Computational Modeling of Molecular Pathways Regulating Cell Survival and Death by the *SigFlux* Algorithm

A. Abdi
Department of Electrical Engineering and Computer Science
New Jersey Institute of Technology
Aa675@njit.edu

Computational modeling has been increasingly used to model cell signaling networks in recent years. In particular, logical modeling is becoming a well-recognized necessary and valuable tool to understand the dynamical features of complex biological networks and has been proved to be particularly fruitful to model and analyze the regulatory and signaling networks [1-3].

Molecular networks that regulate cell fate, more specifically, cell death or cell survival, have been extensively studied by molecular biologists. Cell death is studied under two major categories of apoptotic or non-apoptotic cell death. Apoptosis often referred to as the programmed cell death, is a cell suicide mechanism. During apoptosis, a cell turns on its regulatory mechanisms to play a coordinated program to disrupt its cellular components and pack them into specialized vesicles that can be easily removed from the environment. In non-apoptotic cell death, referred to as necrosis, intracellular components are released in the surrounding tissues that evoke some inflammatory response that gathers several elements of the immune system in the tissue and cause severe injury. Interestingly, there are significant overlaps between the molecular pathways that regulate these three important cell fates, i.e., apoptosis, non-apoptotic cell death (NonACD) or cell survival.

The study by Calzone et al (2010) is one of the publicly available studies that has integrated the three major cellular pathways that lead to apoptosis, NonACD or survival. This study used a discrete modeling formalism to present a mathematical model of cell fate decision making that shows the interplays between pro-survival, apoptotic and non-apoptotic pathways in response to death receptor-mediated signals [4]. The present study uses a *SigFlux* algorithm based on the model proposed by Calzone et al., and develops a software to rank the significance of each one of the molecules in the same integrated network studied by Calzone et al. The main aim is to determine which molecules are the most important ones in causing the three cell fates: survival, apoptosis or NonACD.

To achieve the main aim of this study, the *SigFlux* algorithm was used to rank the significance of every node in the network [1]. The biological cell network was represented by a NetworkX graph in Python. To give some brief background, NetworkX is a “Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks”. It represents the cell structure superbly and has the functionality to delimit certain paths as being inhibitor or activator with edge weighting. However, for the *SigFlux* algorithm, path type does not matter. We show that this software makes some biologically relevant predictions and can correctly identify molecules that are known to be critical for apoptosis versus NonACD.

The results obtained are not only biologically relevant to the role of each molecule for the regulation of a given outcome, but also, they can be explained to some degree by their connectivity to other molecules in the network. While this study is a small-scale testing of this algorithm, it can potentially be used for the analysis of larger molecular networks. It also has some predictive value, as it may identify some previously unknown molecules in a larger network that can control the function of an output. Such previously unknown molecules would be more interesting since the predictive power of this algorithm can be tested by designing appropriate molecular biology experiments for identification of novel regulators of different cellular functions.

REFERENCES

[1] Liu W, Li D, Zhang J, Zhu Y, He F. SigFlux: A novel network feature to evaluate the importance of proteins in signal transduction networks. BMC Bioinformatics. 7:515, 2006.

[2] Albert R., Thakar J. Boolean modeling: a logic-based dynamic approach for understanding signaling and regulatory networks and for making useful predictions. Wiley Interdisc. Rev. Syst. Biol. Med. 6, pp. 353–369, 2014.

[3] Abou-Jaoudé W, Traynard P, Monteiro PT, et al. Logical Modeling and Dynamical Analysis of Cellular Networks. Frontiers in Genetics. 2016.

[4] Calzone L, Tournier L, Fourquet S, Thieffry D, Zhivotovsky B, Barillot E, et al. Mathematical Modelling of Cell-Fate Decision in Response to Death Receptor Engagement. PLoS Comput Biol 6(3): 2010.

Computational Modeling of Molecular Pathways Regulating Cell Survival and Death by the *SigFlux* Algorithm

A. Abdi
Department of Electrical Engineering and Computer Science
New Jersey Institute of Technology
Aa675@njit.edu

ABSTRACT

Development of new computational tools to better understand the complex behavior of molecular networks in a cell, has recently become one of the most attractive areas of research at the interface of computer science and molecular biology. In this study, the molecular pathways that regulate cell survival, apoptotic or non-apoptotic cell death is studied. More specifically, a *SigFlux* algorithm is used to develop a software to rank the significance of every molecule (node) for causing three different cell fates, i.e., cell survival, apoptosis (programmed cell death) and non-apoptotic cell death (NonACD). The ranking obtained from applying this algorithm identified some molecules that are biologically relevant to the induction of cell survival, apoptosis or NonACD. For example, this algorithm found that Caspase-3, one of the most important molecules in the initiation and progression of apoptosis, has the highest rank for causing apoptosis but shows a lower rank in NonACD or cell survival. This finding is consistent with empirical findings obtained from extensive experimentation on a wide variety of cells.

BACKGROUND

Computational modeling has been increasingly used to model cell signaling networks in recent years. In particular, logical modeling is becoming a well-recognized necessary and valuable tool to understand the dynamical features of complex biological networks and has been proved to be particularly fruitful to model and analyze the regulatory and signaling networks (Albert and Thakar, 2014, Abou-Jaoudéet al., 2016).

Molecular networks that regulate cell fate, more specifically, cell death or cell survival, have been extensively studied by molecular biologists. Cell death is studied under two major categories of apoptotic or non-apoptotic cell death. Apoptosis often referred to as the programmed cell death, is a cell suicide mechanism. During apoptosis, a cell turns on its regulatory mechanisms to play a coordinated program to disrupt its cellular components and pack them into specialized vesicles that can be easily removed from the environment. In non-apoptotic cell death, referred to as necrosis, intracellular components are released in the surrounding tissues that evoke some inflammatory response that gathers several elements of the immune system in the tissue and cause severe injury. Interestingly, there are significant overlaps between the molecular pathways that regulate these three important cell fates, i.e., apoptosis, non-apoptotic cell death (NonACD) or cell survival.

The study by Calzone et al (2010) is one of the publicly available studies that has integrated the three major cellular pathways that lead to apoptosis, NonACD or survival. This study used a discrete modeling formalism to present a mathematical model of cell fate decision making that shows the interplays between pro-survival, apoptotic and non-apoptotic pathways in response to death receptor-mediated signals (Calzone et al., 2010). The present study uses a *SigFlux* algorithm based on the model proposed by Calzone et al., and develops a software to rank the significance of each one of the molecules in the same integrated network studied by Calzone et al. The main aim is to determine which molecules are the most important ones in causing the three cell fates: survival, apoptosis or NonACD. We show that this software makes some biologically relevant predictions and can correctly identify molecules that are known to be critical for apoptosis versus NonACD.

METHODS

To achieve the main aim of this study, the SigFlux algorithm was used to rank the significance of every node in the network (Liu et al., 2006). The biological cell network was represented by a NetworkX graph in Python. To give some brief background, NetworkX is a "Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks". NetworkX refers to networks as graphs as well, so the terms will be used interchangeably in this paper. It represents the cell structure superbly and has the functionality to delimit certain paths as being inhibitor or activator with edge weighting. However, for the SigFlux algorithm, path type does not matter.

The SigFlux algorithm is calculated as follows:

$$SigFlux = \frac{m_{pi} + m_{fi}}{\sum_{i=1}^{n}(m_{pi} + m_{fi})}$$

where m_{pi} denotes the number of signaling paths from input to output in which protein i is involved, m_{fi} deontes the number of feedback loops including protein i, and n is the number of all proteins in the network. The more paths protein i is involved in, the more important protein i is for the signaling network. The value of *SigFlux* varies between 0 and 1. The extreme value of zero occurs when protein i is not a member of any paths, and the extreme value of one is assigned to the most important protein in the network, i.e. the protein whose removal causes a disruption in the topological structure of the network (Liu et al., 2006).

The greater the *SigFlux* value of a certain node (molecule) in a network, the greater its significance is in the overall network. The values range from 0 to 1. However, *SigFlux* values cannot be compared from network to network. A hypothetical 0.3 *SigFlux* value could be significant in network A, yet be irrelevant in network B.

Input Cell Network: The network we will be calculating SigFlux values for is the death receptor network, also referred to as the death receptor pathways (Figure 1). Cells in different human tissues express different levels of the molecules in this network to regulate cell survival or cell death.

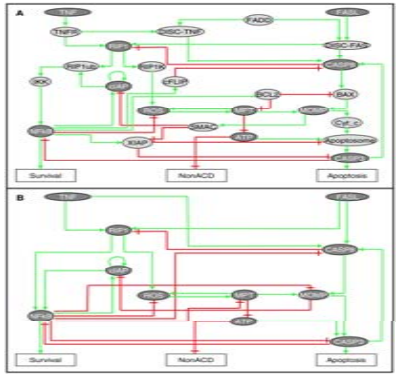
Calculating Signal Paths: As the algorithm for SigFlux states, to calculate the numerator and denominator the number of signals paths for the considered node and for every node must be calculated respectively. This paper defines a "signal path" as all the routes from every given input to every given output in the network that node i is a part of. A simple example, where node 1 is considered input and node 6 is the output, is presented in Figure 3, with blue arrows.

Calculating signal paths was implemented in Python using a function that utilized one of NetworkX's built in algorithms, called "all_simple_paths". This algorithm is a depth-first search that when given a graph, input, and output, and node, returns all the paths from input to output that contains the given node. Using a function which allows for the addition of multiple inputs and outputs, the number of signal paths for a molecule in a cell network can be calculated using it's graph representation in NetworkX and the all_simple_paths algorithm.

Calculating Feedback Loops: Similar to signal paths, calculating feedback loops is included in both the numerator and denominator of the SigFlux value for a node. In this paper, a feedback loop is defined as a as a path from node A back to node A; a simple cycle. The number of feedback loops that include node i is the second value that goes into the numerator of SigFlux, the mfi. An example of a feedback loop on a simple network, where node 1 is the input, and node 6 is the output is shown in Figure 3, by the red arrows.

However, unlike signal paths, calculating these values isn't as simple when it comes to implementation. There are two common ways of calculating. First, more difficult method is to begin a depth-first search starting from node i, the node you are calculating feedback loops for, and remembering which nodes you already saw and the path to them. If you happen to visit a node you already saw, then there is a cycle, and you can find it by concatenating paths. You take a simple count of the number of paths you create via concatenation, and then you have the number of feedback loops that node i is a part of. The second method is to use the built-in algorithm in NetworkX mentioned previously; all_simple_paths. Consider node i, the node we are calculating the number of feedback loops for. If we consider all the neighbors of node i to be input nodes, and node i itself to be an output node, then we can use the all_simple_paths algorithm with the aforementioned input and output to calculate feedback loops that include node i.

Calculations: The SigFlux parameter was calculated three times for every node in the death receptor network; one time for every cell fate. The inputs were constant across all three calculation runs. Molecules TNF and FASL were considered inputs against the outputs of Survival, NonACD, and Apoptosis, where the cell fates were each considered nodes for the purposes of calculating SigFlux. Connections between molecules in the death receptor were handled as NetworkX edges with all the same weight, meaning the only differences in every edge were its source and target.



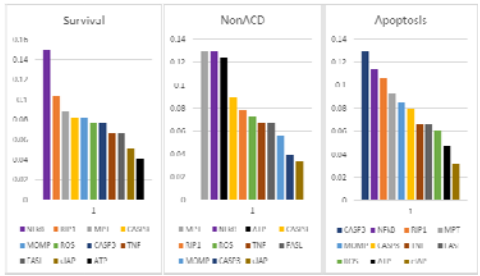
RESULTS

Tables 1-3 present the *SigFlux* for every molecule in the death receptor network, with survival, NonACD, and apoptosis as the outputs respectively. The *SigFlux* values of the molecules changed drastically depending on which cell fate was being considered the output, meaning that certain molecules are more significant to certain cell fates. Across all three tables, NFkB had a high *SigFlux* value, meaning the molecule is important in the entire network regardless of which cell fate is being orchestrated. Note: all *SigFlux* values were rounded to 4 digits to avoid redundancy. None of the values were the same prior to the rounding, and are ordered accordingly.

Survival	
NFkB	0.1495
RIP1	0.1031
MPT	0.0876
CASP8	0.0825
MOMP	0.0825
ROS	0.0773
CASP3	0.0773
TNF	0.067
FASL	0.067
cIAP	0.0515
ATP	0.0412

NonACD	
MPT	0.1292
NFkB	0.1292
ATP	0.1236
CASP8	0.0899
RIP1	0.0787
ROS	0.073
TNF	0.0674
FASL	0.0674
MOMP	0.0562
CASP3	0.0393
cIAP	0.0337

Apoptosis	
CASP3	0.1298
NFkB	0.1138
RIP1	0.1058
MPT	0.0926
MOMP	0.0847
CASP8	0.0794
TNF	0.0661
FASL	0.0661
ROS	0.0608
ATP	0.0476
cIAP	0.0317



DISCUSSION

The results were calculated on the graph model of the death receptor network reported by Calzone et al., after it was transcribed into a manipulatable data structure in Python. As briefly explained in Section II, Calzone et al. created two mathematical models of the death receptor network. The first, called the master model, contained every molecule in the network and every connection and pathway between said molecules. The second, named the reduced model, removes the redundant molecules and connections. This reduced model is what was used to calculate the SigFlux parameter for the remaining molecules, to achieve a better understanding of which molecules played a more significant role in certain cell fates.

A simple examination of the results can confirm that the SigFlux values directly correlate with the biological role of each molecule for a given outcome. For example, Caspase-3 is the most important molecule in apoptosis. In fact, apoptosis is often called Caspase-3 induced death. This molecule is showing the highest SigFlux value for apoptosis, despite the fact that it directly connects to NFkB, the most important molecule for the cell survival. Besides the biological relevance of the findings, the topology of the network, as discussed below, can also explain the results very well.

By looking and the reduced mathematical model of the death receptor pathways in Figure 2b, one can see that there is one molecule right above every cell fate, with multiple connections. By logic, one could guess that these molecules will be more important than others to their nearest output. However, the number of edges each node in the network has is also an indicator of their significance in the network. As illustrated previously in the SigFlux formula, the greater the number of signal paths a node is a part of, the greater its SigFlux value. More edges mean a greater likelihood for more paths. Glancing at Figure 2a, it is apparent that NFkB is the most significant molecule in the network. Visual depiction agrees; NFkB has 10 edges and acts as a "gateway node" for Survival. It is impossible for the cell survival fate to be reached without passing through NFkB, which is another reason for its high SigFlux value. RIP1 also has a high SigFlux of 0.1031, which is a result of its early role in connecting the input TNF to nodes further in the network (NFkB being one of them).

CONCLUSIONS

In this study, a SigFlux algorithm is used to rank the importance of every molecule (node) for causing three different cell fates, i.e., cell survival, apoptosis or non-apoptotic cell death. The results obtained are not only biologically relevant to the role of each molecule for the regulation of a given outcome, but also, they can be explained to some degree by their connectivity to other molecules in the network. While this study is a small-scale testing of this algorithm, it can potentially be used for the analysis of larger molecular networks. It also has some predictive value, as it may identify some previously unknown molecules in a larger network that can control the function of an output. Such previously unknown molecules would be more interesting since the predictive power of this algorithm can be tested by designing appropriate molecular biology experiments for identification of novel regulators of different cellular functions.

References

[1] Liu W, Li D, Zhang J, Zhu Y, He F. SigFlux: A novel network feature to evaluate the importance of proteins in signal transduction networks. BMC Bioinformatics. 2006;7:515. doi:10.1186/1471-2105-7-515.

[2] Albert R., Thakar J. (2014). Boolean modeling: a logic-based dynamic approach for understanding signaling and regulatory networks and for making useful predictions. Wiley Interdisc. Rev. Syst. Biol. Med. 6, 353–369. 10.1002/wsbm.1273 [PubMed] [Cross Ref]

[3] Abou-Jaoudé W, Traynard P, Monteiro PT, et al. Logical Modeling and Dynamical Analysis of Cellular Networks. Frontiers in Genetics. 2016;7:94. doi:10.3389/fgene.2016.00094.

[4] Calzone L, Tournier L, Fourquet S, Thieffry D, Zhivotovsky B, Barillot E, et al. (2010) Mathematical Modelling of Cell-Fate Decision in Response to Death Receptor Engagement. PLoS Comput Biol 6(3): e1000702.