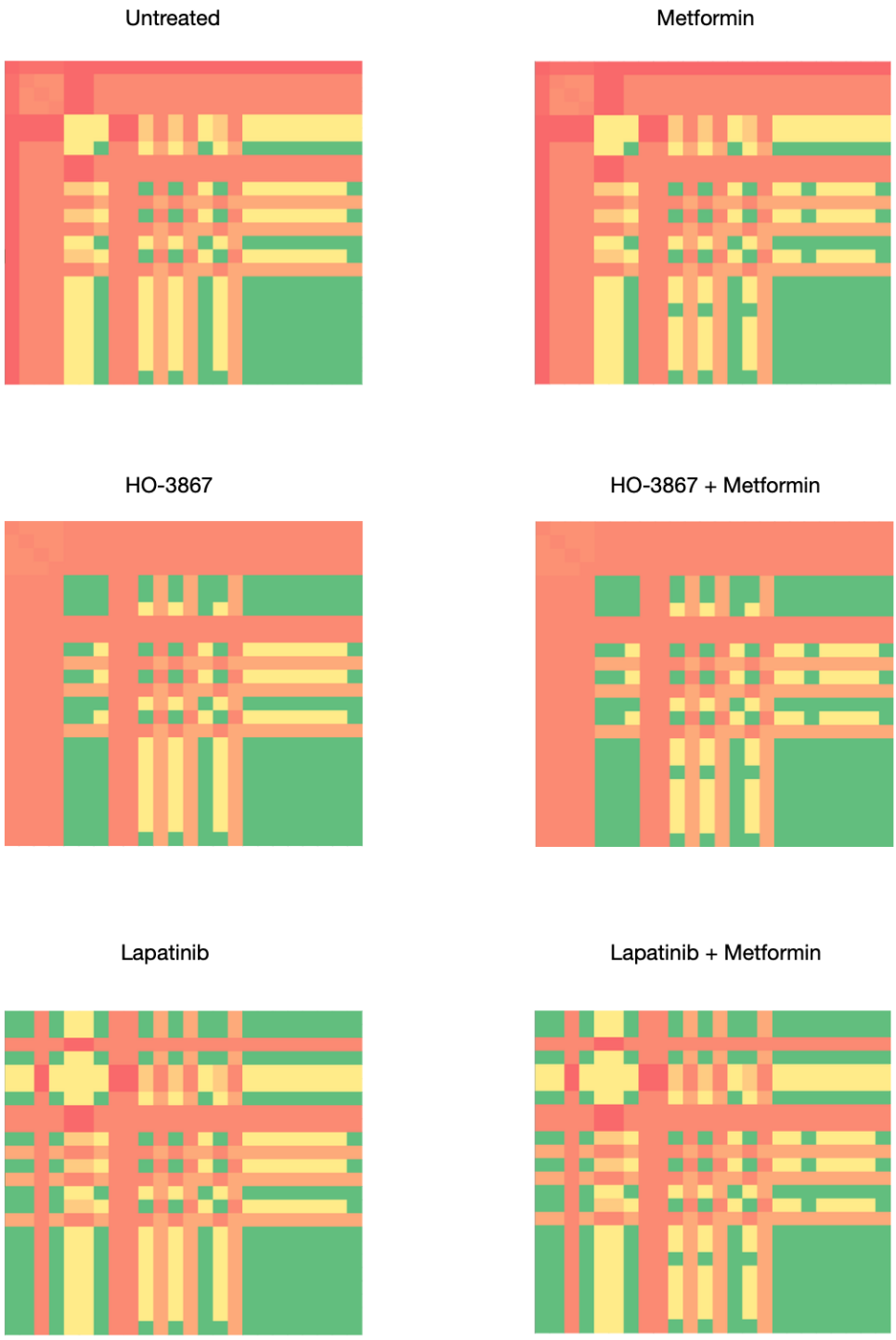


Figure 4.6: Heat map of AUC values for two faults occurring simultaneously for different drug combinations. Here, a color closer to red in the spectrum represents a higher AUC value and a color closer to green in the spectrum represents a lower AUC value. (Reprinted from [3])

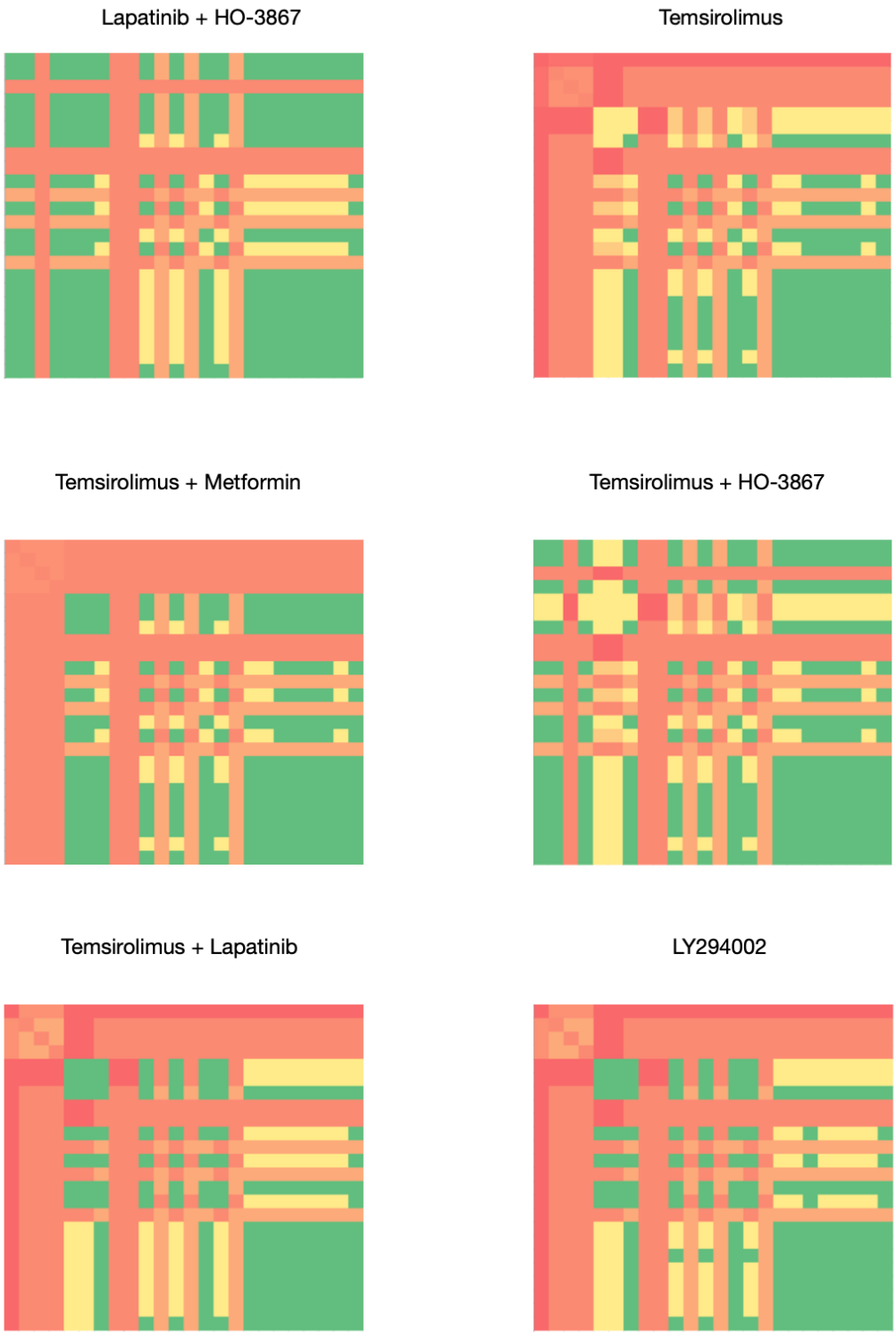
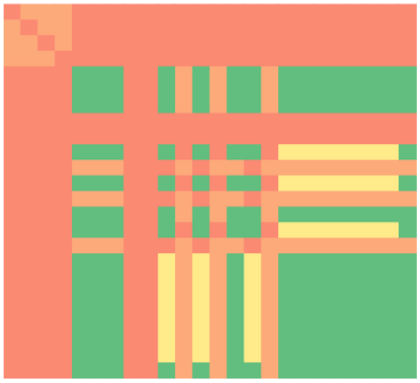
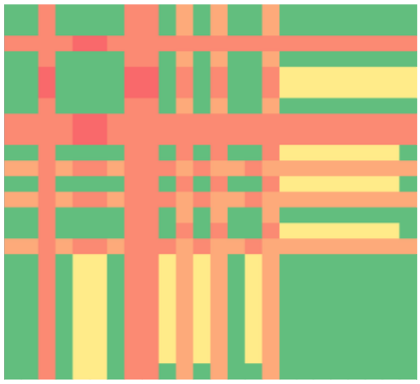


Figure 4.6 Continued

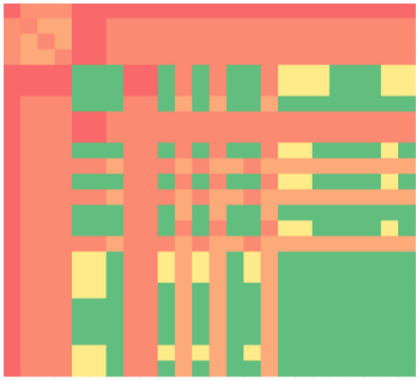
LY294002 + Metformin



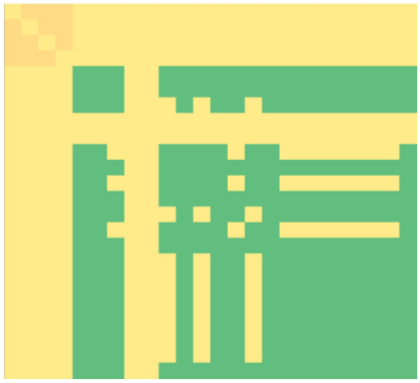
LY294002 + HO-3867



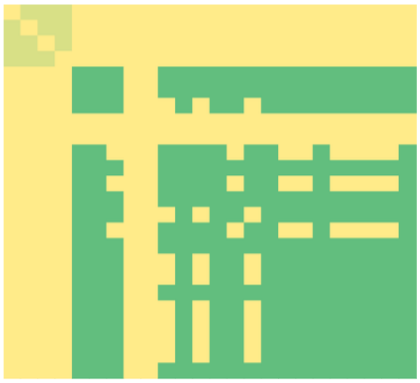
LY294002 + Lapatinib



Cryptotanshinone



Cryptotanshinone + Metformin

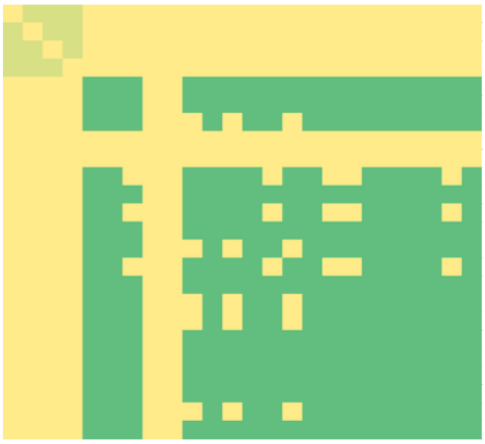


Cryptotanshinone + HO-3867

Figure 4.6 Continued

Cryptotanshinone + Lapatinib

Cryptotanshinone + Temsirolimus



Cryptotanshinone + LY294002

Figure 4.6 Continued

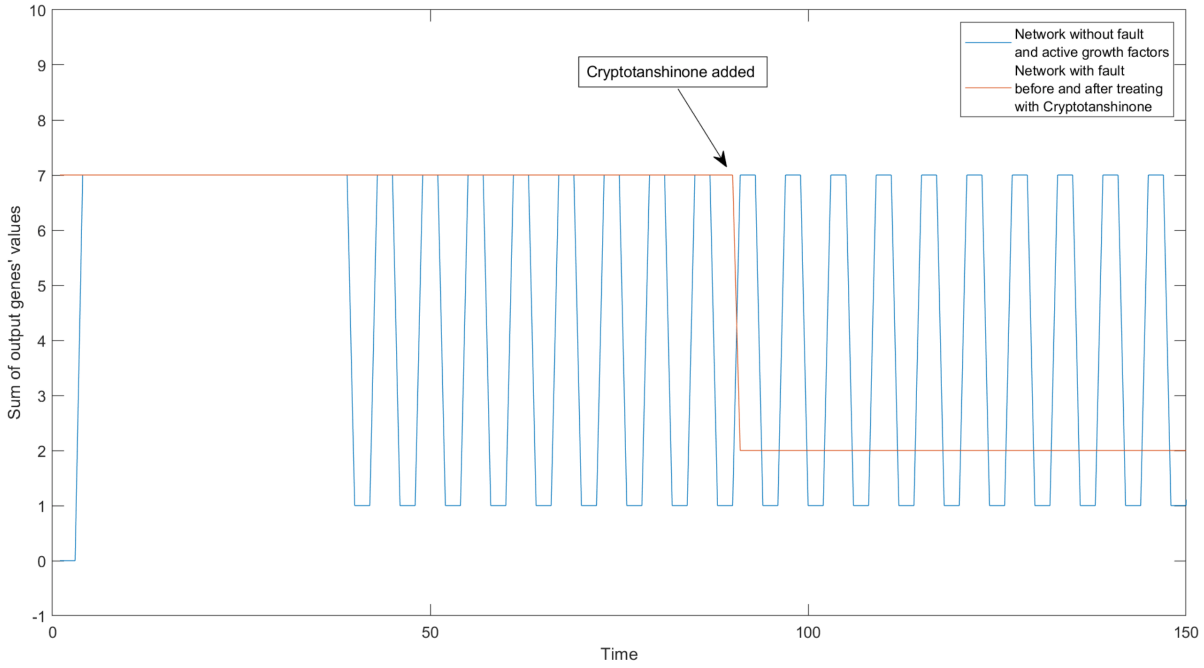


Figure 4.7: The plot of the sum of output genes' values for the fault free network (with active growth factors) and the network with fault at ERK1/2 before and after Cryptotanshinone is added. (Reprinted from [3])

and we have around 85% or more apoptosis in 24 hours. Hence, our computational predictions made using the modified boolean approach seem to be in line with the experimentally obtained ones.

In order to further confirm the efficacy of Cryptotanshinone, we carried out experiments with and without Cryptotanshinone on SW900 lung cancer cell line. From Figure 4.9, it is clear that the drugs (Metformin and HO-3867) are rather ineffective by themselves, but upon the addition of Cryptotanshinone in the mixture, we observe a remarkable increase in the efficacy of inducing cell death. These results further strengthen our argument that Cryptotanshinone substantially enhances cell death. As a side remark, we also note that the average AUC values of Metformin (compare rows 2, 28) and HO-3867 (compare rows 3, 29) from Table 4.2 are in line with our experimental results.

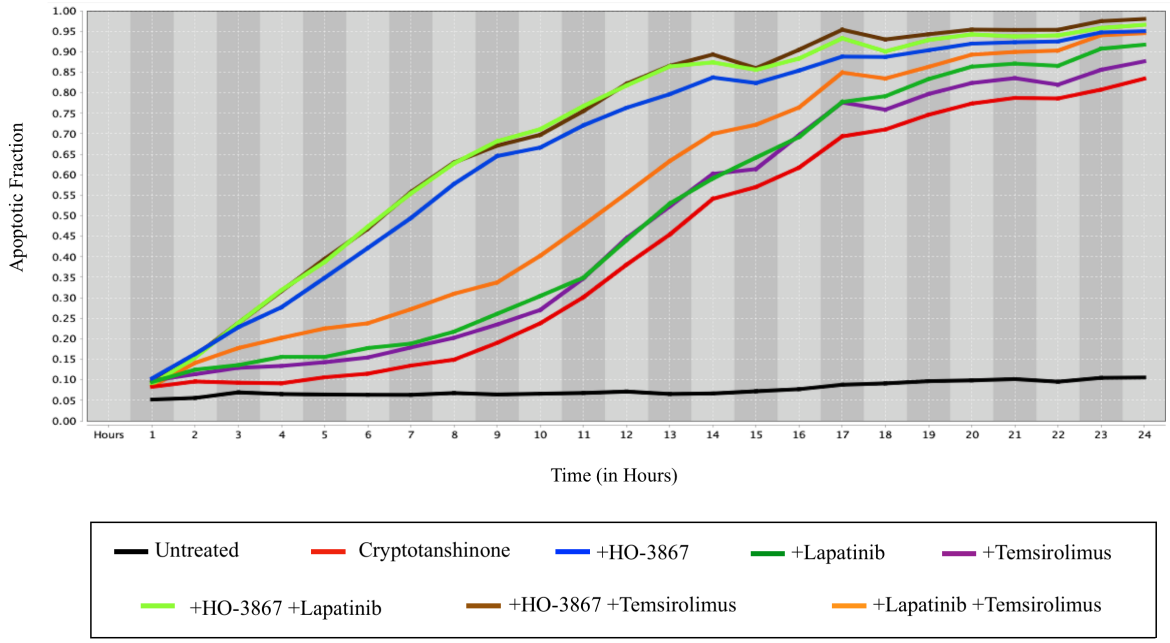


Figure 4.8: Apoptosis fraction versus time (in hours) for different drug combinations on H2073 cancer cell line. The drug combinations in the legend from left to right are Untreated cell line, Cryptotanshinone, Cryptotanshinone + HO-3867, Cryptotanshinone + Lapatinib, Cryptotanshinone + Temsirolimus, Cryptotanshinone + HO-3867 + Lapatinib, Cryptotanshinone + HO-3867 + Temsirolimus, and Cryptotanshinone + Lapatinib + Temsirolimus. (Reprinted from [3])

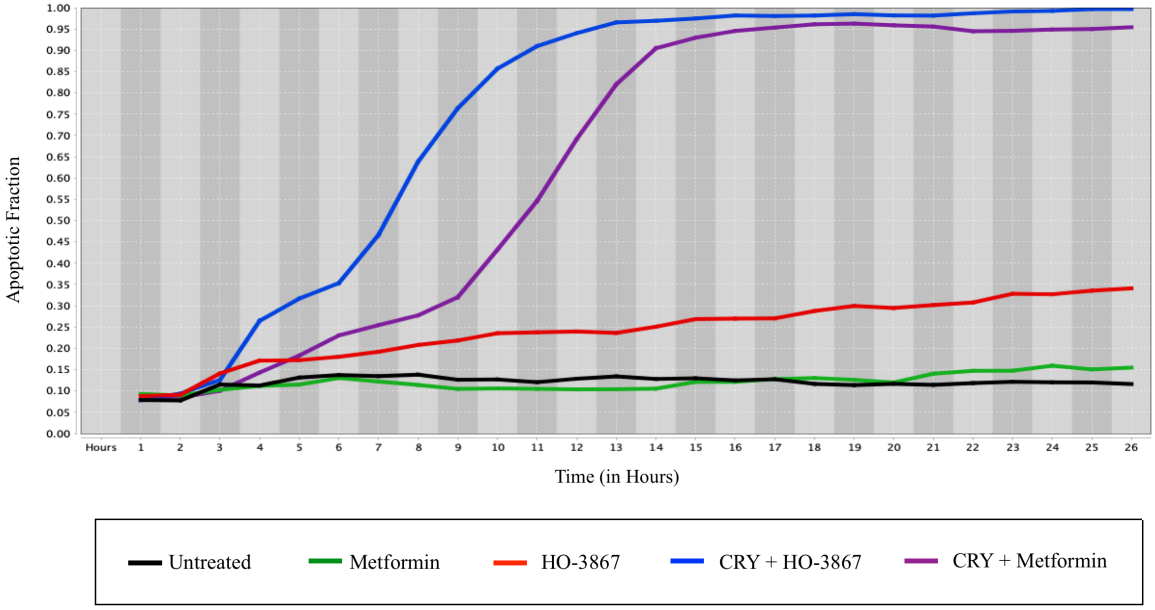


Figure 4.9: Apoptosis fraction versus time (in hours) for different drug combinations on SW900 cancer cell line. The drug combinations in the legend from left to right are Untreated cell line, Metformin, HO-3867, Cryptotanshinone + HO-3867, and Cryptotanshinone + Metformin. (Reprinted from [3])

5. DISCUSSION *

Biologists are always interested in identifying a few pivotal genes that can be controlled rather than a large group of genes, since it is both tedious and expensive to test the efficacy of the latter. In Section 2, we presented a mathematical framework to deduce the significant modulator genes in a biological signaling pathway and applied it to the Breast Cancer pathway. For a single gene intervention, our results showed mTOR as the most favourable target to achieve cell death. In the case of simultaneous intervention using two genes, a combination therapy targeting mTOR and STAT3 emerged as an outstanding modulator of cell death.

Most of the cancer treatments to date, for instance the drug Gleevec used to treat Chronic Myeloid Leukemia (CML), have utilized kinase inhibitors to keep cell proliferation in check. Although such a treatment does provide good results when the inhibitory action of the drug matches the particular mutation present, the success is usually short-lived as the cell figures out a mechanism to bypass the activity of that drug and/or additional mutations develop. On the other hand, if one could bring about a robust induction of cell death, independent of the particular mutations involved, then this might provide a very effective strategy for killing cancer cells. We believe that a robust modulator of this type should be capable of reinforcing its own activity via a positive feedback type of mechanism. This is similar to what happens in a servomechanism in control theory where to robustly achieve zero steady-state error in the presence of a persistent disturbance, one includes an *internal model* of the disturbance in the feedback loop [98, 99]. Although the idea of such an engineering approach to cancer therapy needs more thorough exploration, intuitively it is quite appealing as the imbalance between cell proliferation and cell death in a mature multicellular

*Parts of this section are reprinted with permission from H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "Bayesian Inference Identifies Combination Therapeutic Targets in Breast Cancer," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 9, pp. 2684-2692, 2019, doi: 10.1109/TBME.2019.2894980 © 2019 IEEE; and H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "In Silico Design and Experimental Validation of Combination Therapy for Pancreatic Cancer," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 3, pp. 1010-1018, 2020, doi: 10.1109/TCBB.2018.2872573 © 2020 IEEE; and H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "Cryptotanshinone Induces Cell Death in Lung Cancer by Targeting Aberrant Feedback Loops," *in IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 8, pp. 2430-2438, 2020, doi: 10.1109/JBHI.2019.2958042 © 2020 IEEE.

organism is what leads to tumorigenesis, and this imbalance could be considered to be the error signal that needs to be zeroed out to restore homoeostasis of the cell numbers.

Pancreatic cancer has a 5 year survival rate of 9%, whereas, that of breast cancer is about 90% [100]. This appallingly low rate of success is commonly attributed to its late detection and the existence of multiple activated pathways with cross-talk. With the possibility for so many different mutations, combination therapy appears to be a promising approach. However, since the potential number of drug combinations is large, and conducting biological experiments is complex and costly, an exhaustive method might not be a perfect strategy. Hence, it is crucial to devise mathematical models and methods that can curtail the search space.

In Section 3, we presented a Boolean framework to deduce the effective drug combinations in a biological signaling pathway. The approach discussed was then applied to the Pancreatic Cancer pathway. For the cancer pathway with three potential mutations, our results showed that Cryptotanshinone in combination with LY294002 resulted as the most favourable drug therapy to attain apoptosis.

Cancer is a disease characterized by unsupervised cell growth and it often progresses by the failure of the body's natural control system [101]. Using negative feedback loops, cells regulate proliferation, and a breakdown of this system leads to unchecked cell proliferation which may result in the formation of tumors. The primary reason for this uncontrolled growth is generally associated with mutations in genes, and diverse activated pathways with interference make the regulation additionally difficult. Hence, to simultaneously intervene in multiple pathways, combination therapy appears to be an attractive choice [102]. However, just with six drugs, the number of experiments to be conducted to decide the best combination is $2^6 - 1 = 63$, which is a prohibitively large number, both from the point of view of expense and the associated manual labor. Thus, we need to develop methods that can predict via simulations the combinations that are promising.

In Section 4, we presented a Modified Boolean model to theoretically infer the potent drug combinations to affect the time evolution of a biological network. We then applied the framework to the Lung Cancer pathway. Our results showed that Cryptotanshinone in itself or in combination

with other drugs resulted in significant improvement in terms of promoting apoptosis.

Three critical pathways, the JAK/STAT, the PI3K/mTOR, and the MAPK pathway interact with one another and play significant roles in cell growth, survival, and differentiation in several human cancers. The Phosphatidylinositol-3-kinase (PI3K)/Akt/mTOR pathway has been strongly implicated to play a key role in the promotion of cell survival and its alteration in multiple cancers and is associated with resistance to several types of therapy. In advanced cancers, the PI3K mutation rate can increase remarkably in different tumour types. In ovarian cancer, for example, PI3K/Akt/mTOR pathway is activated in a staggering 70% of the cases [103].

PI3K activates the downstream kinase AKT which activates the mTOR protein, an essential node in cell growth. As a serine/threonine kinase and a downstream member of the PI3K/Akt pathway, mTOR is an essential regulator of cell growth and its survival. PI3K is an enzyme that phospohorylates certain components of the cell membrane. Upon phosphorylation, these components bind to the protein Akt which becomes phosphorylated and activated. This upregulated Akt then activates mTOR, which promotes cell growth and proliferation by stimulating protein synthesis. In addition to receiving signals from Akt, mTOR keeps track of the environment of the cell for the presence of nutrients and growth factors. mTOR pathway has the capacity to control growth factor, estrogen-dependent and estrogen-independent, pathways which contribute to the pathological process and advancement of tumors. Hence, in recent years, a new approach in breast cancer therapy has been to obstruct this mTOR pathway. To inhibit this pathway, research has focused on developing mTOR inhibitors as therapeutic agents for patients with breast cancer. Pre-clinical reports back inhibition of this pathway, and stage I–III trials associating inhibitors of the mTOR pathway have been carried out in solid tumors. The drug Everolimus is a well known mTOR inhibitor, and recently, FDA approved it for the treatment of HR-positive, HER2-negative breast cancer [104].

Additionally, mutated PI3K pathway along with other activated pathways avert drugs from carrying out the inhibitory effects by devising an escape mechanism that leads to resistance. In the majority of cancers, one such pathway is the MAPK pathway (RAS/RAF/MEK/ERK cascade)

where abnormal KRAS activity prompts a cascade of up-regulated genes that contribute to the progression of cancer [105]. Abnormal KRAS activity has been shown to trigger many downstream signaling pathways and plays a decisive role in the development and progression of multiple human cancers including pancreatic cancer [106]. On account of these diverse activated pathways, a combination of agents targeting multiple steps of the intracellular machinery has a better chance of yielding a successful therapeutic outcome. In order to affect the PI3K pathway, we used LY294002, a well known PI3K inhibitor. LY294002 was shown to exclusively block the PI3K pathway in different cancers and indicated a good potential when used in combination therapy to fight gastric cancer [107].

The Janus kinase (JAK)/STAT signaling plays an important role in cell growth, differentiation and tumor invasion in diverse human cancers. This pathway is an epitome of biological interactions where extracellular factors regulate gene expression [108]. The STAT family of transcription factors integrate Cytokine and growth factor signaling to supervise a diverse array of cellular processes. STAT3, a member of the STAT protein family of transcription factors, forms dimers in the nucleus of cells and supervises the gene expressions of its target genes. The JAK/STAT pathway is mostly activated in all human cancers including breast, pancreatic, and lung cancers.

Under ordinary physiological habitat, STAT3 activation is tightly controlled, but strong evidence indicates that STAT3 is constitutively activated in many cancers and is a key transcription factor that is oncogenic in human cells [109]. STAT3 was recently found to have a key role in cultivating cancer stem cells in both in-vitro and in-vivo mouse tumor models. This result indicates its crucial involvement in tumor initiation, progression and maintenance [110]. In pancreatic cancer, activated STAT3 forms dimers, and this exhibited an advancement of pancreatic intraepithelial neoplasia growth and PDAC development [111]. Moreover, a down-regulation of STAT3 significantly abated the invasion ability of pancreatic cancer cells in vitro [112].

Further, aberrant STAT3 is expressed in about 55% of Non-small cell lung cancer (NSCLC) tumors and this evidence of STAT3's indispensable role in the initiation and progression of tumors makes it a pivotal target [113]. As a result, it is only natural to look for STAT3 inhibitors that can

obstruct the dimerization of STAT3 and restrict cell growth and differentiation. Recently, the drug Pyrimethamine, a STAT3 inhibitor was approved by the FDA. Khan [114] reported inhibition of STAT3 activity by Pyrimethamine in breast cancer cells.

Cryptotanshinone is one of the active components of Salvia miltiorrhiza Bunge and a naturally occurring compound derived from a traditional Chinese herb, and has been previously shown to possess anti-tumor properties in various types of human cancer cells [115]. Chen [116] recently showed that Cryptotanshinone suppressed the mTOR signaling pathway in MCF-7 breast cancer cell lines. Cryptotanshinone inhibited the expression of cyclin D1 thereby arresting cells in the $\mathrm{G}_1 - \mathrm{G}_0$ phase of the cell cycle and preventing proliferation.

Shin et. al. [48] corroborated the effectiveness of Cryptotanshinone on prostate cancer cell lines, where it inhibited STAT3 signaling through blocking its dimerization and decreasing the expression of its downstream target proteins. In pancreatic cancer cells, Ge et. al. [71] recently demonstrated that Cryptotanshinone inhibited proliferation and significantly induced apoptosis and cell cycle arrest via inhibition of the STAT3 pathway.

Cryptotanshinone has also been shown to decrease the expression of its downstream target proteins such as cyclin D1, survivin, and Bcl-xL. Evista (Raloxifene·HCl), another well known drug, is a selective estrogen receptor modulator, and is used for preventing osteoporosis and treating cases of ER positive breast cancer in postmenopausal women with high risk of invasive breast cancer. Shi [49] reported the inhibition of STAT3 phosphorylation by Evista in multiple cell lines including MCF-7. In the same vein our theoretical and experimental results demonstrated that Cryptotanshinone when used in combination undoubtedly boosts cell death [117].

In view of the preceding discussion, the literature on cancer signaling and Cryptotanshinone backs our computational result that Cryptotanshinone by itself and when used in combination is a promising drug in multiple cancers including breast, pancreatic, and lung cancers. The computational predictions made in this work agree with the past experimental results, thereby demonstrating the effectiveness of our models. We conclude that applying these methodologies to various biological signaling pathways could help the medical community in designing effective drugs,

without having to rely solely on conducting costly experiments. Finally, we believe that these findings can form a basis for the advancement of new and better methodologies for the drug design and treatment of other cancers.

REFERENCES

[1] H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "Bayesian inference identifies combination therapeutic targets in breast cancer," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 9, pp. 2684–2692, 2019.

[2] H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "In silico design and experimental validation of combination therapy for pancreatic cancer," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 17, no. 3, pp. 1010–1018, 2018.

[3] H. Vundavilli, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "Cryptotanshinone induces cell death in lung cancer by targeting aberrant feedback loops," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 8, pp. 2430–2438, 2019.

[4] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA: a cancer journal for clinicians*, vol. 68, no. 1, pp. 7–30, 2018.

[5] S. A. Kauffman, "Metabolic stability and epigenesis in randomly constructed genetic nets," *Journal of theoretical biology*, vol. 22, no. 3, pp. 437–467, 1969.

[6] O. A. Arshad and A. Datta, "Towards targeted combinatorial therapy design for the treatment of castration-resistant prostate cancer," *BMC bioinformatics*, vol. 18, no. 4, p. 134, 2017.

[7] T. Mestl, E. Plahte, and S. W. Omholt, "A mathematical framework for describing and analysing gene regulatory networks," *Journal of theoretical Biology*, vol. 176, no. 2, pp. 291–300, 1995.

[8] I. Shmulevich, E. R. Dougherty, S. Kim, and W. Zhang, "Probabilistic boolean networks: a rule-based uncertainty model for gene regulatory networks," *Bioinformatics*, vol. 18, no. 2, pp. 261–274, 2002.

[9] N. Friedman, M. Linial, I. Nachman, and D. Pe'er, "Using bayesian networks to analyze expression data," *Journal of computational biology*, vol. 7, no. 3-4, pp. 601–620, 2000.

[10] F. Liu, S.-W. Zhang, W.-F. Guo, Z.-G. Wei, and L. Chen, "Inference of gene regulatory network based on local bayesian networks," *PLoS computational biology*, vol. 12, no. 8, p. e1005024, 2016.

[11] N. Omranian, J. M. Eloundou-Mbebi, B. Mueller-Roeber, and Z. Nikoloski, "Gene regulatory network inference using fused lasso on multiple data sets," *Scientific reports*, vol. 6, p. 20533, 2016.

[12] C. M. Bishop, "Pattern recognition and machine learning (information science and statistics) springer-verlag new york," *Inc. Secaucus, NJ, USA*, 2006.

[13] G. Zweig and S. Russell, "Speech recognition with dynamic bayesian networks," 1998.

[14] P. S. Venkat, K. R. Narayanan, and A. Datta, "A bayesian network-based approach to selection of intervention points in the mitogen-activated protein kinase plant defense response pathway," *Journal of Computational Biology*, vol. 24, no. 4, pp. 327–339, 2017.

[15] H. Njah and S. Jamoussi, "Weighted ensemble learning of bayesian network for gene regulatory networks," *Neurocomputing*, vol. 150, pp. 404–416, 2015.

[16] N. X. Vinh, M. Chetty, R. Coppel, and P. P. Wangikar, "Gene regulatory network modeling via global optimization of high-order dynamic bayesian network," *BMC bioinformatics*, vol. 13, no. 1, p. 131, 2012.

[17] P. D. Hoff, *A first course in Bayesian statistical methods*. Springer Science & Business Media, 2009.

[18] A. Vazquez, E. E. Bond, A. J. Levine, and G. L. Bond, "The genetics of the p53 pathway, apoptosis and cancer therapy," *Nature reviews Drug discovery*, vol. 7, no. 12, p. 979, 2008.

[19] S. E. Shimony, "Finding maps for belief networks is np-hard," *Artificial Intelligence*, vol. 68, no. 2, pp. 399–410, 1994.

[20] J. Pearl, "Probabilistic reasoning in intelligent systems: Networks of plausible reasoning," 1988.

[21] R. E. Neapolitan *et al.*, *Learning bayesian networks*, vol. 38. Pearson Prentice Hall Upper Saddle River, NJ, 2004.

[22] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Elsevier, 2014.

[23] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[24] R. Garcia, T. L. Bowman, G. Niu, H. Yu, S. Minton, C. A. Muro-Cacho, C. E. Cox, R. Falcone, R. Fairclough, S. Parsons, *et al.*, "Constitutive activation of stat3 by the src and jak tyrosine kinases participates in growth regulation of human breast carcinoma cells," *Oncogene*, vol. 20, no. 20, p. 2499, 2001.

[25] P. Sansone and J. Bromberg, "Targeting the interleukin-6/jak/stat pathway in human malignancies," *Journal of Clinical Oncology*, vol. 30, no. 9, p. 1005, 2012.

[26] B. D. Fonseca, E. M. Smith, V. H.-Y. Lee, C. MacKintosh, and C. G. Proud, "Pras40 is a target for mammalian target of rapamycin complex 1 and is required for signaling downstream of this complex," *Journal of Biological Chemistry*, 2007.

[27] E. Paplomata and R. O'Regan, "The pi3k/akt/mtor pathway in breast cancer: targets, trials and biomarkers," *Therapeutic advances in medical oncology*, vol. 6, no. 4, pp. 154–166, 2014.

[28] A. S. Strimpakos, E. M. Karapanagiotou, M. W. Saif, and K. N. Syrigos, "The role of mtor in the management of solid tumors: an overview," *Cancer treatment reviews*, vol. 35, no. 2, pp. 148–159, 2009.

[29] A. R. Hussain, S. O. Ahmed, M. Ahmed, O. S. Khan, S. Al AbdulMohsen, L. C. Platanias, K. S. Al-Kuraya, and S. Uddin, "Cross-talk between nfkb and the pi3-kinase/akt pathway

can be targeted in primary effusion lymphoma (pel) cell lines for efficient apoptosis," *PLoS One*, vol. 7, no. 6, p. e39945, 2012.

[30] J. R. Testa and A. Bellacosa, "Akt plays a central role in tumorigenesis," *Proceedings of the National Academy of Sciences*, vol. 98, no. 20, pp. 10983–10985, 2001.

[31] P. J. Roberts and C. J. Der, "Targeting the raf-mek-erk mitogen-activated protein kinase cascade for the treatment of cancer," *Oncogene*, vol. 26, no. 22, p. 3291, 2007.

[32] G. Wilson, A. Cramer, A. Welman, F. Knox, R. Swindell, H. Kawakatsu, R. B. Clarke, C. Dive, and N. J. Bundred, "Activated c-src in ductal carcinoma in situ correlates with high tumour grade, high proliferation and her2 positivity," *British journal of cancer*, vol. 95, no. 10, p. 1410, 2006.

[33] M. M. Moasser, "The oncogene her2: its signaling and transforming functions and its role in human cancer pathogenesis," *Oncogene*, vol. 26, no. 45, p. 6469, 2007.

[34] T.-J. Kim, S. Mitsutake, and Y. Igarashi, "The interaction between the pleckstrin homology domain of ceramide kinase and phosphatidylinositol 4, 5-bisphosphate regulates the plasma membrane targeting and ceramide 1-phosphate levels," *Biochemical and biophysical research communications*, vol. 342, no. 2, pp. 611–617, 2006.

[35] A. Senchenkov, D. A. Litvak, and M. C. Cabot, "Targeting ceramide metabolism—a strategy for overcoming drug resistance," *Journal of the National Cancer Institute*, vol. 93, no. 5, pp. 347–357, 2001.

[36] D. Merino, S. Lok, J. Visvader, and G. Lindeman, "Targeting bcl-2 to enhance vulnerability to therapy in estrogen receptor-positive breast cancer," *Oncogene*, vol. 35, no. 15, p. 1877, 2016.

[37] K. Banerjee and H. Resat, "Constitutive activation of stat 3 in breast cancer cells: A review," *International journal of cancer*, vol. 138, no. 11, pp. 2570–2578, 2016.

[38] A. C. Faber, E. M. Coffee, C. Costa, A. Dastur, H. Ebi, A. N. Hata, A. T. Yeo, E. J. Edelman, Y. Song, A. T. Tam, *et al.*, "mtor inhibition specifically sensitizes colorectal cancers with

kras or braf mutations to bcl-2/bcl-xl inhibition by suppressing mcl-1," *Cancer discovery*, pp. CD–13, 2013.

[39] C. Sotiriou, P. Wirapati, S. Loi, A. Harris, S. Fox, J. Smeds, H. Nordgren, P. Farmer, V. Praz, B. Haibe-Kains, *et al.*, "Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis," *Journal of the National Cancer Institute*, vol. 98, no. 4, pp. 262–272, 2006.

[40] S. Loi, B. Haibe-Kains, C. Desmedt, F. Lallemand, A. M. Tutt, C. Gillet, P. Ellis, A. Harris, J. Bergh, J. A. Foekens, *et al.*, "Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade," *Journal of clinical oncology*, vol. 25, no. 10, p. 1239, 2007.

[41] Y. Tsujimoto, "Role of bcl-2 family proteins in apoptosis: apoptosomes or mitochondria?," *Genes to cells*, vol. 3, no. 11, pp. 697–707, 1998.

[42] D. Kim, I.-H. Lee, S. Kim, M. Choi, H. Kim, S. Ahn, P. E. Saw, H. Jeon, Y. Lee, and S. Jon, "A specific stat3-binding peptide exerts anti-proliferative effects and antitumor activity by inhibiting stat3 phosphorylation and signaling," *Cancer research*, pp. canres–2187, 2014.

[43] K. Selvendiran, L. Tong, A. Bratasz, M. L. Kuppusamy, S. Ahmed, Y. Ravi, N. J. Trigg, B. K. Rivera, T. Kálai, K. Hideg, *et al.*, "Anticancer efficacy of a difluorodiarylidenyl piperidone (ho-3867) in human ovarian cancer cells and tumor xenografts," *Molecular cancer therapeutics*, pp. 1535–7163, 2010.

[44] L. M. Witters, A. Witkoski, M. D. Planas-Silva, M. Berger, J. Viallet, and A. Lipton, "Synergistic inhibition of breast cancer cell lines with a dual inhibitor of egfr-her-2/neu and a bcl-2 inhibitor," *Oncology reports*, vol. 17, no. 2, pp. 465–469, 2007.

[45] Y. Hu, R. Guo, J. Wei, Y. Zhou, W. Ji, J. Liu, X. Zhi, and J. Zhang, "Effects of pi3k inhibitor nvp-bkm120 on overcoming drug resistance and eliminating cancer stem cells in human breast cancer cells," *Cell death & disease*, vol. 6, no. 12, p. e2020, 2015.

[46] T. Sadler, M. Gavriil, T. Annable, P. Frost, L. Greenberger, and Y. Zhang, "Combination therapy for treating breast cancer using antiestrogen, era-923, and the mammalian target of rapamycin inhibitor, temsirolimus," *Endocrine-related cancer*, vol. 13, no. 3, pp. 863–873, 2006.

[47] H. An, X. Yu, C. Xiang, Y. Zhang, J. Xia, and Y. Wang, "Baicalein and u0126 suppress human breast cancer cell line mcf-7 through regulating mapk signaling pathway," *Int J Clin Exp Pathol*, vol. 9, no. 10, pp. 10266–10273, 2016.

[48] D.-S. Shin, H.-N. Kim, K. D. Shin, Y. J. Yoon, S.-J. Kim, D. C. Han, and B.-M. Kwon, "Cryptotanshinone inhibits constitutive signal transducer and activator of transcription 3 function through blocking the dimerization in du145 prostate cancer cells," *Cancer research*, vol. 69, no. 1, pp. 193–202, 2009.

[49] W. Shi, D. Yan, C. Zhao, M. Xiao, Y. Wang, H. Ma, T. Liu, H. Qin, C. Zhang, C. Li, *et al.*, "Inhibition of il-6/stat3 signaling in human cancer cells using evista," *Biochemical and biophysical research communications*, vol. 491, no. 1, pp. 159–165, 2017.

[50] J. Hua, C. Sima, M. Cypert, G. Gooden, S. Shack, L. Alla, E. Smith, J. M. Trent, E. R. Dougherty, and M. L. Bittner, "Tracking transcriptional activities with high-content epifluorescent imaging," *Journal of biomedical optics*, vol. 17, no. 4, p. 046008, 2012.

[51] A. D. Seidman, C. A. Hudis, J. Albanell, J. Albanel, W. Tong, I. Tepler, V. Currie, M. Moynahan, M. Theodoulou, M. Gollub, *et al.*, "Dose-dense therapy with weekly 1-hour paclitaxel infusions in the treatment of metastatic breast cancer.," *Journal of Clinical Oncology*, vol. 16, no. 10, pp. 3353–3361, 1998.

[52] L. Rahib, B. D. Smith, R. Aizenberg, A. B. Rosenzweig, J. M. Fleshman, and L. M. Matrisian, "Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the united states," *Cancer research*, vol. 74, no. 11, pp. 2913–2921, 2014.

[53] C. S. Yabar and J. M. Winter, "Pancreatic cancer: a review," *Gastroenterology Clinics*, vol. 45, no. 3, pp. 429–445, 2016.

[54] J. Kleeff, M. Korc, M. Apte, C. La Vecchia, C. D. Johnson, A. V. Biankin, R. E. Neale, M. Tempero, D. A. Tuveson, R. H. Hruban, *et al.*, "Pancreatic cancer," *Nature reviews Disease primers*, vol. 2, p. 16022, 2016.

[55] J. N. Saultz and R. Garzon, "Acute myeloid leukemia: a concise review," *Journal of clinical medicine*, vol. 5, no. 3, p. 33, 2016.

[56] M. J. De Hoon, S. Imoto, K. Kobayashi, N. Ogasawara, and S. Miyano, "Inferring gene regulatory networks from time-ordered gene expression data of bacillus subtilis using differential equations," in *Biocomputing 2003*, pp. 17–28, World Scientific, 2002.

[57] R. Layek, A. Datta, M. Bittner, and E. R. Dougherty, "Cancer therapy design based on pathway logic," *Bioinformatics*, vol. 27, no. 4, pp. 548–555, 2010.

[58] G. Karlebach and R. Shamir, "Modelling and analysis of gene regulatory networks," *Nature Reviews Molecular Cell Biology*, vol. 9, no. 10, p. 770, 2008.

[59] H. De Jong, "Modeling and simulation of genetic regulatory systems: a literature review," *Journal of computational biology*, vol. 9, no. 1, pp. 67–103, 2002.

[60] R. Pal, A. Datta, A. J. Fornace Jr, M. L. Bittner, and E. R. Dougherty, "Boolean relationships among genes responsive to ionizing radiation in the nci 60 acds," *Bioinformatics*, vol. 21, no. 8, pp. 1542–1549, 2004.

[61] M. Korc, "Role of growth factors in pancreatic cancer.," *Surgical oncology clinics of North America*, vol. 7, no. 1, pp. 25–41, 1998.

[62] S. Eser, A. Schnieke, G. Schneider, and D. Saur, "Oncogenic kras signalling in pancreatic cancer," *British journal of cancer*, vol. 111, no. 5, p. 817, 2014.

[63] N. Bardeesy and R. A. DePinho, "Pancreatic cancer biology and genetics," *Nature Reviews Cancer*, vol. 2, no. 12, p. 897, 2002.

[64] A. De Luca, M. R. Maiello, A. D'Alessio, M. Pergameno, and N. Normanno, "The ras/raf/mek/erk and the pi3k/akt signalling pathways: role in cancer pathogenesis and implications for therapeutic approaches," *Expert opinion on therapeutic targets*, vol. 16, no. sup2, pp. S17–S27, 2012.

[65] Y. Hu, C. Zhao, H. Zheng, K. Lu, D. Shi, Z. Liu, X. Dai, Y. Zhang, X. Zhang, W. Hu, *et al.*, "A novel stat3 inhibitor ho-3867 induces cell apoptosis by reactive oxygen species-dependent endoplasmic reticulum stress in human pancreatic cancer cells," *Anti-cancer drugs*, vol. 28, no. 4, pp. 392–400, 2017.

[66] M. M. Javle, R. T. Shroff, H. Xiong, G. A. Varadhachary, D. Fogelman, S. A. Reddy, D. Davis, Y. Zhang, R. A. Wolff, and J. L. Abbruzzese, "Inhibition of the mammalian target of rapamycin (mtor) in advanced pancreatic cancer: results of two phase ii studies," *BMC cancer*, vol. 10, no. 1, p. 368, 2010.

[67] S. Chan, M. E. Scheulen, S. Johnston, K. Mross, F. Cardoso, C. Dittrich, W. Eiermann, D. Hess, R. Morant, V. Semiglazov, *et al.*, "Phase ii study of temsirolimus (cci-779), a novel inhibitor of mtor, in heavily pretreated patients with locally advanced or metastatic breast cancer," *Journal of Clinical Oncology*, vol. 23, no. 23, pp. 5314–5322, 2005.

[68] D. M. Walters, J. M. Lindberg, S. J. Adair, T. E. Newhook, C. R. Cowan, J. B. Stokes, C. A. Borgman, E. B. Stelow, B. T. Lowrey, M. E. Chopivsky, *et al.*, "Inhibition of the growth of patient-derived pancreatic cancer xenografts with the mek inhibitor trametinib is augmented by combined treatment with the epidermal growth factor receptor/her2 inhibitor lapatinib," *Neoplasia*, vol. 15, no. 2, pp. IN8–IN10, 2013.

[69] G. E. Konecny, M. D. Pegram, N. Venkatesan, R. Finn, G. Yang, M. Rahmeh, M. Untch, D. W. Rusnak, G. Spehar, R. J. Mullin, *et al.*, "Activity of the dual kinase inhibitor lapatinib (gw572016) against her-2-overexpressing and trastuzumab-treated breast cancer cells," *Cancer research*, vol. 66, no. 3, pp. 1630–1639, 2006.

[70] A. K. Gupta, G. J. Cerniglia, R. Mick, M. S. Ahmed, V. J. Bakanauskas, R. J. Muschel, and W. G. McKenna, "Radiation sensitization of human cancer cells in vivo by inhibiting the activity of pi3k using ly294002," *International Journal of Radiation Oncology• Biology• Physics*, vol. 56, no. 3, pp. 846–853, 2003.

[71] Y. Ge, B. Yang, Z. Chen, and R. Cheng, "Cryptotanshinone suppresses the proliferation and induces the apoptosis of pancreatic cancer cells via the stat3 signaling pathway," *Molecular medicine reports*, vol. 12, no. 5, pp. 7782–7788, 2015.

[72] L. Lu, C. Li, D. Li, Y. Wang, C. Zhou, W. Shao, J. Peng, Y. You, X. Zhang, and X. Shen, "Cryptotanshinone inhibits human glioma cell proliferation by suppressing stat3 signaling," *Molecular and cellular biochemistry*, vol. 381, no. 1-2, pp. 273–282, 2013.

[73] Y. Ge, R. Cheng, Y. Zhou, J. Shen, L. Peng, X. Xu, Q. Dai, P. Liu, H. Wang, X. Ma, *et al.*, "Cryptotanshinone induces cell cycle arrest and apoptosis of multidrug resistant human chronic myeloid leukemia cells by inhibiting the activity of eukaryotic initiation factor 4e," *Molecular and cellular biochemistry*, vol. 368, no. 1-2, pp. 17–25, 2012.

[74] V. Heinemann, "Gemcitabine: progress in the treatment of pancreatic cancer," *Oncology*, vol. 60, no. 1, pp. 8–18, 2001.

[75] J. Li, J. Kleeff, N. Giese, M. W. Büchler, M. Korc, and H. Friess, "Gefitinib ('iressa', zd1839), a selective epidermal growth factor receptor tyrosine kinase inhibitor, inhibits pancreatic cancer cell growth, invasion, and colony formation," *International journal of oncology*, vol. 25, no. 1, pp. 203–210, 2004.

[76] G. Fountzilas, M. Bobos, A. Kalogera-Fountzila, N. Xiros, S. Murray, H. Linardou, G. Karayannopoulou, A. K. Koutras, D. Bafaloukos, E. Samantas, *et al.*, "Gemcitabine combined with gefitinib in patients with inoperable or metastatic pancreatic cancer: a phase ii study of the hellenic cooperative oncology group with biomarker evaluation," *Cancer investigation*, vol. 26, no. 8, pp. 784–793, 2008.

[77] J. E. Larsen and J. D. Minna, "Molecular biology of lung cancer: clinical implications," *Clinics in chest medicine*, vol. 32, no. 4, pp. 703–740, 2011.

[78] A. Carracedo, L. Ma, J. Teruya-Feldstein, F. Rojo, L. Salmena, A. Alimonti, A. Egia, A. T. Sasaki, G. Thomas, S. C. Kozma, *et al.*, "Inhibition of mtorc1 leads to mapk pathway activation through a pi3k-dependent feedback loop in human cancer," *The Journal of clinical investigation*, vol. 118, no. 9, pp. 3065–3074, 2008.

[79] H. Kitano, "A robustness-based approach to systems-oriented drug design," *Nature reviews Drug discovery*, vol. 6, no. 3, p. 202, 2007.

[80] G. P. Figueredo, P.-O. Siebers, M. R. Owen, J. Reps, and U. Aickelin, "Comparing stochastic differential equations and agent-based modelling and simulation for early-stage cancer," *PloS one*, vol. 9, no. 4, p. e95150, 2014.

[81] A. Lahiri, P. S. Venkatasubramani, and A. Datta, "Bayesian modeling of plant drought resistance pathway," *BMC plant biology*, vol. 19, no. 1, p. 96, 2019.

[82] R. Pal, A. Datta, and E. R. Dougherty, "Robust intervention in probabilistic boolean networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1280–1294, 2008.

[83] I. Shmulevich, E. R. Dougherty, and W. Zhang, "Control of stationary behavior in probabilistic boolean networks by means of structural intervention," *Journal of Biological Systems*, vol. 10, no. 04, pp. 431–445, 2002.

[84] T.-H. Chung, M. Brun, and S. Kim, "Quantization of global gene expression data," in *2006 5th International Conference on Machine Learning and Applications (ICMLA'06)*, pp. 187–192, IEEE, 2006.

[85] R. S. Saraf, A. Datta, C. Sima, J. Hua, R. Lopes, and M. Bittner, "An in-silico study examining the induction of apoptosis by cryptotanshinone in metastatic melanoma cell lines," *BMC cancer*, vol. 18, no. 1, p. 855, 2018.

[86] O. A. Arshad, P. S. Venkatasubramani, A. Datta, and J. Venkatraj, "Using boolean logic modeling of gene regulatory networks to exploit the links between cancer and metabolism

for therapeutic purposes," *IEEE journal of biomedical and health informatics*, vol. 20, no. 1, pp. 399–407, 2016.

[87] S. L. Harris and A. J. Levine, "The p53 pathway: positive and negative feedback loops," *Oncogene*, vol. 24, no. 17, p. 2899, 2005.

[88] Y.-K. Yoon, H.-P. Kim, S.-W. Han, D. Y. Oh, S.-A. Im, Y.-J. Bang, and T.-Y. Kim, "Kras mutant lung cancer cells are differentially responsive to mek inhibitor due to akt or stat3 activation: implication for combinatorial approach," *Molecular Carcinogenesis: Published in cooperation with the University of Texas MD Anderson Cancer Center*, vol. 49, no. 4, pp. 353–362, 2010.

[89] V. Papadimitrakopoulou, "Development of pi3k/akt/mtor pathway inhibitors and their application in personalized therapy for non–small-cell lung cancer," *Journal of Thoracic Oncology*, vol. 7, no. 8, pp. 1315–1326, 2012.

[90] J. Gandhi, J. Zhang, Y. Xie, J. Soh, H. Shigematsu, W. Zhang, H. Yamamoto, M. Peyton, L. Girard, W. W. Lockwood, *et al.*, "Alterations in genes of the egfr signaling pathway and their relationship to egfr tyrosine kinase inhibitor sensitivity in lung cancer cell lines," *PLoS one*, vol. 4, no. 2, p. e4576, 2009.

[91] C. Fumarola, M. A. Bonelli, P. G. Petronini, and R. R. Alfieri, "Targeting pi3k/akt/mtor pathway in non small cell lung cancer," *Biochemical pharmacology*, vol. 90, no. 3, pp. 197–207, 2014.

[92] R. Vander Broek, S. Mohan, D. Eytan, Z. Chen, and C. Van Waes, "The pi 3 k/a kt/m tor axis in head and neck cancer: functions, aberrations, cross-talk, and therapies," *Oral diseases*, vol. 21, no. 7, pp. 815–825, 2015.

[93] Y. Hu, Y. Hong, Y. Xu, P. Liu, D.-H. Guo, and Y. Chen, "Inhibition of the jak/stat pathway with ruxolitinib overcomes cisplatin resistance in non-small-cell lung cancer nsclc," *Apoptosis*, vol. 19, no. 11, pp. 1627–1636, 2014.

[94] D. F. Calvisi, S. Ladu, A. Gorden, M. Farina, E. A. Conner, J.-S. Lee, V. M. Factor, and S. S. Thorgeirsson, "Ubiquitous activation of ras and jak/stat pathways in human hcc," *Gastroenterology*, vol. 130, no. 4, pp. 1117–1128, 2006.

[95] R. Hilger, M. Scheulen, and D. Strumberg, "The ras-raf-mek-erk pathway in the treatment of cancer," *Oncology Research and Treatment*, vol. 25, no. 6, pp. 511–518, 2002.

[96] A. Giubellino, T. R. Burke, and D. P. Bottaro, "Grb2 signaling in cell motility and cancer," *Expert opinion on therapeutic targets*, vol. 12, no. 8, pp. 1021–1033, 2008.

[97] Y. Zhuang and W. Miskimins, "Cell cycle arrest in metformin treated breast cancer cells involves activation of ampk, downregulation of cyclin d1, and requires p27 kip1 or p21 cip1," *Journal of molecular signaling*, vol. 3, no. 1, p. 18, 2008.

[98] B. A. Francis and W. M. Wonham, "The internal model principle of control theory," *Automatica*, vol. 12, no. 5, pp. 457–465, 1976.

[99] G. J. Silva, A. Datta, and S. P. Bhattacharyya, *PID controllers for time-delay systems*. Springer Science & Business Media, 2007.

[100] B. J. Kenner, S. T. Chari, A. Maitra, S. Srivastava, D. F. Cleeter, V. L. W. Go, L. J. Rothschild, and A. E. Goldberg, "Early detection of pancreatic cancer—a defined future using lessons from other cancers: a white paper," *Pancreas*, vol. 45, no. 8, p. 1073, 2016.

[101] S. Huang, I. Ernberg, and S. Kauffman, "Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective," in *Seminars in cell & developmental biology*, vol. 20, pp. 869–876, Elsevier, 2009.

[102] X. Xu, W. Ho, X. Zhang, N. Bertrand, and O. Farokhzad, "Cancer nanomedicine: from targeted delivery to combination therapy," *Trends in molecular medicine*, vol. 21, no. 4, pp. 223–232, 2015.

[103] M. L. Gasparri, E. Bardhi, I. Ruscito, A. Papadia, A. A. Farooqi, C. Marchetti, G. Bogani, I. Ceccacci, M. D. Mueller, and P. B. Panici, "Pi3k/akt/mtor pathway in ovarian cancer

treatment: Are we on the right track?," *Geburtshilfe und Frauenheilkunde*, vol. 77, no. 10, p. 1095, 2017.

[104] M. E. Royce and D. Osman, "Everolimus in the treatment of metastatic breast cancer," *Breast cancer: basic and clinical research*, vol. 9, pp. BCBCR–S29268, 2015.

[105] H. Ji, Z. Wang, S. A. Perera, D. Li, M.-C. Liang, S. Zaghlul, K. McNamara, L. Chen, M. Albert, Y. Sun, *et al.*, "Mutations in braf and kras converge on activation of the mitogen-activated protein kinase pathway in lung cancer mouse models," *Cancer Research*, vol. 67, no. 10, pp. 4933–4939, 2007.

[106] K. Giehl, B. Skripczynski, A. Mansard, A. Menke, and P. Gierschik, "Growth factor-dependent activation of the ras-raf-mek-mapk pathway in the human pancreatic carcinoma cell line panc-1 carrying activated k-ras: implications for cell proliferation and cell migration," *Oncogene*, vol. 19, no. 25, p. 2930, 2000.

[107] J.-Y. Shin, J.-O. Kim, S. K. Lee, H.-S. Chae, and J.-H. Kang, "Ly294002 may overcome 5-fu resistance via down-regulation of activated p-akt in epstein-barr virus-positive gastric cancer cells," *Bmc Cancer*, vol. 10, no. 1, p. 425, 2010.

[108] H. Yu and R. Jove, "The stats of cancer—new molecular targets come of age," *Nature Reviews Cancer*, vol. 4, no. 2, p. 97, 2004.

[109] H. Yu, H. Lee, A. Herrmann, R. Buettner, and R. Jove, "Revisiting stat3 signalling in cancer: new and unexpected biological functions," *Nature reviews Cancer*, vol. 14, no. 11, p. 736, 2014.

[110] A. Xiong, Z. Yang, Y. Shen, J. Zhou, and Q. Shen, "Transcription factor stat3 as a novel molecular target for cancer prevention," *Cancers*, vol. 6, no. 2, pp. 926–957, 2014.

[111] M. Lesina, M. U. Kurkowski, K. Ludes, S. Rose-John, M. Treiber, G. Klöppel, A. Yoshimura, W. Reindl, B. Sipos, S. Akira, *et al.*, "Stat3/socs3 activation by il-6 transsignaling promotes progression of pancreatic intraepithelial neoplasia and development of pancreatic cancer," *Cancer cell*, vol. 19, no. 4, pp. 456–469, 2011.

[112] H. dong Li, C. Huang, K. jian Huang, W. dong Wu, T. Jiang, J. Cao, Z. zhong Feng, and Z. jun Qiu, "Stat3 knockdown reduces pancreatic cancer cell invasiveness and matrix metalloproteinase-7 expression in nude mice," *PLoS One*, vol. 6, no. 10, p. e25941, 2011.

[113] R. Catlett-Falcone, W. S. Dalton, and R. Jove, "Stat proteins as novel targets for cancer therapy," *Current opinion in oncology*, vol. 11, no. 6, p. 490, 1999.

[114] M. W. Khan, A. Saadalla, A. H. Ewida, K. Al-Katranji, G. Al-Saoudi, Z. T. Giaccone, F. Gounari, M. Zhang, D. A. Frank, and K. Khazaie, "The stat3 inhibitor pyrimethamine displays anti-cancer and immune stimulatory effects in murine models of breast cancer," *Cancer Immunology, Immunotherapy*, vol. 67, no. 1, pp. 13–23, 2018.

[115] B.-Q. Wang, "Salvia miltiorrhiza: Chemical and pharmacological review of a medicinal plant," *Journal of Medicinal Plants Research*, vol. 4, no. 25, pp. 2813–2820, 2010.

[116] W. Chen, Y. Luo, L. Liu, H. Zhou, B. Xu, X. Han, T. Shen, Z. Liu, Y. Lu, and S. Huang, "Cryptotanshinone inhibits cancer cell proliferation by suppressing mammalian target of rapamycin–mediated cyclin d1 expression and rb phosphorylation," *Cancer Prevention Research*, pp. 1940–6207, 2010.

[117] L. Chen, H.-J. Wang, W. Xie, Y. Yao, Y.-S. Zhang, and H. Wang, "Cryptotanshinone inhibits lung tumorigenesis and induces apoptosis in cancer cells in vitro and in vivo," *Molecular medicine reports*, vol. 9, no. 6, pp. 2447–2452, 2014.