Nonparametric Simulation of Signal Transduction Networks with Semi-Synchronized Update

Isar Nassiri[1], Ali Masoudi-Nejad[1]*, Mahdi Jalili[2], Ali Moeini[3]

1 Laboratory of System Biology and Bioinformatics (LBB), Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran, 2 Department of Computer Engineering, Sharif University of Technology, Tehran, Iran, 3 Department of Algorithms and Computation, College of Engineering, University of Tehran, Tehran, Iran

Abstract

Simulating signal transduction in cellular signaling networks provides predictions of network dynamics by quantifying the changes in concentration and activity-level of the individual proteins. Since numerical values of kinetic parameters might be difficult to obtain, it is imperative to develop non-parametric approaches that combine the connectivity of a network with the response of individual proteins to signals which travel through the network. The activity levels of signaling proteins computed through existing non-parametric modeling tools do not show significant correlations with the observed values in experimental results. In this work we developed a non-parametric computational framework to describe the profile of the evolving process and the time course of the proportion of active form of molecules in the signal transduction networks. The model is also capable of incorporating perturbations. The model was validated on four signaling networks showing that it can effectively uncover the activity levels and trends of response during signal transduction process.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: amasoudin@ibb.ut.ac.ir

Introduction

Cells sense, monitor and process signals originating from the environment. Monitoring external conditions requires the signal transduction to the site of action and thereby triggers various biological responses. Cells have developed signal transduction pathways, which facilitate signal transmission from the receptors to the target molecules by cascades of modifications to cellular molecules such as phosphorylation [1]. Evolving proteomic approaches to network biology have focused on measuring the changes in abundances of signaling proteins in active forms (e.g. phosphorylated form) under different experimental conditions [2]. Examples of such protein-based datasets becoming more frequent in the literature [3,4,5,6,7]. At the present, this type of studies on protein functional status focuses on frequent sampling of a limited number of key molecules [4,8]. This necessitates developing mathematical methods to prioritize selection of the molecules such that measuring their activity to be informative and capable of predicting the outcome of new experiments [9,10].

Evidently, it is important to apply adequate updating rules to choose biologically correct model [11]. During the signal transduction process, always time delays are associated with the transport of signaling molecules to reach the site of action [12]. For example, the transport of signaling molecules over intracellular distances of more than a few micrometers normally requires facilitated transport mechanisms such as movement of phosphorylated kinases on the scaffolds [13,14,15]. The other issue is the stochasticity in signaling processes at the molecular level. The origin of stochasticity can be due to chaotic births and deaths of individual molecules as well as the extracellular noise due to environmental fluctuations [16]. Such delays and stochastic noises

are source of inherent fluctuations in transferred signal [17]. Generally, asynchronous and continuous models, in which all edges demonstrate different efficiencies for making signal transition, are closer to biological phenomena [18].

A number of studies have shown close relation between structure of biological networks and their functional phenotypes [19,20]. Therefore, scholars struggled to develop nonparametric methods that are based on only network topology [9,10,21,22]. However, no model has been reported to compute both the proportion of active form and the trend of activation of given molecule in signal transduction networks [9,10]. The present study proposes an efficient nonparametric method for accurate identification of activity trends and the proportion of active form for each member of the signaling networks. We validated our approach for logical modeling on four signaling networks. We show that activity level and activation trend of signaling proteins observed through the proposed model have significant correlations with experimental results.

Materials and Methods

Signaling networks are represented as directed graphs where nodes denote signaling components and edges represent the direction of information flow. Edges are labeled with positive (+1) or negative (−1) signs, which define activation or inhibition, respectively. The input (source) nodes represent the ligands or their receptors, the intermediate nodes consist of various kinases and second messengers, and the output (sink) nodes represent transcription factors, channels, cytoskeleton, motility components, or cellular responses [23]. Through the interactions in the

network, signals propagate from the receptor (source node) to the downstream and target molecules (sink nodes).

The proposed dynamic model follows the changes in the activity levels of signaling proteins (between 0 and 1) in time steps and the way that signals propagate through molecular interactions [24]. Thus, in some sense, the predictions returned by our simulator can be interpreted like the normalized results of bead-based micro-ELISA assay [4,25]. In the following sections, we describe different parts of our method to simulate signaling networks.

Edge weighting

Matching connection is a measure used to establish similarity between two nodes (proteins) [26,27]. If nodes i and j have a number of common neighbors, they may be related to each other, even though they are not directly connected [28,29,30]. Here in order to quantify the efficiency of signal transduction between the two nodes adjacent to the given edge, we used similarity index and took into account the existence of matching connections. The extension is called Normalized Similarity Index (NSI). Let us consider a directed network with adjacency matrix A, and n nodes that are denoted from 1 to n. Let there also be a $n \times n$ similarity matrix $\mathbf{S} = (NSI_{ij})$ satisfying $0 < NSI_{ij} \leq 1$, $NSI_{ii} = 0$, and $NSI_{ij} = NSI_{ji}$ for all $i, j \in \{1, …, n\}$. The NSI of edge between nodes i and j is computed considering the number of matching connections of i and j to any other node m ($A_{i,m}A_{j,m}, A_{i,m}A_{m,j}, A_{m,i}A_{j,m}, A_{m,i}A_{m,j}$), and the number of connections (edges) between the node i and j (C_{ij}). More precisely,

$$NSI_{ij} = \frac{\sum_{m \neq i,j} (A_{i,m}A_{j,m}, A_{i,m}A_{m,j}, A_{m,i}A_{j,m}, A_{m,i}A_{m,j}) + C_{i,j}}{k_i + k_j - \left[\sum_{m \neq i,j} (A_{i,m}A_{j,m}, A_{i,m}A_{m,j}, A_{m,i}A_{j,m}, A_{m,i}A_{m,j}) + C_{i,j} \right]} \quad (1)$$

where k_i is degree of node i. Figure 1 shows a graphics for computing NSI. Each edge is associated with a weight (NSI_{ij}) ranging from 0 to 1 that represents the efficiency of signals passing along the edge. Edge weighting by NSI suggests that inside the highly connected parts of network, the efficiency of signal propagation along the edges is mainly, due to the facility posed by high NSI values. Data from previous studies supported the idea that cellular signals are transmitted dominantly through pathways of highly connected proteins [31].

The similarity index is often used to cluster different components of biological networks that are functionally similar [32,33]. The definition draws upon the notion that two molecules (nodes) could be considered similar if they perceive the rest of the nodes within the network in a similar way [26,27,34]. Previous similarity indices are based on the number of common neighbors shared by two non-adjacent nodes i and j [28,32,33,34]. However, NSI quantifies the similarity between two connected nodes. The VOSveiwer program was used for visulization of weighted networks by similarity measure [35].

Simulation of signal transduction and weighting the nodes

The simulation starts from signal ligand and iteratively traverses the whole network by a breadth-first-search (BFS). As we are only interested in the connected sub-network that begins with the first node, the nodes that are not reachable from the first node are simply ignored. The state of a node j, (X_j), is defined as a proportion of molecules in active form (e.g. phosphorelated) and gets a value between 0 and 1. At each time interval t, the $X(t)$ is changed or updated based on its value in the previous time step, $X(t-1)$, the statuses of the set of activator ($S^A_{i \rightarrow j}$), and those of
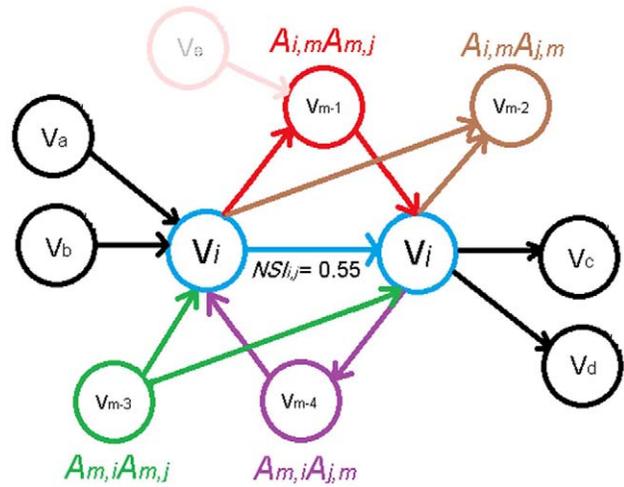
Figure 1. Weighting of the edge by Normalize Similarity Index (NSI). V_i and V_j are connected with one directed edge ($C_{ij} = 1$). Node V_i and V_j are also connected with four common node ($V_{m-1}, V_{m-2}, V_{m-3}, V_{m-4}$), and in total with eight distinct neighbors ($V_{m-1}, V_{m-2}, V_{m-3}, V_{m-4}, V_a, V_b, V_c, V_d$). The NSI will then be $NSI_{ij} = 5/9 = 0.55$.
doi:10.1371/journal.pone.0039643.g001

repressor nodes feeding into it ($S^R_{(i \rightarrow j)}$) according to the following formula:

$$X_j(t) = \left[X_j(t-1) \times (0.7 \leq R \leq 1) \right] + \\ \left[1 - \Pi_{i \in S^A}(1 - S^A_{ij} \times NSI^A_{ij}) \right] \times \\ \Pi_{i \in S^R}(1 + S^R_{ij} \times NSI^R_{ij}) \times [1 - X_j(t-1)] \quad (2)$$

where, each directed edge (i, j) models the transduction of signals from node i to node j of activation ($NSI^A_{ij} = NSI_{ij}$) or repressor ($NSI^R_{ij} = -NSI_{ij}$) type. In order to explain clearly the equation 2, let $X = (X_A, X_I) = (\{\#\text{Active form}\}, \{\#\text{Inactive form}\})$ be the states of a given molecule at time t. If $P(X_A, t)$ is the probability function for the state ($\{\#\text{Active form}\}) = X_A$ at time t, in order to simplifying the expression we can use $P(X_A, t) = X(t)$. Therefore, $X(t)$ is the probability or proportion of the active form of given molecule at time t.

Although we know a great deal about global protein stability profiling, little is known about global stability of signaling proteins in the active form (e.g. phosphorylated) [36,37]. In the proposed model, the relative stability (R) of proteins in active form is chosen randomly and changed after each time step ($X_j(t-1) \times R$) [20]. We adopted equation 2 according to the method proposed by *Van Kampen* for selecting the subset of molecules in the active form in preceding iteration to be passed to next iteration [38]. First, we developed a deterministic equations of the experimental system, and then added a noise term ($0 \leq R \leq 1$), and adjusted a specified random function to reproduce a correct mean as taken from the experimental results ($0.7 \leq R \leq 1$). The noise term is independent and uncorrelated with the time span of the system. We used $[1 - X(t-1)]$ in formula to ensure that the activation of a node has bounded in the interval [0–1], regardless of the number of iterations.

In the model, time is quantized into regular intervals (time-steps) as the longest duration required for state of all components to be updated [39]. We extended this basic model to account for variability in the duration of signal transmission by performing

signal transduction to some nodes in an asynchronous update fashion, called semi-synchronous update [20,21,40]. If a directed network is imaged in hierarchical layout, in semi-synchronous update, each node in layer d at time-step t receive signals from its regulators in layer d and layer d-1 in t, and its regulators in downstream of layer d in t+1. Therefore, the state of a node in each time step depends on the signals transmitted to the nodes at time-step t or t-1. The rules for semi-synchronous algorithm can be written as:

$$X_j(t) = f_j\Big(X_j(t-1), X_{i_1,j}(m), X_{i_2,j}(m), \ldots, X_{i_k,j}(m)\Big),$$
$$m \in \{t-1, t\} \tag{3}$$

where $X_j(t)$ is the state of node j at time-step t, f_j is the function associated with state of node j at time-step t-1 ($X_j(t-1)$) and its regulators ($X_{i_k,j}$) in previous time step ($t-1$) or current round of updates t.

Equation 2 is derived from a specific form of non-linear dynamic representation of the regulatory networks of transcription. It is assumed that all transcription factors independently bind at the available distinct sites of promoter region without interacting with each other [41]. In updating the nodes, we also assumed that the order of incoming signals does not affect the outcome, i.e., $\big[f_j\big(X_j(t-1), X_{i_1,j}, X_{i_2,j}\big) = f_j\big(X_j(t-1), X_{i_2,j}, X_{i_1,j}\big)\big]$ in Equation (3).

In the first step of simulation, user can specify the activity of first node (proportion of active form) as a source of signal. The activity of first node (w_f) changes systematically according to the following relation:

$$w_{f,\,t} = \left(t \times \frac{w_f}{T \times 10} \right) \tag{4}$$

where, T is a user-specified number of iterations, w_f is a user-specified activity of first node, and w_{f,t_n} is the weight of first node in iteration t. The term 10 in the denominator confines the proportion of receptors in active form (binds its ligand), that is less than 1, because almost 10% of the total receptors pool are usually in high-affinity form [42,43]. The activities of other nodes are initially equal to 0.

As the main focus of this study was designing a purely non-parametric simulation method, we used network topology to predict the half-lives of proteins. Wang et al. (2009) systematically examined the entire human signaling network and mapped the enrichment of proteins with different half-lives in the groups of nodes that were categorized based on their degree in network and cellular location [44]. According to their results, 77.5% of the nodes with degree equal to or bigger than 6, 73.6% of nuclear proteins, and 48.9% of ligands had long half-lives. Therefore, for the nodes with a degree equal to or bigger than 6, and out-degree (sink nodes) or in-degree (source nodes) equal to 0, we defined two kinds of proportion of activity: 1- if node has short half-life and 2- if it has long half-life [44,45].

Simulation of perturbation experiments

Typically, the functional importance of a molecule in signal transduction network is determined by cell response after its inhibition by specific chemical inhibitors or interference with the siRNA [4]. In many existing methods, molecular inhibition is simulated by deleting the corresponding nodes and edges incident on it [46,47]. However, the disruption of any molecule in cell signaling network may lead to a cascade breakdown of other

downstream molecules. We briefly considered some examples from the perturbation results of the phosphoproteins in the inflammatory and growth signaling network in human hepato-blastoma cells [4]. Figure 2 depicts the inflammatory and growth signaling network in human hepatocyte cells. The experimental results are plotted in Figure 3, which is organized into panels, corresponding to the various ligands [4,48]. If we simulate molecular inhibition of MEK and IKK by deleting the corresponding nodes and the edges incident on them, the signal flow to the ERK and IκB is broken (Figure 2). However, according to the experimental results, treatment of MEK and IKK with their corresponding inhibitors and the measurements of the activity of their downstream proteins (ERK and IκB) have shown that a little signal flow can pass through the inhibited node, and the signals sent by upstream nodes can have inductive effect on its target molecules (Figure 3A and 3B) [4].

To take into account the gain of function or inhibitory effects on the molecule(s), we used a complementary node for target node [39,49,50]. To simulate the influence of the inhibitory effect on the signaling component, target node receives constant inhibitory messages from the complementary node during simulation; however, signal flow can pass through the inhibited node in some iteration. When the measured molecule is inhibited (e.g. MEK), its measurement cannot be used in experimental studies (Figure 3C) [4]. Therefore, in our method the proportion of active form of inhibited node is denoted with the input value specified by the user (often 0).

Overview of the tool

The Nonparametric Simulator of Signaling Networks (NSIN) program was written in C^{++}, and is freely available at http://lbb.ut.ac.ir/Download/LBBsoft/NSIN. The inputs to NSIN are text files specifying a directed graph in which each node represents a specific molecule (Input A), and each directed edge models the transduction of signals of activation (+1) or blockage (−1) type (Input B). The simulator provides both single and set of running modes. In single mode, users can specify the activity of the input node (between 0 and 1), while the activities of the other nodes change as a function of initial stimulus. A set-mode run, consists of multiple inputs (up to 10), each with the same or different activity. The simulation stops after a number of iterations specified by the user. As output, the program provides a set of continues values for proportion of molecules in active form, and the weight of edges in the network according to the NSI formula. For some nodes, two kinds of proportion of activity including short (X_s) and long half-lives(X_L), are defined. In addition, the in-degree and out-degree of each node is specified and can be used as an indicator to select the proper half-lives for molecules.

Validating the model

Experiments for model validation can be divided into three major types. The first type involves stimulation of system with a step change in the input and measuring the changes in activities of various downstream species as a function of time [9]. Second type of experiment for model validation involves changes in the input by adding a high concentration of ligand and then measuring activity changes of various downstream species in several intervals [51]. Third type is used to provide insight of molecular activity trends in response to the external stimulus and perturbation [10]. We validated our model on four signaling networks including the EGFR/IGF-1R/IR (type I experiment for model validation), inflammatory and growth signaling networks in human hepatocyte cells (type II), and MAPK1,2 and AKT signaling network downstream from EGFR in MDA231 breast cancer cell line (type
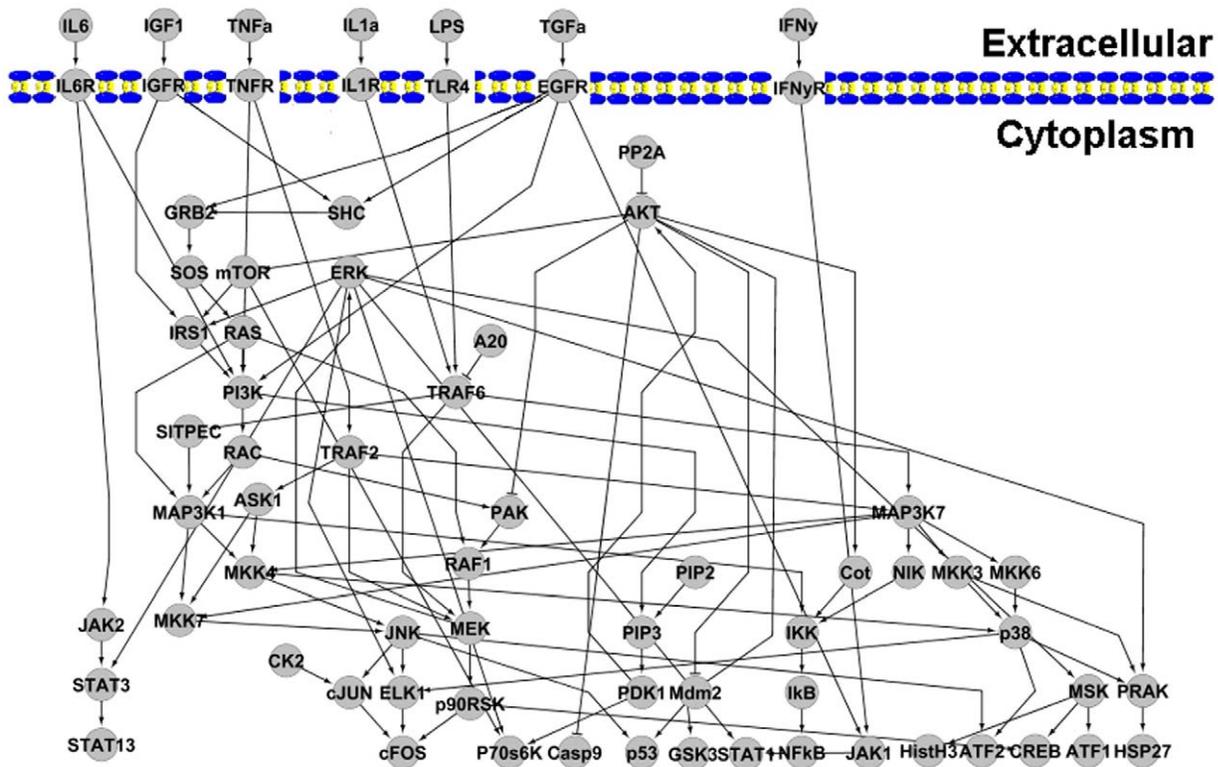
Figure 2. Inflammatory and growth signaling network in human hepatocyte cells. Activator (→) and inhibitor (⊣) reactions are indicated as edges. Nodes represented in the extracellular space are ligands.
doi:10.1371/journal.pone.0039643.g002

III). We also used human signaling network (containing 1634 nodes and 5089 edges), and mouse hippocampal CA1 neuron network (containing 545 nodes and 1259 edges) for visualization of organization in signal transduction networks weighted by similarity measure [31,52].

Signaling network of EGFR/IGF-1R/IR

Zielinski *et al.* (2009) assembled canonical pathways, downstream the three major receptors, including insulin receptor (IR), IGF-1R, and EGFR in SKOV3 cell line [9]. The network has 82 nodes and 128 edges, of which EGF, IGF, and insulin can be considered as the input nodes (Table S1).

Signaling networks of human hepatocyte cells

We obtained the connectivity for two signaling networks of human hepatocyte cells from published literatures (Table S2 and S3) [53,54]. Network of inflammatory and growth signaling is depicted in Figure 2.

MAPK1,2 and AKT signaling network downstream from EGFR in breast cancer cell line

We obtained the connectivity MAPK1,2 and AKT signaling network downstream from EGFR in MDA231 breast cancer cell line from a published literature [10]. There is a gain of function mutation in Ras for the MDA231 cell line [10].
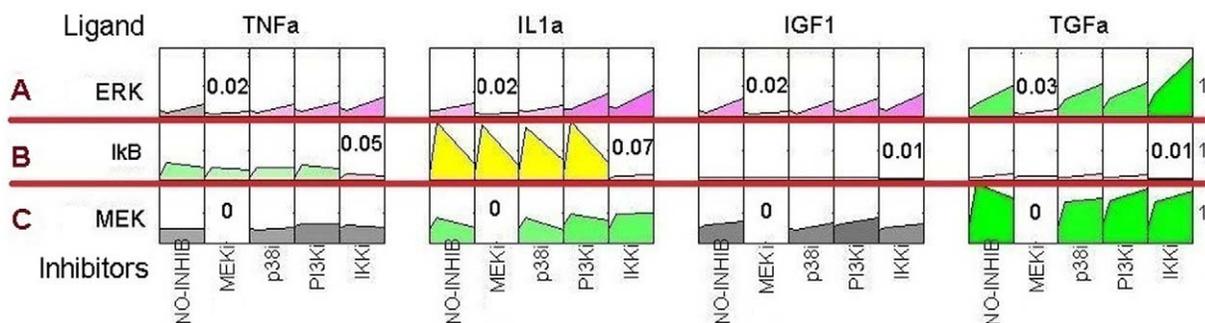
Figure 3. Signaling dataset of three proteins from human hepatocyte cells. Rows represent the measures of three assayed intracellular molecules, and columns represent 4 different ligands. For each combination of ligand and target molecule, one of 4 kinase inhibitors was applied, as indicated in the schematic diagram below the data. The numbers on the right of the figure indicate the maximum values of the corresponding row. The Figure was created using Data-Rail program [4,48].
doi:10.1371/journal.pone.0039643.g003

Results

In this study, we developed a computational framework to model the concentration of the active form of given proteins and obtained the quantities of proteins involved in cell-specific signaling networks. In the proposed modeling method the weight of a node was the product of its activity at time t-1, the amounts of signals received by upstream nodes and the weight of activating or blocking edges. We calculated the edge weights through the NSI formula (see Methods), as the efficiency of transmission signal between the nodes. As the signal transduction network was given appropriate weights using the similarity measure, cluster patterns became visible (Figure 4A and 4B). The cluster patterns corresponded to different parts of eukaryotic cells including cell membrane, cytoplasm and nucleus. These profiles suggest that inside the clusters, the efficiency of signal propagation along the edges is more than their outside that is primarily due to the facility posed by high NSI values.

Case studies for model validation

We chose four signaling networks as the EGFR/IGF-1R/IR, inflammatory and growth signals in human hepatocyte cells, and MAPK1,2 and AKT signaling network downstream from EGFR as the benchmarks for validating our method [9,10,53].

Signaling network of EGFR/IGF-1R/IR

The ability of the model to predict the proportion of active molecules in response to different levels of stimulation was the first question that we considered. Since the output of the model is continuous, it is possible to evaluate the activation of target molecules in response to different concentrations of ligands.

We used the network constructed by combining EGF, IGF and insulin signaling pathways and related experimental results including the responses of selected molecules to the specific stimuli [9]. Four proteins with network crosstalk, ERK, AKT, p70s6K, and JNK, were selected. In order to compare the experimental results with those obtained through computations, four sets of simulations were performed by 25% step-wise increase in the receptor activation [9]. This set of simulations made possible to have translation of the molecular response into activation level. The experimental and simulation results are reported as proportion of each examined molecule in active form. According to the results, our model was in line with experimentally observed values with a Pearson correlation of 0.742 ($P<10^{-15}$) (Figure 5) (Table S1).

We observed that computer simulation was able to recapitulate all the trends observed in the experimental studies. As expected, co-stimulation of the signaling with insulin and EGF led to the activation of JNK and MAPK1,2. Also, the activation of insulin and IGF-1 was translated to the activation of p70s6K during stimulation (Table S1) [9].

As mentioned before, our simulation method does not need experimental data such as the reaction rates or association constants. The performance of our method depends on how the underlying network is completed. For example, the discrepancy between the predictions of proportion of AKT molecules in active form with the experimental results, in almost all cases, can reveal that our information about the regulation of AKT in SKOV3 cell line was incomplete (Figure 5) (Table S1).

Signaling networks of inflammatory and growth signals in hepatocyte cells

We further evaluated the ability of the model in predicting the outcome of experiments by using two datasets consisting of the activity levels of signaling proteins of primary and transformed hepatocytes (Huh7) under various perturbations [4,7].

Inflammatory and growth signaling network in human hepatocyte cells included the pathways downstream the seven receptors, including IL6R, IGFR, TNFR, IL1R, TLR4, EGFR, and IFNγR (Figure 2). We simulated the activity levels of 14 signaling proteins (AKT, CREB, ERK, GSK3, HistH3, HSP27, IκB, IRS1s, JNK, MEK, p38, p70s6K, p90 & STAT3) under activation of IL1α, TGFα, IL6, TNFα and perturbations by two inhibitors (MEKi & PI3Ki) for all 168 possible pair-wise combinations at 100 iterations. Our tool showed that, predictions of activity levels of phosphoproteins under various perturbations agreed with those obtained experimentally with a Pearson correlation of 0.831 ($P<10^{-44}$) (Table S2). We simulated the activation state of CREB, ERK, HistH3, HSP27, p90, IκB, JNK, MEK, p38, and p70s6 proteins in human hepatocyte cells under activation of IL1b, TGFa, HER, INS, TNFα and perturbations by cMETi, MEKi and PI3Ki for all 200 possible pair-wise combinations at 100 iterations. Simulation of activity levels of 10 signaling proteins under various perturbations agreed with experimental data (Pearson correlation of 0.817; $P<10^{-46}$) (Table S3).

MAPK1,2 and AKT signaling network downstream from EGFR

We tested the accuracy and performance of our method to simulate the effects of targeted manipulation in MAPK1,2 and AKT signaling network in MDA231 breast cancer cell line [10]. In the MDA231 cell line, there is a gain of function mutation in Ras. This is modeled using fixed activity assignments on Ras during the simulation [10].

We compared the simulation and experimental results through computing the proportion of active form of mTOR, GSK3β, p70s6K, AKT, and MAPK1,2 molecules under activation of EGF and inhibition of TSC2 [10]. Simulation of normal and perturbed signal transduction was performed in 100 iterations. The Wilcoxon test was used to study changes in the signal propagation between the simulation results before and after perturbation. The results produced through the simulation agreed with those obtained through the experiments (Table 1). We did not expect that the TSC2 perturbation can have a significant effect on the activity of GSK3β and AKT, which was what the statistical test indicated (Table 1) [10]. The mTOR, p70S6K, and MAPK1,2 showed a significant response to the perturbation, and the changes in mean activity were beyond the significance level of 0.01.

Predictions made by the model

Activation of a T cell by exposure to specific agonist may lead to cytotoxic attacks on target cells, cytokine production or cell proliferation [55]. The process of T cell activation can be separated into a hierarchy of thresholds. In general, a hierarchy of thresholds is observed for T cell responses with the relative threshold order: $S_{cytotoxicity} << S_{cytokine\ production} < S_{cell\ proliferation}$ (where S stands for potency of stimulus) [56]. These thresholds depend on the stimulus conditions with the most significant changes occurring in the presence of co-stimulation of receptors [57,58]. The final level of this signaling process is the integration of signals to regulation of the gene transcription [59]. This process has been studied for its possible role in the diseases such as of autoimmune disorders, atopic dermatitis and fibrotic diseases [60,61]. In order to have an example of how our modeling approach might be used for novel predictions, we used threshold concept in T cell activation problem.

We considered activation model motivated by the dependence of cell responses on the potency of the signal and simulated the
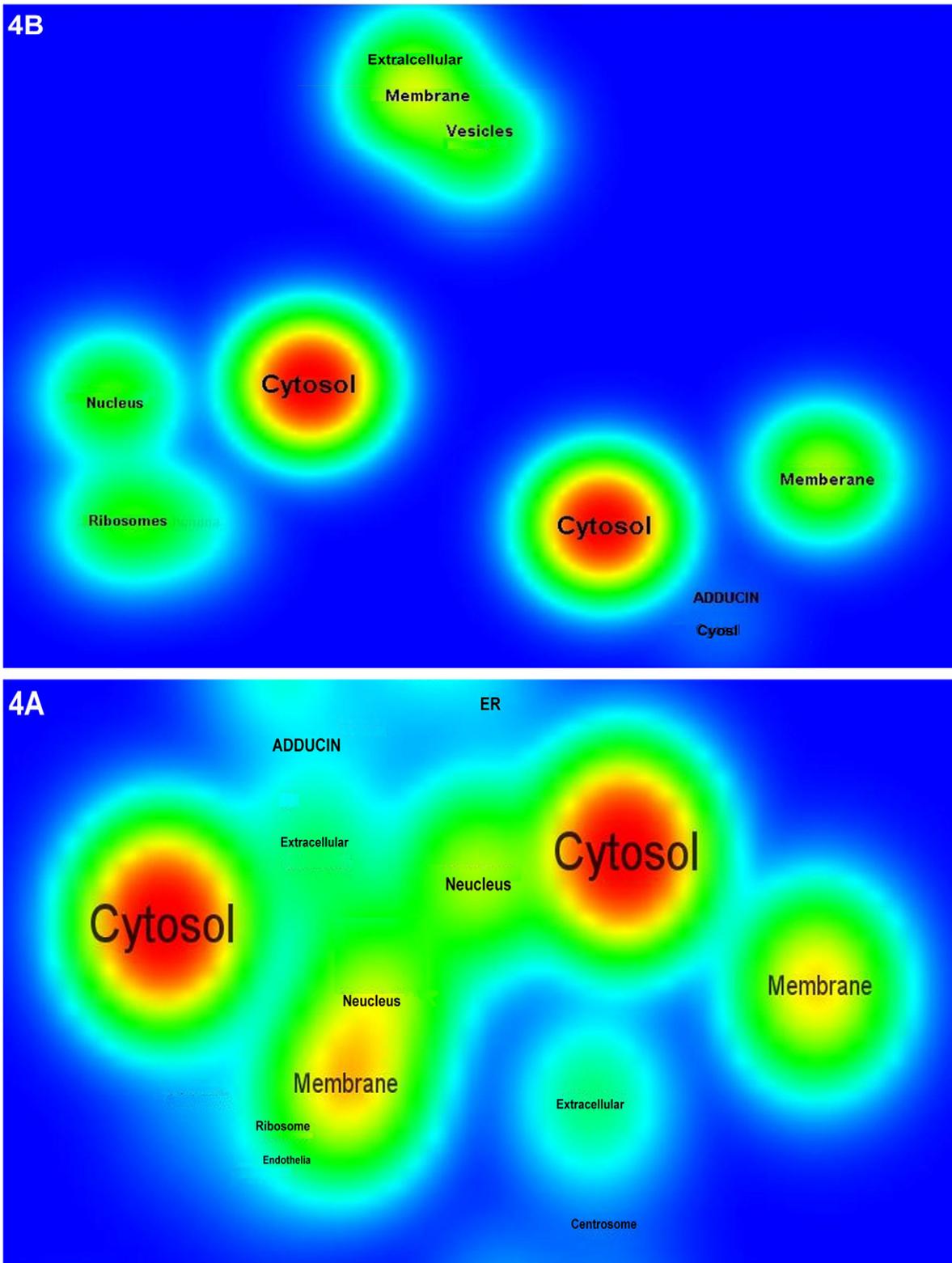
Figure 4. Cluster density view of the weighted cellular signaling networks by applying similarity measure. (A) human cancer cell signaling, and (B) mouse hippocampus CA1 neural networks were weighted by similarity measure. Nodes were labeled according to their position in the cell, including cell membrane, adducin, cell adhesion, centrosome, cytoskeleton, endothelial, endoplasmic reticulum, cytosolic, extracellular space, golgi apparatus, lysosome, mitochondria, nucleus, ribosome and vesicles. VOSveiwer program was used for visualizing connectivity-based clustering patterns [35]. This tool provides visualization of similarities, where objects with high similarity are located close to each other and those with low similarity are located far from each other.
doi:10.1371/journal.pone.0039643.g004
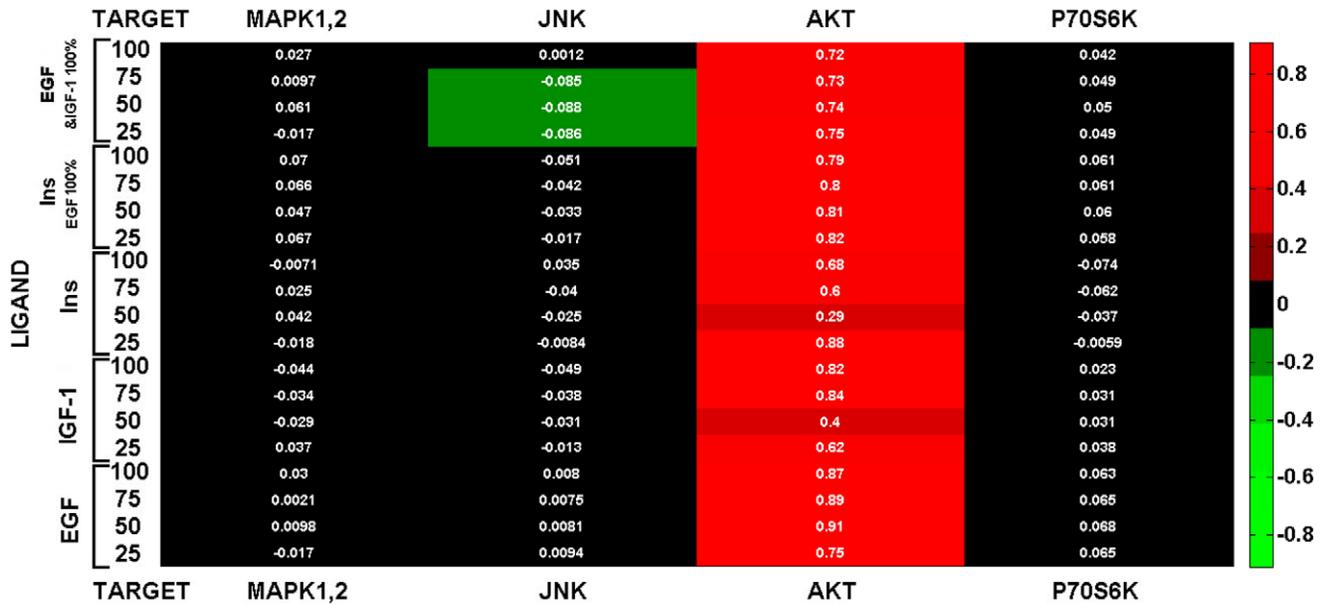
Figure 5. Comparison between the experimental results with those obtained through simulation. The activity level of four signaling molecules (MAPK1,2, JNK, AKT, and p70s6K) proteins in EGF/IGF/Insulin cell signaling network were simulated under activation of three ligand (EGF, insulin, and IGF-1) at 100 iterations. If the value of difference between the experimental and simulation data is bigger than zero, the corresponding box is colored in red; if best agreement, the box is black; but if the value of the difference between simulation and experimental data be smaller than zero, the box is green.
doi:10.1371/journal.pone.0039643.g005

effects of different treatments on NF-κB activation (Figure 2) [62]. Table 2 shows the results of simulating proportion of NF-κB molecules in active form were analyzed under combined and only treatment effects of five ligands including: IL1α, IL6, TNFα, IGF1, and TGFα. For each treatment condition, we simulated 100 time-steps. With respect to simulation results, NF-κB molecules appear to respond stronger under IL1α & IL6, IL1α & TNFα, IL1α & IGF1, and IL1α & TGFα treatments (Table 2). Hence, co-stimulation with IL1α may sufficiently stimulate the T cell to reach the proliferation threshold in response to treatment, as well as suggest effective target to perturb the progression of T cells to proliferation phase.

Table 1. The comparison between the simulation and experimental results of five signaling molecules under activation of EGF and inhibition of TSC2.

Molecule	Simulation Results		Change in Activity After TSC2 Inhibition		Wilcoxon-test	
	Normal	TSC2 inhibited	Experiment	Simulation	Z	P-value
mTOR	0.0038	0.022	↑	↑	−8.68	$<10^{-18}$
MAPK1,2	0.017	0.014	↓ or -	↓	−4.21	$<10^{-5}$
p70s6K	0.029	0.035	↑	↑	−8.68	$<10^{-18}$
AKT	0.072	0.071	↓ or -	-	−1.26	0.206
GSK3β	0.025	0.026	-	-	−3.14	0.016

The P-values were calculated by Wilcoxon test and used to test the changes in simulation results after perturbation. The upward arrow (↑) indicates that the perturbation caused a rise in the proportion of molecules in active form; the straight line (-) indicates no change, and the downward arrow (↓) indicates decrease in the proportion of molecules in active form.
doi:10.1371/journal.pone.0039643.t001

Table 2. Proportion of NF-κB molecules in active form under combined and only treatment effects of five ligands.

Signal		IL1α	IL6	TNFα	IGF1	TGFα
IL1α	NF-kβ	0.17				
IL6		0.26	0.10			
TNFα		0.25	0.18	0.18		
IGF1		0.26	0.22	0.24	0.20	
TGFα		0.25	0.20	0.24	0.23	0.18

The rows and columns of the table correspond to the treatment groups, and cells to proportion of NF-κB molecules in active form.
doi:10.1371/journal.pone.0039643.t002

Discussion

Modeling function of cellular networks in a dynamic fashion provides an optimal basis for elaborate study of cellular signal transduction [63]. However, lack of detailed concentration and kinetic data may make it difficult to use some modeling methods [64]. Thus, nonparametric modeling provides an alternative modeling approach to test hypothetical signaling networks [20,65]. In this work we developed a non-parametric computational framework to describe the proportion of active form and the trend of activation of given molecules in signaling network. In the previous nonparametric methods, simulation has been performed in such a way that all the components changed their states simultaneously in a unit of time based on the assumption that every reaction in the network takes exactly single unit of time in the signaling process [9,10]. Derek *et al.* (2008) presented a non-parametric and a Petri net-based model of cellular signaling networks [10]. This method provides insights into the trends of molecular activity-levels in response to an external stimulus, based

on the network's connectivity. Despite this success, the model could not predict the proportion of molecules in active form [10]. Zielinski *et al.* (2009) developed a network-specific model for dynamic simulation of signal transduction, and tested it on a network constructed by combining EGF, IGF and insulin signaling pathways [9]. The model agreed with many of the experimentally observed trends; however, it was notable to recapitulate the proportion of molecules in active form with significant correlation to the observed values in the experimental study [9]. In our proposed method, the defining dynamics occurred at the level of interactions among molecules, and coarse properties were computed by aggregating local quantities. This approach provides access to the microscopic dynamics which is hindered by the complexity of the system. The result was a fast method, which can provide insights into the proportion of molecules in active form and trends of molecular activity level in response to external stimuli. In our simulation method, node values were produced by combining two processes including edge weighting and simulation of signal flow from the initial node(s). During simulation, we used three updating strategies to reduce the artifacts due to the assumption of uniformity in reactions arising from synchronous updating methods. First, we employed a semi-synchronous updating scheme [40]. In our method, some of the incoming messages influence the updating state immediately, while others took longer to affect. Second, the level of activity for each node was transferring to next time-step with efficiency less than 1. This allowed us to take into account the relative stability of proteins in active form. Third, we used a specific function to weight the network edges and calibrated their efficiency for signal transition in the network. The outcome of the procedure used for edge weighting was a fractioned network to several clusters. These clusters were grouping molecules with the same place in the cell (*e.g.* nucleus). By coupling edge weighting with the statistical methodology of node weighting, we obtained a method capable of characterizing dynamic properties of signaling networks while using only network's connectivity information. Validation on several signaling networks showed that our method could effectively simulate both inhibited and constitutive activation of signal transduction components. Our simulation results were in strong agreements with the experimental results. Therefore, the present simulation method not only reproduces experimental data but also can predict non-intuitive and previously unknown responses. Also, the simulation results are capable of linking signal transduction to any type of quantifiable cellular responses such as cell growth, survival, apoptosis, necrosis, cytokine secretion, or transcriptional activity.

Our principal conclusion is that the dynamical phenotype possesses the ability of design according to the network topology. This finding corresponds conceptually to proteins where a two- and three-dimensional structure usually possesses design abilities according to the sequence of amino acids [66]. Our investigation showed that the dynamical phenotypes arise via the presence of conserved network links and could reflect wide variations in the level of activity at different positions. In summary, with the networks involved in case studies, our approach has proved itself as a promising tool to analyze signal transduction, effect of drugs and network modifications.

Supporting Information

Table S1 Simulation results of activity levels of four signaling proteins in signaling network of EGFR/IGF-1R/IR under various perturbations with comparison with experimental results. (A) The proportion of MAPK1,2, JNK, AKT, and p70s6K molecules in active forms in response to the activation of EGFR, IGF-1R, and IR. (B) Activation of MAPK1,2, JNK, AKT, and p70s6K in response to co-stimulation of EGFR/IR and IGF-1R/EGFR/IR.
(XLSX)

Table S2 The simulation results of activity levels of 14 signaling proteins in human hepatocyte cells under various perturbations, in comparison with experimental results.
(XLSX)

Table S3 The simulation results of activity levels of 10 signaling proteins in the Huh7 cell line under various perturbations, in comparison with experimental results.
(XLSX)

Author Contributions

Conceived and designed the experiments: IN AMN. Performed the experiments: IN. Analyzed the data: IN AMN MJ AM. Contributed reagents/materials/analysis tools: IN. Wrote the paper: IN AMN.

References

1. Olsen J, Blagoev B, Macek B, Mann M (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell 127: 635–648.
2. Tyers M, Mann M (2003) From genomics to proteomics. Nature 422: 193–197.
3. Gaudet S, Janes KA, Albeck JG, Pace EA, Lauffenburger DA, et al. (2005) A compendium of signals and responses triggered by prodeath and prosurvival cytokines. Mol Cell Proteomics 4: 1569–1590.
4. Alexopoulos L, Saez-Rodriguez J, Cosgrove B, Lauffenburger D, Sorger P (2010) Networks Inferred from Biochemical Data Reveal Profound Differences in Toll-like Receptor and Inflammatory Signaling between Normal and Transformed Hepatocytes. Molecular & Cellular Proteomics 15 840–847.
5. Pritchard JR, Cosgrove BD, Hemann MT, Griffith LG, Wands JR, et al. (2009) Three-kinase inhibitor combination recreates multipathway effects of a geldanamycin analogue on hepatocellular carcinoma cell death. Mol Cancer Ther 8: 2183–2192.
6. Cosgrove BD, King BM, Hasan MA, Alexopoulos LG, Farazi PA, et al. (2009) Synergistic drug-cytokine induction of hepatocellular death as an in vitro approach for the study of inflammation-associated idiosyncratic drug hepatotoxicity. Toxicol Appl Pharmacol 237: 317–330.
7. Melas IN, Mitsos A, Messinis DE, Weiss TS, Alexopoulos LG (2011) Combined logical and data-driven models for linking signalling pathways to cellular response. Bmc Systems Biology 5: 107.
8. White FM (2008) Quantitative phosphoproteomic analysis of signaling network dynamics. Curr Opin Biotechnol 19: 404–409.
9. Zielinski R, Przytycki PF, Zheng J, Zhang D, Przytycka TM, et al. (2009) The crosstalk between EGF, IGF, and Insulin cell signaling pathways--computational and experimental analysis. BMC Syst Biol 3: 88.
10. Ruths D, Muller M, Tseng JT, Nakhleh L, Ram PT (2008) The signaling petri net-based simulator: a non-parametric strategy for characterizing the dynamics of cell-specific signaling networks. PLoS Comput Biol 4: e1000005.
11. Riel V (2006) Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. Brief Bioinform 7: 364–374.
12. Chen L, Aihara K (2002) Stability of genetic regulatory networks with time delay. IEEE Trans Circuits Syst I: 602–608.
13. Klann M, Lapin A, Reuss M (2009) Stochastic simulation of signal transduction: impact of the cellular architecture on diffusion. Biophys J 96: 5122–5129.

14. vSun J, Weinstein H (2007) Toward realistic modeling of dynamic processes in cell signaling: Quantification of macromolecular crowding effects. J Chem Phys 127: 155105–155110.

15. Schnapp B (2003) Trafficking of signaling modules by kinesin motors. Cell Sci 116: 2125–2135.

16. Jenkins DJ, Stekel DJ (2010) Stochasticity versus determinism: consequences for realistic gene regulatory network modelling and evolution. J Mol Evol 70: 215–231.

17. Abhishek Garg KM, Alessandro Di Cara, Giovanni De Micheli, Ioannis Xenarios (2008) Modeling stochasticity and robustness in gene regulatory networks. Bioinformatics 25 101–109.

18. Garg A, Di Cara A, Xenarios I, Mendoz L, De Micheli G (2008) Synchronous versus asynchronous modeling of gene regulatory networks. Bioinformatics 24: 1917–1925.

19. Riel NA (2006) Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. Brief Bioinform 7: 364–374.

20. Li S, Assmann SM, Albert R (2006) Predicting essential components of signal transduction networks: a dynamic model of guard cell abscisic acid signaling. PLoS Biol 4: e312.

21. Jack J, Wambaugh JF, Shah I (2011) Simulating quantitative cellular responses using asynchronous threshold boolean network ensembles. BMC Syst Biol 5: 109.

22. Aldridge BB, Saez-Rodriguez J, Muhlich JL, Sorger PK, Lauffenburger DA (2009) Fuzzy logic analysis of kinase pathway crosstalk in TNF/EGF/insulin-induced signaling. PLoS Comput Biol 5: e1000340.

23. Zhang R, Shah M, Yang J, Nyland S, Liu X, et al. (2008) Network model of survival signaling in large granular lymphocyte leukemia. Proc Natl Acad Sci USA 105: 16308–16313.

24. Albert R, Wang RS (2009) Discrete dynamic modeling of cellular signaling networks. Methods Enzymol 467: 281–306.

25. Janes K, Albeck J, Peng L, Sorger P, Lauffenburger D, et al. (2003) A high-throughput quantitative multiplex kinase assay for monitoring information flow in signaling networks: application to sepsis-apoptosis. Mol Cell Proteomics 2: 463–473.

26. Brun C, Chevenet F, Martin D, Wojcik J, Guenoche A, et al. (2003) Functional classification of proteins for the prediction of cellular function from a protein-protein interaction network. Genome Biol 5: R6.

27. Samanta MP, Liang S (2003) Predicting protein functions from redundancies in large-scale protein interaction networks. Proc Natl Acad Sci U S A 100: 12579–12583.

28. Kaiser M, Hilgetag C (2004) Edge vulnerability in neural and metabolic networks. Biological Cybernetics 90: 311–317.

29. Leydesdorff L, Vaughan L (2006) Co-occurrence matrices and their applications in information science: Extending ACA to the Web environment. Journal of the American Society for Information Science and Technology 57: 1616–1628.

30. Pavlopoulos G, Secrier M, Moschopoulos CN, Soldatos TG, Kossida S et al (2011) Using graph theory to analyze biological networks. BioData Mining 28: 10.

31. Ma'ayan A, Jenkins S, Neves S (2005) Formation of regulatory patterns during signal propagation in a mammalian cellular network. Science 309: 1078–1083.

32. Costa Lda F, Kaiser M, Hilgetag CC (2007) Predicting the connectivity of primate cortical networks from topological and spatial node properties. BMC Syst Biol 1: 16.

33. Nisbach F (2007) Developmental time windows for spatial growth generate multiple-cluster small-world networks. European Physics Journal 58: 185–191.

34. Li A, Horvath S (2007) Network neighborhood analysis with the multi-node topological overlap measure. Bioinformatics 23: 222–231.

35. Nees J, Waltman L (2009) Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 5: 14–17.

36. Muller M, Obeyesekere M, Mills GB, Ram PT (2008) Network topology determines dynamics of the mammalian MAPK1,2 signaling network: bifan motif regulation of C-Raf and B-Raf isoforms by FGFR and MC1R. Faseb Journal 22: 1393–1403.

37. Yen HC, Xu Q, Chou DM, Zhao Z, Elledge SJ (2008) Global protein stability profiling in mammalian cells. Science 322: 918–923.

38. Kampen V (1992) Stochastic processes in physics and chemistry. Elsevier Science Publishing Company, North-Holland: 450p.

39. Albert R, Othmer HG (2003) The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in Drosophila melanogaster. J Theor Biol 223: 1–18.

40. Assmann SM, Albert R (2009) Discrete dynamic modeling with asynchronous update, or how to model complex systems in the absence of quantitative information. Methods Mol Biol 553: 207–225.

41. Chen L, Wang R, Zhang X (2009) Biomolecular networks: methods and applications in systems biology. John Wiley & Sons: 421p.

42. Lemmon MA (2009) Ligand-induced ErbB receptor dimerization. Exp Cell Res 315: 638–648.

43. Krall JA, Beyer EM, MacBeath G (2011) High- and low-affinity epidermal growth factor receptor-ligand interactions activate distinct signaling pathways. Plos One 6: e15945.

44. Fu C, Li J, Wang E (2009) Signaling network analysis of ubiquitin-mediated proteins suggests correlations between the 26S proteasome and tumor progression. Mol Biosyst 5: 1809–1816.

45. Belle A, Tanay A, Bitincka L, Shamir R, O'Shea EK (2006) Quantification of protein half-lives in the budding yeast proteome. Proc Natl Acad Sci U S A 103: 13004–13009.

46. Chaouiya C (2007) Petri net modelling of biological networks. Brief Bioinform 8: 210–219.

47. Klamt S, Saez-Rodriguez J, Gilles ED (2007) Structural and functional analysis of cellular networks with CellNetAnalyzer. BMC Syst Biol 1: 2.

48. Saez-Rodriguez J, Goldsipe A, Muhlich J, Alexopoulos L, Millard B, et al. (2008) Flexible informatics for linking experimental data to mathematical models via Datarail. Bioinformatics 15: 840–847.

49. Wang RS, Albert R (2011) Elementary signaling modes predict the essentiality of signal transduction network components. BMC Syst Biol 5: 44.

50. Saadatpour A, Wang RS, Liao A, Liu X, Loughran TP, et al. (2011) Dynamical and structural analysis of a T cell survival network identifies novel candidate therapeutic targets for large granular lymphocyte leukemia. Plos Computational Biology 7: e1002267.

51. Apgar JF, Toettcher JE, Endy D, White FM, Tidor B (2008) Stimulus design for model selection and validation in cell signaling. PLoS Comput Biol 4: e30.

52. Cui Q, Ma Y, Jaramillo M, Bari H, Awan A, et al. (2007) A map of human cancer signaling. Mol Syst Biol 3: 152.

53. Saez-Rodriguez J, Alexopoulos l, Morris M, Lauffenburger D, Sorger P (2011) Hepatocytes Using Discrete Logical Models Comparing Signaling Networks between Normal and Transformed. Cancer Res 71: 5400-5425.

54. Saez-Rodriguez J, Alexopoulos LG, Epperlein J, Samaga R, Lauffenburger DA, et al. (2009) Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction. Mol Syst Biol 5: 331.

55. Doherty DG, Norris S, Madrigal-Estebas L, McEntee G, Traynor O, et al. (1999) The human liver contains multiple populations of NK cells, T cells, and CD3+CD56+ natural T cells with distinct cytotoxic activities and Th1, Th2, and Th0 cytokine secretion patterns. J Immunol 163: 2314–2321.

56. Valitutti S, Lanzavecchia A (1997) Serial triggering of TCRs: a basis for the sensitivity and specificity of antigen recognition. Immunol Today 18: 299–304.

57. Iezzi G, Karjalainen K, Lanzavecchia A (1998) The duration of antigenic stimulation determines the fate of naive and effector T cells. Immunity 8: 89–95.

58. Wedagedara J (2011) Some large deviation estimates for an Erlang queue of phosphorylated T-cell receptors. Journal of National Science Foundation of Sri Lanka 39: 03–12.

59. Van Den Berg HA, Rand DA, Burroughs NJ (2001) A reliable and safe T cell repertoire based on low-affinity T cell receptors. Journal of Theoretical Biology 209: 465–486.

60. Cope AP (2004) Altered signalling thresholds in T lymphocytes cause autoimmune arthritis. Arthritis Res Ther 6: 112–116.

61. Cevikbas F, Steinhoff M (2012) IL-33: A Novel Danger Signal System in Atopic Dermatitis. Journal of Investigative Dermatology 132: 1326–1329.

62. Jost PJ, Ruland J (2007) Aberrant NF-kappaB signaling in lymphoma: mechanisms, consequences, and therapeutic implications. Blood 109: 2700–2707.

63. Morris M, Saez-Rodriguez J, Sorger P, Lauffenburger D (2010) Logic-based models for the analysis of cell signaling networks. Biochem Biophys Res Commun 49 3216–3224.

64. Breitling R, Hoeller D (2005) Current challenges in quantitative modeling of epidermal growth factor signaling. FEBS Lett 579: 6289–6294.

65. Klamt S, Saez-Rodriguez J, Lindquist JA, Simeoni L, Gilles ED (2006) A methodology for the structural and functional analysis of signaling and regulatory networks. BMC Bioinformatics 7: 56.

66. Wallner B, Elofsson A (2005) All are not equal: a benchmark of different homology modeling programs. Protein Sci 14: 1315–1327.