

2016

# Systems Perturbation Analysis of a Large-Scale Signal Transduction Model Reveals Potentially Influential Candidates for Cancer Therapeutics

Bhanwar Lal Puniya

*University of Nebraska-Lincoln*, puniya.bl@gmail.com

Laura Allen

*University of Nebraska at Omaha*

Colleen Hochfelde

*Albert Einstein College of Medicine, New York*, hochfeld@mail.einstein.yu.edu

Mahbubul Majumder

*University of Nebraska at Omaha*, mmajumder@unomaha.edu

Tomáš Helikar

*University of Nebraska-Lincoln*, thelikar2@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/biochemfacpub>

 Part of the [Biochemistry Commons](#), [Biotechnology Commons](#), and the [Other Biochemistry, Biophysics, and Structural Biology Commons](#)

---

Puniya, Bhanwar Lal; Allen, Laura; Hochfelde, Colleen; Majumder, Mahbubul; and Helikar, Tomáš, "Systems Perturbation Analysis of a Large-Scale Signal Transduction Model Reveals Potentially Influential Candidates for Cancer Therapeutics" (2016). *Biochemistry -- Faculty Publications*. 163.

<https://digitalcommons.unl.edu/biochemfacpub/163>

This Article is brought to you for free and open access by the Biochemistry, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Biochemistry -- Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



# Systems Perturbation Analysis of a Large-Scale Signal Transduction Model Reveals Potentially Influential Candidates for Cancer Therapeutics

Bhanwar Lal Puniya<sup>1</sup>, Laura Allen<sup>2</sup>, Colleen Hochfelder<sup>3</sup>, Mahbulul Majumder<sup>2</sup> and Tomáš Helikar<sup>1\*</sup>

<sup>1</sup>Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, NE, USA, <sup>2</sup>Department of Mathematics, University of Nebraska at Omaha, Omaha, NE, USA, <sup>3</sup>Albert Einstein College of Medicine, New York, NY, USA

## OPEN ACCESS

### Edited by:

David A. Rosenblueth,  
Universidad Nacional Autónoma de  
México, México

### Reviewed by:

Reka Albert,  
Pennsylvania State University, USA  
Ovidiu Radulescu,  
Université de Montpellier 2, France

### \*Correspondence:

Tomáš Helikar  
thelikar2@unl.edu

### Specialty section:

This article was submitted to  
Bioinformatics and  
Computational Biology,  
a section of the journal  
Frontiers in Bioengineering and  
Biotechnology

**Received:** 01 November 2015

**Accepted:** 25 January 2016

**Published:** 11 February 2016

### Citation:

Puniya BL, Allen L, Hochfelder C,  
Majumder M and Helikar T (2016)  
Systems Perturbation Analysis of a  
Large-Scale Signal Transduction  
Model Reveals Potentially Influential  
Candidates for Cancer Therapeutics.  
*Front. Bioeng. Biotechnol.* 4:10.  
doi: 10.3389/fbioe.2016.00010

Dysregulation in signal transduction pathways can lead to a variety of complex disorders, including cancer. Computational approaches such as network analysis are important tools to understand system dynamics as well as to identify critical components that could be further explored as therapeutic targets. Here, we performed perturbation analysis of a large-scale signal transduction model in extracellular environments that stimulate cell death, growth, motility, and quiescence. Each of the model's components was perturbed under both loss-of-function and gain-of-function mutations. Using 1,300 simulations under both types of perturbations across various extracellular conditions, we identified the most and least influential components based on the magnitude of their influence on the rest of the system. Based on the premise that the most influential components might serve as better drug targets, we characterized them for biological functions, housekeeping genes, essential genes, and druggable proteins. The most influential components under all environmental conditions were enriched with several biological processes. The inositol pathway was found as most influential under inactivating perturbations, whereas the kinase and small lung cancer pathways were identified as the most influential under activating perturbations. The most influential components were enriched with essential genes and druggable proteins. Moreover, known cancer drug targets were also classified in influential components based on the affected components in the network. Additionally, the systemic perturbation analysis of the model revealed a network motif of most influential components which affect each other. Furthermore, our analysis predicted novel combinations of cancer drug targets with various effects on other most influential components. We found that the combinatorial perturbation consisting of PI3K inactivation and overactivation of IP3R1 can lead to increased activity levels of apoptosis-related components and tumor-suppressor genes, suggesting that this combinatorial perturbation may lead to a better target for decreasing cell proliferation and inducing apoptosis. Finally, our approach shows a potential to identify and prioritize therapeutic targets through systemic perturbation analysis of large-scale computational models of signal transduction. Although some components of the presented computational results have been validated against independent gene expression data sets, more laboratory experiments are warranted to more comprehensively validate the presented results.

**Keywords:** computational modeling, *in silico* perturbation analysis, signal transduction, cancer, therapeutic targets

## INTRODUCTION

Recent advances in systems biology and computational biology have introduced methods for the visualization, comprehension, and interpretation of big data in biomedical research. These fields provide an array of methodologies including computer simulations that can be used to generate new hypotheses and identify which hypotheses might be more productive to undertake experimentally, and eliminate hypotheses with little chance of success (Kitano, 2002a,b; Ghosh et al., 2011). These methods can be effective in navigating complex network problems associated with diseases. Many diseases and pathologies can be characterized by the dysregulation or dysfunction of multiple molecular components that are connected within these highly intertwined biological and biochemical networks (Loscalzo and Barabasi, 2011). Biological networks, including biochemical signal transduction networks, consist of a large number of highly interconnected pathways that give rise to complex, non-linear dynamics governing various cellular functions (Helikar et al., 2008; Helikar and Rogers, 2009). Disruptions of these networks, such as mutations or disease states can have drastic effects upon the whole system. These effects are difficult to predict from static network diagrams.

However, understanding the hierarchy of these changes remains a paramount problem. Often the specific causal interactions of the disease state are hidden within the massive cell-wide alterations, making attempts to reverse a disease state more challenging. In addition, the specific causal interactions are difficult to predict making the development of a potential therapeutic target results in unforeseen side effects (Singh and Singh, 2012). The unwanted effects of these drugs are often drastic as seen with many cancer medications (Kayl and Meyers, 2006; Lotfi-Jam et al., 2008; Singh and Singh, 2012). These challenges are further exacerbated by drug resistance that can render therapies ineffective. Therefore, it is necessary to gain a systems level understanding of the components associated with the disease states.

In recent years, targeted therapy has been used for multiple diseases, e.g., cancer (Vanneman and Dranoff, 2012), and often involve the activation or inactivation of a specific component in a biological network by a small molecule or drug, for instance. Perturbation analyses allow one to interrogate the structure and dynamic footprint of the underlying biological system. Perturbation biology has been proposed as an approach to reduce the collateral damage caused by non-specific drugs. Computational network perturbations and new methods to evaluate the robustness of a given network can help identify more effective network components to target in order to obtain desired outcomes with minimal disruption to the rest of the network (Molinelli et al., 2013).

In order to fully leverage the potential of computational network perturbation analyses large dynamical models are necessary. A wide spectrum of modeling approaches exists ranging from detailed (but less scalable) differential equation-based systems to large (but not dynamic) static networks. In the middle are approaches such as logical modeling that are relatively scalable while capable of capturing the dynamic nature of biological systems (Le Novère, 2015). Logical networks, namely Boolean networks, have been used to describe and simulate a

wide spectrum of biological systems ranging in size as well as contextual application (Naldi et al., 2010; Helikar et al., 2012; Madrahimov et al., 2013; Rocha et al., 2013; Conroy et al., 2014). Thus, applying perturbation analysis to large-scale logical models may provide new insights into the system, which could be used to identify novel therapeutic targets.

Herein, we present results from a system-wide perturbation analysis of a large-scale Boolean model of a signal transduction network widely present in many types of cells. Specifically, the model previously described in Helikar et al. (2008) represents signaling events within the integrated epidermal growth factor (EGF), G protein-coupled receptor, and integrin signaling network. The model consists of 137 components (mostly proteins) and 557 biochemical interactions. The simulation-based, system-wide perturbation analyses enabled us to identify the most and least influential components (ones with the most and least impact on the rest of the network). To explore the role and effects of these perturbations in the context of the complex extracellular environment, the simulations and analyses were conducted under four biologically relevant environmental conditions known to stimulate cell growth, cell death, motility, and quiescence (in addition to a set of randomly generated environmental stimuli). In order to investigate potential therapeutic targets, we performed functional annotation and analysis of the most influential signal transduction components under both inactivating (e.g., knockout) and activating (e.g., overexpression) perturbations. The most influential components were found to be enriched with many biological processes and druggable targets. Also, the most influential components under activating perturbations were enriched with more essential genes than the least influential components. We used the most influential components and their upstream regulators to identify novel interactions. We also identified a network of the most influential components consisting of drug targets considered in multiple cancer types. The highest ranked among the most influential components were already explored as drug targets against cancer, including EGFR, PI3K, Raf, Ras, and Erk. Because some of these targets have been reported to be associated with drug resistance (Holohan et al., 2013; Rodon et al., 2013; Wagle et al., 2014), we analyzed additional components of the signal transduction network that could potentially complement drug-resistant targets. As a result of the systemic analysis, we identified one novel combinatorial target, PI3K-IP3R1, with consistent occurrence in all simulated environmental conditions. This combination could be used to suppress cell proliferation while increasing the rate of apoptosis. We simulated the effect of combinatorial perturbation and the results were correlated with the literature, further supporting our predictions.

## MATERIALS AND METHODS

### Computational Model

The computational model analyzed in this work is a Boolean model of signal transduction in a generic cell type. In Boolean models, each component can assume an active (1) or inactive

(0) state at any time  $t$ . The activity state of the model's internal components is determined by the regulatory mechanisms of other directly interacting components. These regulatory mechanisms are described with Boolean functions (in the form of truth tables or Boolean expressions). To represent the milieu of stimuli in the extracellular environment, the model contains external components that represent various ligands. The activity level of these components is specified as a probability to simulate different levels of concentrations. This methodology was previously detailed and exemplified in Helikar et al. (2008, 2012), Helikar and Rogers (2009), and Todd and Helikar (2012).

The signal transduction model, previously detailed in Helikar et al. (2008), was constructed manually from around 500 published papers. The model consists of several main signaling pathways, including the receptor tyrosine kinase (EGF receptor), G protein-coupled receptors (G-alpha  $i$ , G-alpha  $q$ , G-alpha  $s$ , and G-alpha 12/13), and the integrin signaling pathways. Each of the 130 components in the model corresponds to a signaling molecule (mainly protein). The model also contains nine external components that represent the extracellular environment (mostly composed of receptor ligands). These external components include the EGF, extracellular matrix (ECM), calcium pump, interleukin 1, and tumor necrosis factor (TNF), ligands for four types of G protein-coupled receptors ( $\alpha_i$ ,  $\alpha_q$ ,  $\alpha_s$ , and 12/13), and a general stress signal. The final model consists of 137 components (130 internal and 7 external) connected with 557 interactions. The model is fully annotated and freely available *via* the Cell Collective software (Helikar et al., 2012, 2013) at [www.thecellcollective.org](http://www.thecellcollective.org) (under Published Models). Cell Collective, an interactive modeling environment, can be used to download the model (and other logical models published by the community) in several file formats (SBML qual, text file of logical functions, truth tables, etc.), as well as simulate directly on the platform. For convenience, the model SBML file is provided as File S1 (Data Sheet 1) in Supplementary Material.

## Model Simulations

The Cell Collective platform was used to perform all computational simulations of the model. Although the model is built by using discrete mathematics the output activity levels (AL) can be continuous (ranging from 0 to 100) as previously described in Helikar et al. (2008) and Helikar and Rogers (2009). Each simulation is synchronous and consists of 800 steps, where the activity level of the measured output component is calculated as the fraction of ones (active states) over the last 300 iterations that describe the network's steady behavior (Helikar et al., 2008; Helikar and Rogers, 2009).

Let  $x_j(t_i)$  denotes a node's activity on the  $i$ th iteration and  $j$ th simulation where  $i = 1, 2, \dots, T$  and  $j = 1, 2, \dots, N$ , total is the simulation out of  $N$  total simulations. We obtain AL as below.

$$AL = \frac{1}{300N} \sum_{j=1}^N \sum_{i=T-300}^T x_j(t_i)$$

The model was simulated and analyzed under four biologically relevant environmental conditions that stimulate cell growth, cell

death, quiescence, motility (and randomly generated behaviors), as established and detailed in Helikar et al. (2008). The environmental conditions that stimulate each of these cellular responses were obtained based on Helikar et al. (2008) where the model's responses were characterized based on 10,000 combinations of randomly generated environmental signals. For example, cell growth behavior is characterized by higher AL of Erk (marker for proliferation) and Akt (marker for anti-apoptosis). Cell motility behavior was characterized by higher AL of Cdc42 and Rac. Quiescence response is considered when the activity level of Akt is medium to low, and proliferation (Erk) and motility (Cdc42 and Rac) are low or inactive (Helikar et al., 2008). Each environmental condition is defined by different combinations of AL of external components (ligands). The activity level ranges of the environmental conditions were further determined by an optimization method whereby 2,000 simulations were run with all external stimuli ranging from 0 to 100 (except for IL1\_TNF and Stress that were limited to low AL). Subsequently, environmental activity level combinations that stimulated cell growth, cell death, motility, and quiescence most effectively were selected as the corresponding environmental conditions (**Table 1**). This is directly analogous to optimization experiments in laboratory studies (e.g., determining the optimal medium and plating conditions of a cell before performing a growth factor titration).

A wild type (WT) experiment (used as a reference) was conducted under each environmental condition without any perturbations. Subsequently, systematic perturbation experiments were conducted under each condition, whereby each component of the model was constitutively activated (activity stuck at 1; gain-of-function/overexpression) or inactivated (activity stuck at 0; loss-of-function/knockout). Each experiment consisted of randomly selecting 100 combinations of AL of the external stimuli from each condition activity range. (The only exception was the random environmental condition, which was simulated 2,000 times.) Each of the 100 combinations were simulated 30 times (i.e., 30 replicates) to ensure consistency of the dynamics in response to a specific combination of stimuli. These replicates were subjected to a Fligner Killeen test of homogeneity of variances, which confirmed that the measured AL of the network components, were homologous for identical combinations of AL of the environmental stimuli.

## Model Analysis

The Kolmogorov–Smirnov (KS) test (Wang et al., 2003) was used to compare the WT dynamics (under each environmental condition) with the dynamics of each perturbation experiment. If the KS test resulted in a  $p$ -value  $< 0.05$ , then it has a difference value (DV) equal to the test statistic; otherwise, the DV for a component is 0. Because we are looking at how a node's perturbation affects the rest of the network, its DV when it is the perturbed node is set to 0.

## Most and Least Influential Components

The most influential components are defined as components that induce the largest changes in the network under a given perturbation. The ranking of the perturbations is derived by calculating

**TABLE 1 | Activity level ranges of environmental stimuli for cell death, growth, motility, quiescence, and random environments.**

| External                        | Death | Growth | Motility | Quiescence | Random |
|---------------------------------|-------|--------|----------|------------|--------|
| Extracellular matrix (ECM)      | 10–72 | 26–82  | 81–99    | 7–30       | 0–100  |
| Epidermal growth factor (EGF)   | 3–15  | 72–97  | 29–83    | 43–56      | 0–100  |
| Calcium pump (ExtPump)          | 35–87 | 24–83  | 41–92    | 17–82      | 0–100  |
| GPCR q ligand (alpha_qL)        | 13–58 | 18–78  | 17–74    | 4–84       | 0–100  |
| GPCR i ligand (alpha_iL)        | 1–4   | 15–77  | 30–82    | 31–83      | 0–100  |
| GPCR s ligand (alpha_sL)        | 30–87 | 24–80  | 20–77    | 19–46      | 0–100  |
| GPCR 12/13 ligand (alpha_1213L) | 14–65 | 18–78  | 12–77    | 18–67      | 0–100  |
| IL1_TNF                         | 4–13  | 8–15   | 4–13     | 2          | 2      |
| Stress                          | 2–5   | 2–5    | 2–5      | 2–3        | 2      |

an influence score (IS) for the  $i$ th node, which is found by summing the DV for all  $M$  nodes in the network. The top 10% are considered most influential, and the bottom components with IS value 0 were considered the least influential. The cutoffs were set to 10% because only a few components had a high influence on the network.

$$IS_i = \sum_{j=1}^M DV_{ij}$$

$$i = 1, \dots, 130$$

## Most Affected Components to a Specific Perturbation

For each perturbation induced, the components that are most sensitive to that perturbation are ranked in decreasing order to be able to characterize downstream effects of the perturbation on the network.

## Annotation and Biological Relevance of Signal Transduction Components

All model components were first annotated using the appropriate NCBI gene IDs (Pruitt et al., 2007) for associated genes and UniProt IDs (Consortium, 2011) for protein products of the genes. All components were then further characterized using online resources such as DrugBank (Wishart et al., 2006).

The biological process enrichment analysis of the most influential components was done using DAVID (Huang et al., 2008), with high stringency. Gene Ontology (Ashburner et al., 2000), SP\_PIR keywords, and KEGG pathways (Kanehisa, 2002) were obtained using FDR < 5%.

Essentiality data were obtained from the Online GENE Essentiality (OGEE) database and mapped on the most and least influential components (Chen et al., 2012). DrugBank data were used to obtain druggability information for each component in the network. Data on cancer-associated genes were obtained from The Cancer Genome Atlas (TCGA) (Weinstein et al., 2013) and mapped on the most influential components to identify cancer-associated most influential components. The enrichment of essential genes and druggable proteins was computed based on the number of genes mapped on most or least influential components out of the total number of most and least influential components.

## Network Motif Analysis

Network motif analysis in the directed signal transduction network was performed using FANMOD tool (Wernicke and Rasche, 2006). The default parameters were used that include 100,000 samples to determine the sub-graphs. The significance of network was computed by comparison with 1,000 random networks. Network motifs that have occurrence more than five times and  $p$ -value < 0.05 were considered as significant. Network motif analysis was previously integrated with logical modeling of signal transduction of epithelial–mesenchymal transition (Steinway et al., 2014).

## Gene Expression Analysis

To investigate the functional activity of the components of signal transduction model, we queried publicly available gene expression data in four different databases (Consortium, 2011); Bgee (gives activity level of genes across different species as well as different developmental stages) (Bastian et al., 2008), CleanEx database (Providing heterogenous data from different) (Praz et al., 2004), Expression Atlas database (gene expression data under different biological conditions) (Petryszak et al., 2014), and GeneVisible database (Gene expression in different tissues) (Zimmermann et al., 2004). Out of 109 signal transduction components i.e., proteins, 107 (~98%) showed expression across different species, developmental stages, organs, and tissues – suggesting the biological activity of signal transduction network. Gene expression status of signal transduction components is shown in the Table S1 in Supplementary Material.

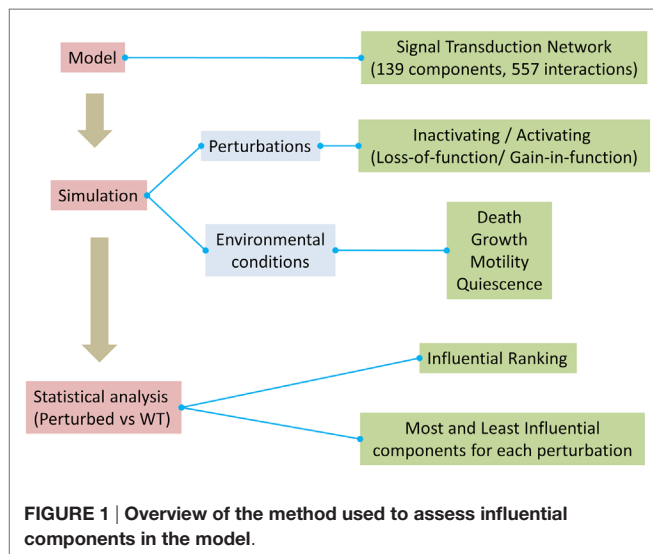
The gene expression dataset GSE53309 was obtained from the GEO database (Barrett et al., 2005; Rosich et al., 2014). We selected samples that were treated with pan-PI3K inhibitor and of normal control. The log<sub>2</sub> RMA signal intensities of samples were transformed into Z-scores (Cheadle et al., 2003). To compare the Z-scores of treated samples with normal control, we used Z-ratio approach. Genes with Z-ratio ≥ 1.50 were considered upregulated and with ≤ -1.50 were considered as downregulated. The Z-ratio cut-off (1.5) is previously found as robust (Cheadle et al., 2003). The genes of signal transduction components whose AL were affected as a result of PI3K inactivation were examined for Z-ratios in both the biological replicates. We used DAVID to perform biological process enrichment analysis of upregulated and downregulated genes. The high stringency and FDR < 5% were used to select significantly enriched biological processes.

## RESULTS

### System-Wide Perturbation Analysis Reveals Core Components of the Signal Transduction Network

A critical objective of biomedical research is the identification and prioritization of novel therapeutic targets. In this context, we performed systematic perturbation analysis in a generic signal transduction model. The workflow used in this work is illustrated in **Figure 1**.

The activating/inactivating perturbation experiments for each component in the model were carried out across four environmental conditions (as described in the Section “Materials and Methods”). Additional randomly generated extracellular conditions were used to check the robustness of the model and results. Perturbation analysis enabled us to identify and rank components of the signaling network that are most and least influential (Table S2 in Supplementary Material). The heatmaps for all the environmental conditions [Figures S1–S10 (Image 1) in Supplementary Material] indicate that a few components had high influence on rest of the system. Therefore, we considered the top 10% of the components from each condition as the most influential. By contrast, the components that had no influence on the system were considered as the least influential (KS = 0).



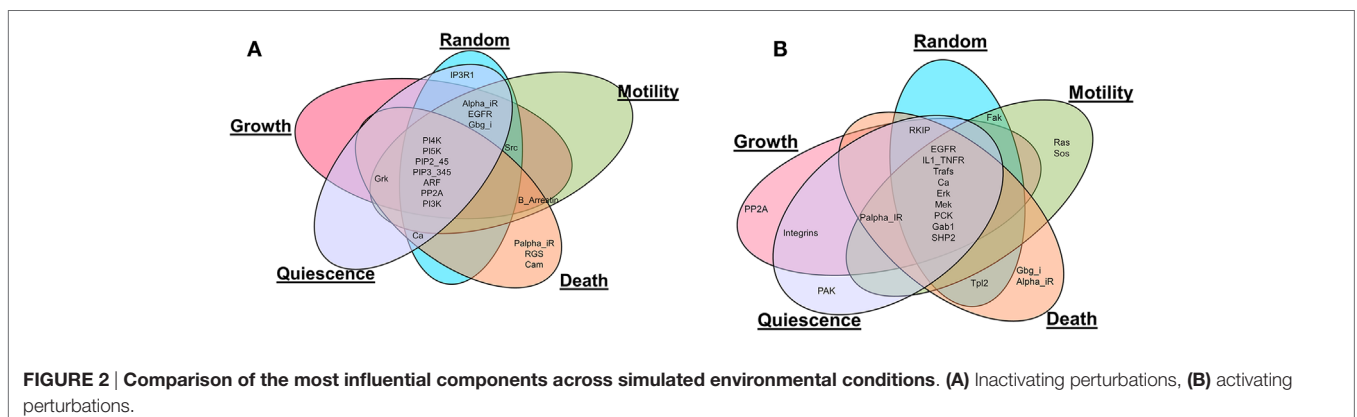
Also, the most influential components correspond to network components that, when perturbed, affect the largest part of the network in terms of the number of affected components and the magnitude of the effect. The most influential components were found for both inactivating (**Figure 2A**) and activating (**Figure 2B**) perturbations under the different environmental conditions. It is interesting to note that many of the most influential components overlap across all environmental conditions. However, the most influential components do not overlap between two types of perturbations (inactivating or activating). We investigated whether the most influential components that spanned different environmental conditions could function as housekeeping genes. Also, the most influential components that are specifically found under one environmental condition should have association with that condition.

### Housekeeping Genes Are Enriched in the Most Influential Components Common in Different Environments

Housekeeping genes are defined as genes expressed at constant level in many cells and under many conditions (Eisenberg and Levanon, 2013). Therefore, components that were identified as most influential under all of the simulated environmental conditions can be hypothesized to have housekeeping function. To investigate this, we compared these most influential components with known housekeeping genes as provided in Eisenberg and Levanon (2013). Under inactivating perturbations, out of the seven components common among the different environmental conditions, PI4K, PI5K, ARF, and PI3K were associated with housekeeping genes (Eisenberg and Levanon, 2013). Under activating perturbations, Trafs, Erk, Mek, and SHP2 (out of nine common components), were associated with housekeeping genes. Housekeeping genes associated with the common components are displayed in **Table 2**. This observation suggests that the most influential components that are common among different environmental conditions are likely to function as housekeeping genes.

### Unique Components Associated with Each Environmental Condition Are Found to Be Condition Specific

Under both types of perturbations, certain environmental conditions had several uniquely associated components (**Figure 2; Table 3**). Under inactivating perturbations, components uniquely



**TABLE 2 | Housekeeping genes in the most influential components overlapped among different environmental conditions.**

| Perturbation | Components | Genes   | Housekeeping genes <sup>a</sup> |
|--------------|------------|---|---------------------------------|
| Inactivating | PI4K       | PI4KA, PI4KB, PIK4CB  | PI4KA, PI4KB                    |
|              | PI5K       | PIP5K1A, PIP5K1B, PIP5K1C                                     | PIP5K1A                         |
|              | ARF        | ARFGAP1, ARFGAP2, ARFGAP3                                     | ARFGAP2, ARFGAP3                |
|              | PP2A       | PPP2CA  | PPP2CA                          |
|              | PI3K       | PIK3CA, PIK3CB, PIK3CD, PIK3CG                                | PIK3C3, PIK3CB                  |
| Activating   | EGFR       | EGFR  | No                              |
|              | IL1_TNFR   | IL1B, TNFRSF1A  | No                              |
|              | TRAF5      | TRAF1, TRAF2, TRAF3, TRAF4, TRAF5, TRAF6, TRAF7               | TRAF6, TRAF7                    |
|              | ERK        | MAPK1 to MAPK15   | MAPK1, MAPK6, MAPK8, MAPK9      |
|              | MEK        | MAP2K1 to MAP2K7  | MAP2K1, MAP2K2, MAP2K5          |
|              | PKC        | PRKCA, PRKCB, PRKCD, PRKCE, PRKCG, PRKCH, PRKCI, PRKCO, PRKCZ | No                              |
|              | GAB1       | GAB1  | No                              |
|              | SHP2       | PTPN11  | PTPN11                          |

<sup>a</sup>List of housekeeping genes were obtained from Eisenberg and Levanon (2013).

**TABLE 3 | Condition-specific components and literature support.**

| Perturbations | Environmental condition | Associated components | Literature   |
|---------------|-------------------------|-----------------------|--|
| Inactivating  | Death                   | CaM, RGS, Palpha_iR   | CaM- and CaM-dependent signaling systems control vertebrate cell proliferation, programmed cell death, and autophagy (Berchtold and Villalobo, 2014). RGS is involved in cell death (Fisher, 2009) |
| Activating    | Death                   | Gbg_i (GNB), Alpha_iR | Gbg_i has been hypothesized to be involved in mTOR-mediated anti-apoptotic pathways. Furthermore, it has been functionally annotated with apoptosis, cell death (Wazir et al., 2013)               |
|               | Growth                  | PP2A                  | Highly regulated family of Ser/Thr phosphatase implicated in cell growth and signaling (Janssens and Goris, 2001)  |
|               | Motility                | KRAS, Sos             | Knockdown of KRAS in pancreatic cancer cell lines leads to decreased motility and proliferation. The Grb2-Sos1 complex may promote cell motility, and tumorigenesis (Qu et al., 2014)              |

associated with the cell death stimulating condition are calmodulin (CaM), RGS, and Palpha\_iR. Out of these, CaM and RGS have been previously associated with cell death and apoptosis (Fisher, 2009; Berchtold and Villalobo, 2014). In fact, CaM plays a central role in the regulation of several cellular functions, including programmed cell death (Berchtold and Villalobo, 2014). It is also known that RGS protein can regulate cell death, cell cycle, and cell division (Fisher, 2009). Under activating perturbations, the most influential components associated with the cell death-inducing condition include Gbg\_i and Alpha\_iR. On the other hand, PP2A was found to be most influential under the growth stimulating condition, Ras and Sos under motility stimulating condition, and PAK under quiescence stimulating condition. These results are also further supported by published studies that reported Gbg\_i (GNB) to be involved in mTOR-mediated anti-apoptotic pathways; Gbg\_i was also functionally annotated with apoptosis and cell death (Wazir et al., 2013). PP2A was reported as a highly regulated Ser/Thr phosphatase involved in cell growth and signaling (Janssens and Goris, 2001). In pancreatic cancer cell lines, the knockdown of KRAS has been found to lead to the decrease in cell motility and proliferation (Rachagani et al., 2011; Birkeland et al., 2012). Furthermore, the Grb2-Sos1 complex has been found to most likely promote cell motility, and tumorigenesis (Qu et al., 2014). These observations suggest that the proteins, which were uniquely associated with simulated environmental conditions, are most likely to have the association with that condition. Finally,

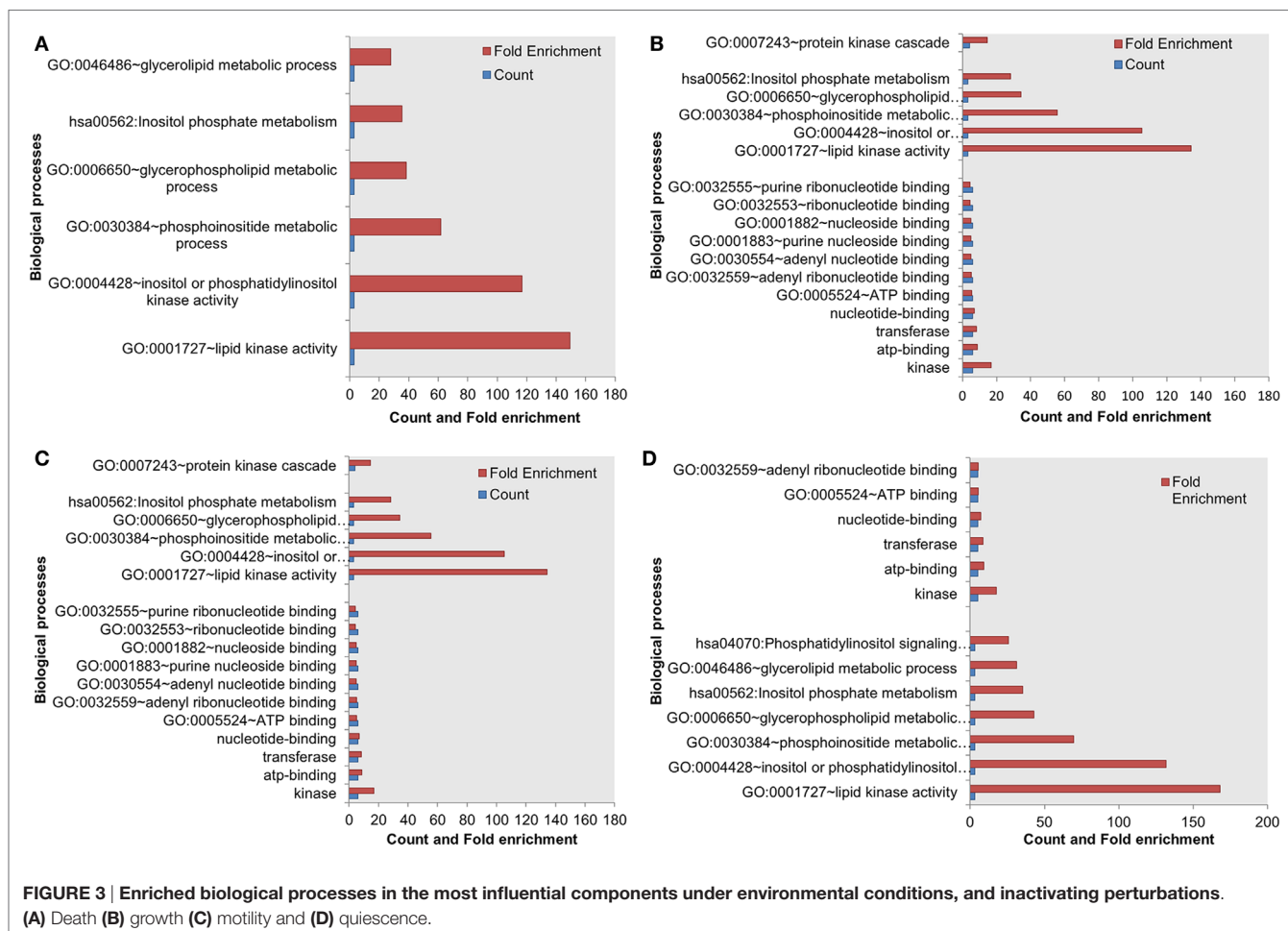
the literature evidence obtained for housekeeping, or condition associated genes, further supports our simulation results.

## Key Biological Processes Are Enriched in the Most Influential Components

Next, we assessed the enrichment of biological processes or pathways in the most influential components. The most influential components across all four conditions under both types of perturbation showed significant enrichment with key biological processes. The counts and fold differences of enriched biological terms in all the conditions are shown in **Figures 3** and **4**. In the case of inactivating perturbations, inositol phosphate metabolism was enriched under all environmental conditions (**Figure 3**). In the case of activating perturbations, the significantly enriched biological processes include phosphate metabolic processes, kinase activity, apoptosis, and, interestingly, the non-small lung cancer pathway (**Figure 4**). These results illustrate that the group of proteins with similar biological functions appear as the influential components under each type of perturbation.

## The Most Influential Components under Activating Perturbations Are Enriched with Essential Genes

Mutations in an essential gene can be lethal. Based on the hypothesis that the influential components might serve as essential for



the survival of the cell, we performed essentiality analysis. We mapped essential genes on the most influential components and on the least influential components. The essential genes mapped on the most influential components were compared with essential genes mapped on the least influential components. Under activating perturbations, more essential genes were found within the most influential components than the least influential components (Figure 5A). Under the cell death stimulating condition, a total of 69% of the most influential components were essential; this is in contrast to the least influential components that contained 31% essential genes. Under other environmental conditions stimulating growth, motility, and quiescence, the difference of essential genes between the most influential and the least influential components are 23, 15, and 32%, respectively.

On the other hand, under inactivating perturbations, we found either an equal or larger number of essential genes in the least influential components (Figure 5B). The most significant differences were observed under the cell death stimulating conditions: the least influential components have 66% of essential genes in contrast to the 46% essential genes in the most influential. Also, under the growth stimulating conditions, 68 and 53% of essential genes were contained within the least and the most influential components, respectively. Under the motility and quiescence stimulating conditions, there were 3 and 9% more essential genes

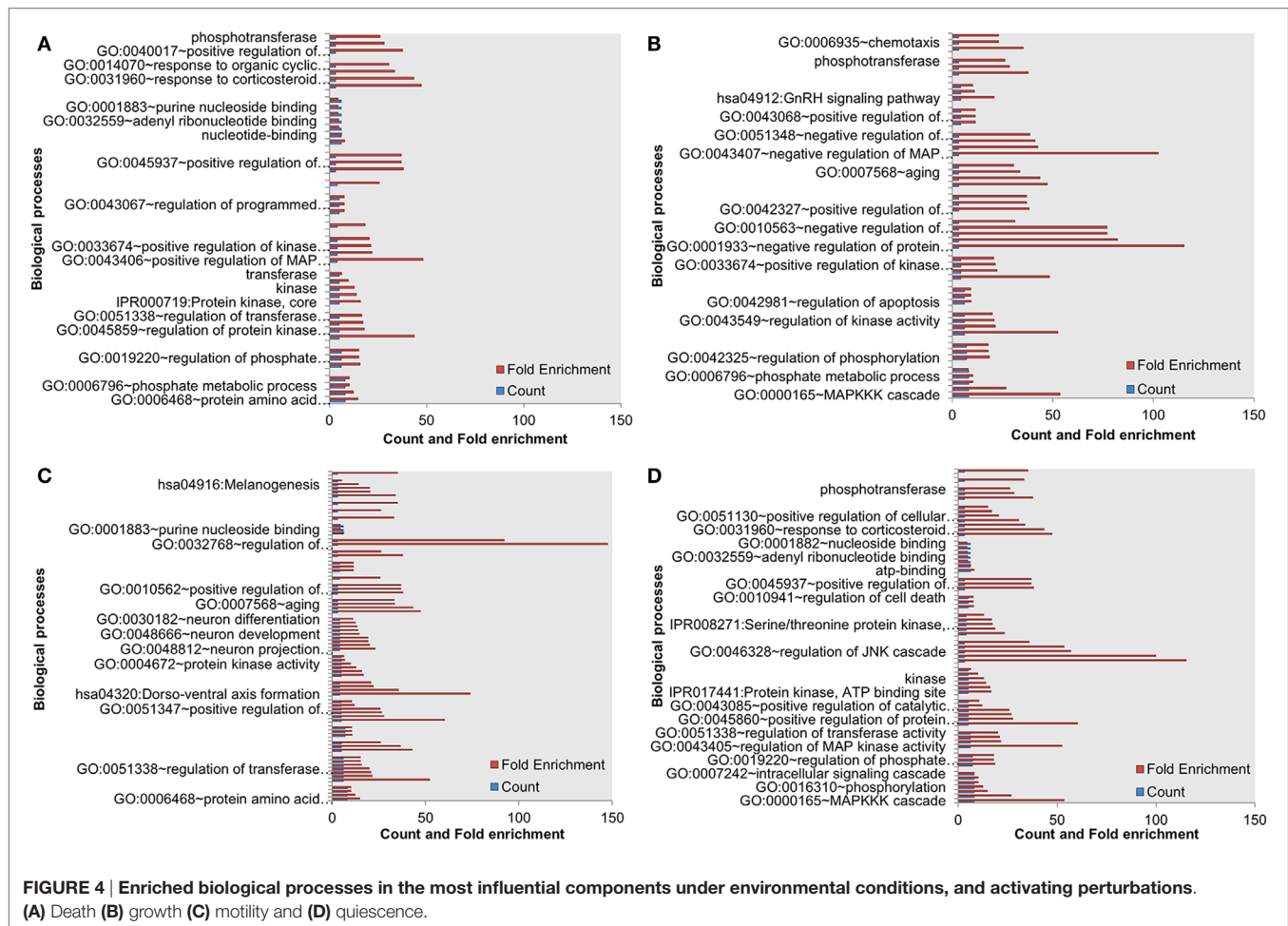
within the least influential components than the most influential components, respectively. We found that under inactivating perturbations, the number of essential genes among the least influential components was slightly larger than the activating perturbation (Figures 5C,D). On the other hand, under activating perturbations, the more essential genes mapped within the most influential components than the least influential components.

Thus, the most influential components are essential under activating perturbations, suggesting an environmental condition-specific essentiality.

## The Most Influential Components Are Enriched with Druggable Proteins

To further investigate the importance of the most influential components, we evaluated the distribution of known druggable targets. We obtained druggability data from the DrugBank database (Wishart et al., 2006) and mapped them on the most and least influential components. A total of 51 components in the whole network were enriched with druggable proteins. We compared druggable proteins within the most influential components with druggable proteins within the least influential components. We found that under both types of perturbations and across all environmental conditions more druggable proteins were found among the most influential than the least influential components





(Figure 6). Druggable proteins are experimentally characterized or predicted to bind to antagonist or agonist drugs with high affinity. Therefore, enrichment of druggable proteins within the most influential components has the potential to suggest important candidates for therapeutic target discovery.

## The Most Influential Components as Drug Targets

### Ranked Most Influential Components Based on Downstream Components

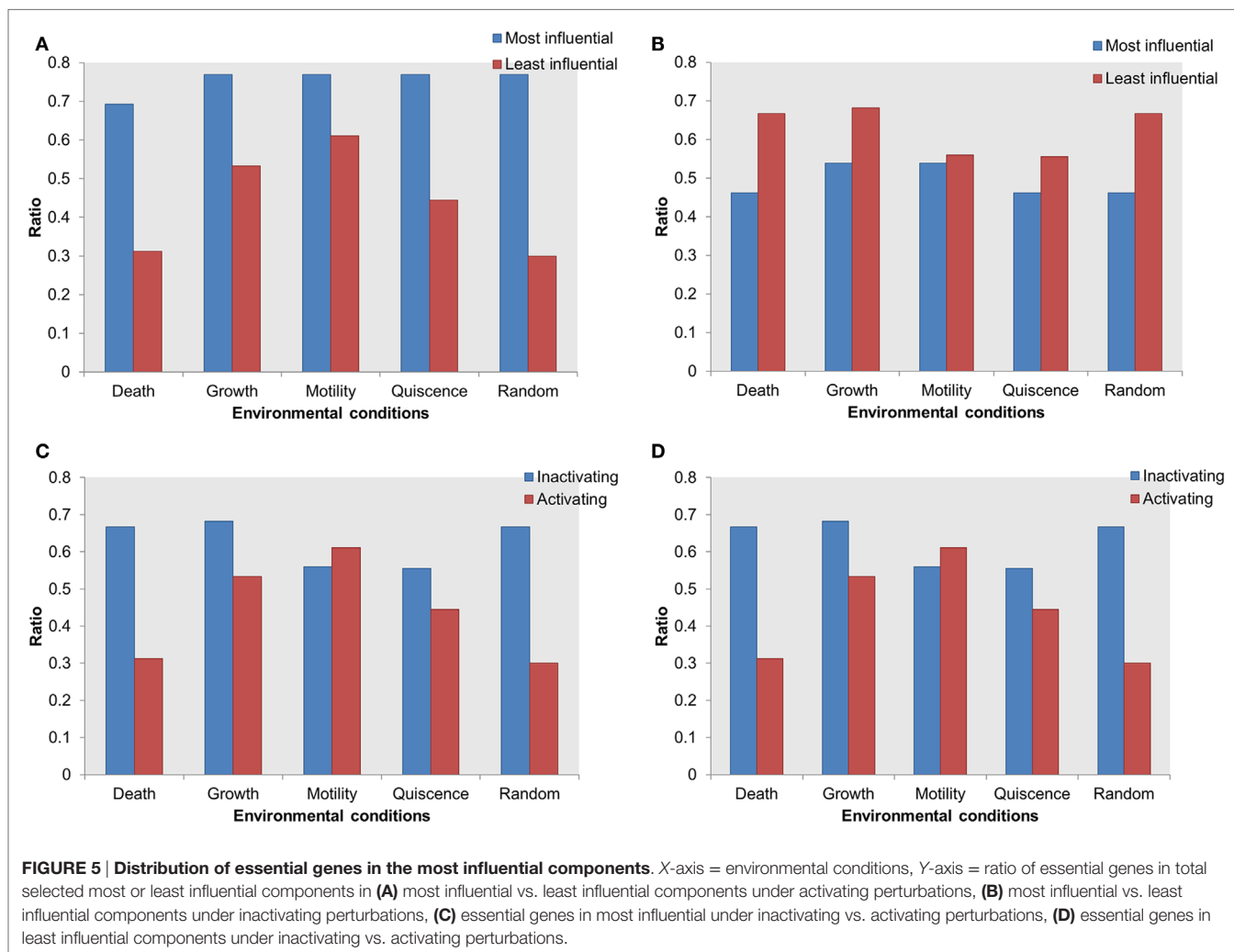
We identified the most affected components of the most influential components under both types of perturbations. We combined all environmental conditions to construct networks of the most influential components with their downstream targets. We subsequently mapped druggable proteins and cancer-associated genes on these networks. Under inactivating perturbations, we obtained a network consisting of the most influential components: PI3K, EGFR, PP2A, GRK, and CaM (Figure 7A). Under activating perturbations, we obtained a network composed of influential components: EGFR, IL1\_TNFR, ERK, SHP2, RKIP, Ras, Gbg\_i, Fak, Integrins, and PP2A (Figure 7B).

The total number of downstream targets for each of the most influential druggable component under both inactivating and activating perturbations is listed in Table 4. We analyzed if these

downstream components also affects their upstream component. In the case of PI3K-out of 42 downstream components, two (PIP3\_345 and RGS) are part of a feedback system. Other feedback components in downstream targets include alpha\_iR for GRK in inactivating perturbations, Gab1 for SHP2, and Palpha\_iR for RKIP under activating perturbations. We observed that EGFR, a validated cancer drug target (Mendelsohn, 2001), affects the largest number of components under activating and inactivating perturbations.

### The Most Influential Components Mainly Affect Other Most Influential Components

Here, we identified all components that directly affect the activity of each most influential component ( $KS = 1$ ). Interestingly, most of these direct upstream components were also ranked as the most influential in at least one environmental condition (Figure 8). Under inactivating perturbations, 22 components were directly upstream of the most influential components. Of these, 19 were the most influential under at least one environmental condition. On the other hand, under activating perturbations, out of 45 upstream components, 19 were also ranked as most influential. Additionally, under inactivating perturbations, 9 (CaM, EGFR, Gbg\_i, GRK, IP3R1, PP2A, PI3K, Ras, and Src) out of total 22 upstream components are druggable. Out of these 22 components, 6 components (CaM, EGFR, Gbg\_i, GRK, IP3R1, and

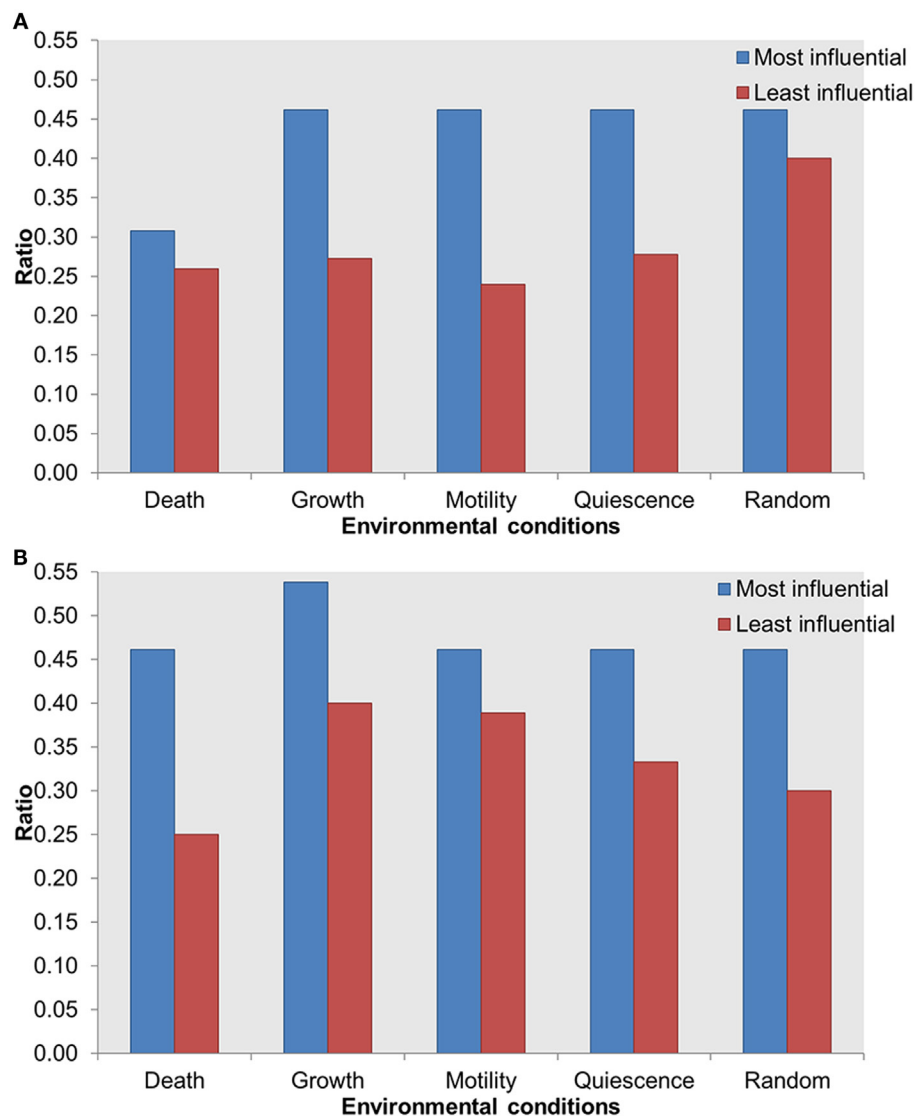


PP2A) were upstream to the most influential druggable components. Under activating perturbations, 21 (CaM, Cdc42, EGFR, Erk, Fak, Gbg\_i, Grb2, GRK, IL1\_TNFR, Integrins, IP3R1, PDK1, PI3K, PKA, PP2A, Rac, Raf, Ras, RKIP, SHP2, and Src) out of 45 upstream to the most influential components are associated with druggable proteins. Out of these 21, 10 components were also the most influential. Under both types of perturbations, a total of 18 (alpha\_iR, ARF, B\_Arrestin, Ca, CaM, EGFR, Gbg\_i, GRK, IP3R1, Palpha\_iR, PI5K, PIP2\_45, PIP3\_345, PP2A, RGS, PI3K, Ras, and Src) upstream components were common. Nine of these components (CaM, EGFR, Gbg\_i, GRK, IP3R1, PP2A, PI3K, Ras, and Src) were druggable or these were used as the drug targets. The important drug targets, such as EGFR, PI3K, Ras, and Raf, are also appeared as influential upstream components. Together, these results suggest that under inactivating perturbations the activity of the most influential components are likely to be modulated by the other most influential components.

### The Most Influential Components as Drug Targets and Drug Resistance

The top most influential components, such as EGFR, PI3K, ERK, and Ras, have been previously explored as drug targets

in multiple cancer types. However, it is also evident from literature that several most influential components have been associated with drug resistance. For example, in non-small cell lung cancer, mutation within the kinase domain of EGFR and epithelial-mesenchymal transition are responsible for the development of resistance to gefitinib (Holohan et al., 2013). In colorectal, and head and neck cancers, KRAS mutation, EGFR-S492R mutation, and increased ErBb signaling are responsible for resistance against Cetuximab (Dienstmann et al., 2012; Holohan et al., 2013). Furthermore, PI3K showed drug resistance in breast cancer against rapamycin through the expression of RSK3 and RSK4 (Rodon et al., 2013). Mutations in ERK1 or ERK2 have shown resistance against ERK inhibitors or RAF/MEK inhibitors (Wagle et al., 2014). Tumors with mutation in BRAF V600E can adapt to the RAF inhibitors (Lito et al., 2013; Perna et al., 2015). As such, the identification and prediction of drug targets alone are not sufficient to identify completely useful drug targets. Investigation of the interactions and feedback of these most influential components could be useful to modulate the activity of the most influential component. Thus, we explored the regulatory interactions to investigate the effect of combinatorial perturbations on cell's behavior.

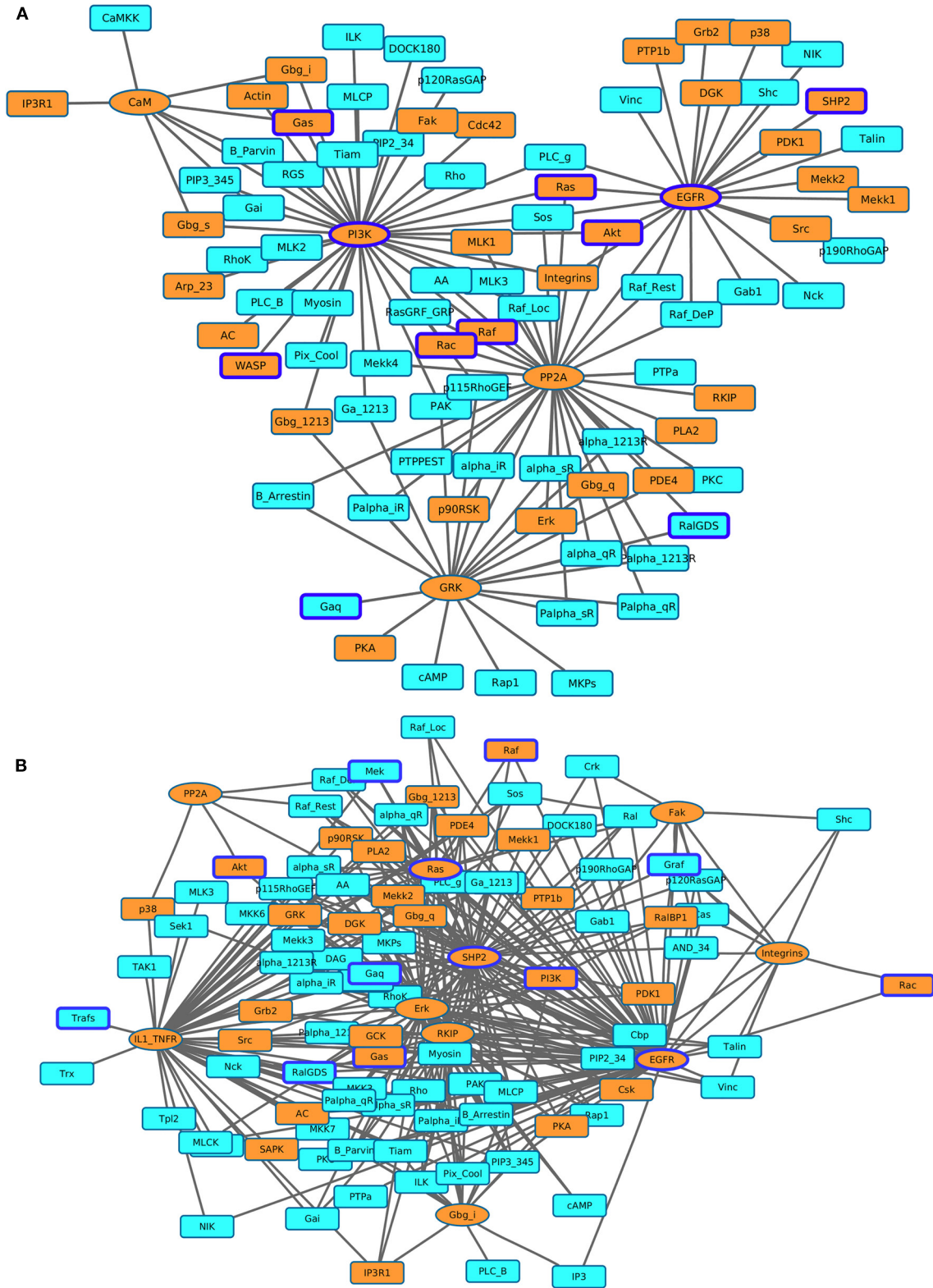


**FIGURE 6 | Distribution of druggable proteins within the most influential vs. least influential components. (A)** Inactivating perturbations, **(B)** activating perturbations. X-axis = environmental conditions, Y-axis = ratio of druggable proteins in total most or least influential components.

## Regulatory Interactions between the Most Influential Components and Their Upstream Components

To develop a better strategy that can account for drug resistance of the most important drug targets, we sought to investigate novel regulatory interactions. We analyzed the previously described interactions between the most influential components and their direct upstream components. We found that some interactions consistently occur in more than one environmental condition. For example, the inactivation of IP3R1 increases the activity of PI3K under all four environmental conditions. However, the maximal effect was observed under the death environmental condition. Additionally, the inactivation of IP3R1 leads to inactive RGS under three environmental conditions stimulating cell growth, motility,

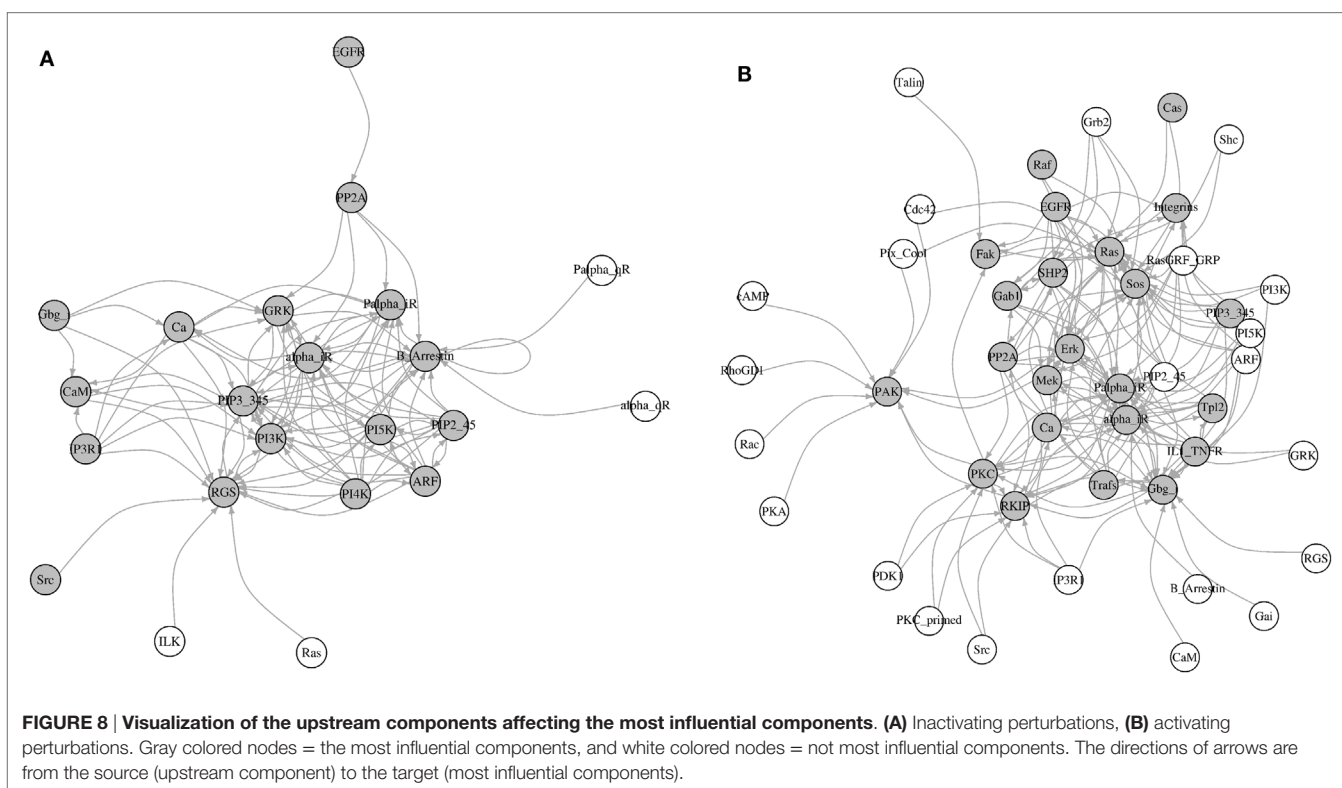
and quiescence. These findings also correlate with published studies that found that RGS positively regulates apoptosis (Fisher, 2009). Other examples of consistently occurred interactions include: the activation of Grb2 leads to increased AL of Ras under all four environmental conditions, and increased Sos activity under two environmental conditions stimulating death and quiescence. The activation of Rac increases the activation of PAK under environmental conditions stimulating cell death and growth. Overall, we found three types of interactions: inactivation of one component leads to the increase of activity of another component (PI3K–IP3R1, IP3R1–PI3K, and RGS–IP3R1), inactivation of a component leads to decreased activity of another component (IP3R1–RGS), and activation of a component leads to increased activity of another component (Grb2–Ras, Grb2–Sos, and Rac–PAK).



**FIGURE 7 | Visualization of the most affected components (KST value = 1) as a result of perturbing the most influential druggable components. (A)** Inactivating perturbations, **(B)** activating perturbations. Orange colored eclipses = most influential druggable components; squares = affected components; orange colored squares = affected druggable components; components with blue borders = experimentally found to be associated with cancer.

**TABLE 4 |** Number of downstream targets of the most influential druggable components.

|           | Number of affected components | Number of affected druggable components | Number of cancer-associated components | Feedback components | Perturbation |
|-----------|-------------------------------|---|--|---------------------|--------------|
| EGFR      | 70                            | 25                                      | 8                                      |                     | Inactivating |
| EGFR      | 24                            | 13                                      | 3                                      |                     | Activating   |
| IL1_TNFR  | 54                            | 14                                      | 5                                      |                     | Activating   |
| Erk       | 54                            | 21                                      | 8                                      |                     | Activating   |
| SHP2      | 53                            | 17                                      |  | 1 (Gab1)            | Activating   |
| RKIP      | 43                            | 12                                      | 4                                      | 1 (PalphaiR)        | Activating   |
| PI3K      | 42                            | 17                                      | 7                                      | 2 (PIP3_345, RGS)   | Inactivating |
| PP2A      | 36                            | 14                                      | 6                                      |                     | Inactivating |
| PP2A      | 5                             | 3                                       | 2                                      |                     | Activating   |
| Ras       | 30                            | 13                                      | 5                                      |                     | Activating   |
| GFRK      | 22                            | 5                                       | 2                                      | 1 (alpha_iR)        | Inactivating |
| Gbg_i     | 15                            | 5                                       | 1                                      |                     | Activating   |
| Fak       | 14                            | 6                                       | 4                                      |                     | Activating   |
| Integrins | 11                            | 3                                       | 3                                      |                     | Activating   |
| CaM       | 8                             | 5                                       | 2                                      |                     | Inactivating |



The fold differences of all these interactions are displayed in the **Table 5**. Under the cell death condition, the inactivation of IP3R1 results in PI3K activity increase by 2.38-fold. Similarly, PI3K inactivation leads to a 5.42-fold increase in IP3R1 activity. In the case of other interactions, the inactivation of IP3R1 leads to inactive RGS under the cell growth, motility, and quiescence stimulating conditions. Under the motility and quiescence stimulating conditions, the inactivation of Gbg\_i leads to inactive CaM. The activation of Grb2 increases the activity of Ras 7.40-fold

under the cell death stimulating conditions, and 2.13-fold under the quiescence stimulating conditions. Grb2 activation also affects Sos 7.8-fold under the cell death stimulating conditions and 2.18-fold under the quiescence stimulating conditions. An activating perturbation of Rac increases the activity of PAK more than 18-fold under the cell death stimulating conditions, and 5.59-fold under the growth stimulating conditions.

To investigate if these interactions are part of any network motifs in the signal transduction network, we performed a

**TABLE 5 | Fold differences of the affected most influential component when the upstream component was perturbed.**

| Perturbed component  | Affected component | Fold differences (perturbed/WT) |                       |                       |                       |
|----------------------|--------------------|---------------------------------|-----------------------|-----------------------|-----------------------|
|                      |                    | Death                           | Growth                | Motility              | Quiescence            |
| IP3R1 (inactivation) | PI3K               | 2.38-Fold <sup>a</sup>          | 1.03-Fold             | 1.04-Fold             | 1.14-Fold             |
| PI3K (inactivation)  | IP3R1              | 5.42-Fold <sup>a</sup>          | 1.18-Fold             | 1.15-Fold             | 1.24-Fold             |
| IP3R1 (inactivation) | RGS                | NSA                             | Complete inactivation | Complete inactivation | Complete inactivation |
| RGS (inactivation)   | IP3R1              | NSA                             | 1.21-Fold             | 1.18-Fold             | 1.24-Fold             |
| Gbg_i (inactivation) | CaM                | NSA                             | NSA                   | Complete inactivation | Complete inactivation |
| CaM (inactivation)   | Gbg_i              | NSA                             | NSA                   | 1.30-Fold             | 1.43-Fold             |
| Grb2 (activation)    | Ras                | 7.40-Fold <sup>a</sup>          | 1.32-Fold             | 1.39-Fold             | 2.13-Fold             |
| Ras (activation)     | Grb2               | 0.99-Fold                       | 0.97-Fold             | 0.99-Fold             | 1.01-Fold             |
| Grb2 (activation)    | Sos                | 7.87-Fold <sup>a</sup>          | 1.39-Fold             | 1.53-Fold             | 2.18-Fold             |
| Sos (activation)     | Grb2               | 1-Fold                          | 0.97-Fold             | 0.99-Fold             | 1.01-Fold             |
| Rac (activation)     | PAK                | 18.41-Fold <sup>a</sup>         | 5.69-Fold             | NSA                   | NSA                   |
| PAK (activation)     | Rac                | 1.18-Fold                       | 1.24-Fold             | NSA                   | NSA                   |

NSA, not significantly affected (*KST* value <1).

<sup>a</sup>Twofold or above change.

network motif analysis. We found that all interactions discussed above were part of network motifs (*p*-value <0.05). IP3R1–PI3K is found in 3 significantly occurred 4-node network motifs and in 15 significantly occurred 5-node network motifs. The other interactions are also found in significantly occurred 4 and 5-node network motifs (Table S3 in Supplementary Material).

These results suggest different types of regulatory effects of activating and inactivating perturbations of direct upstream components of the most influential components.

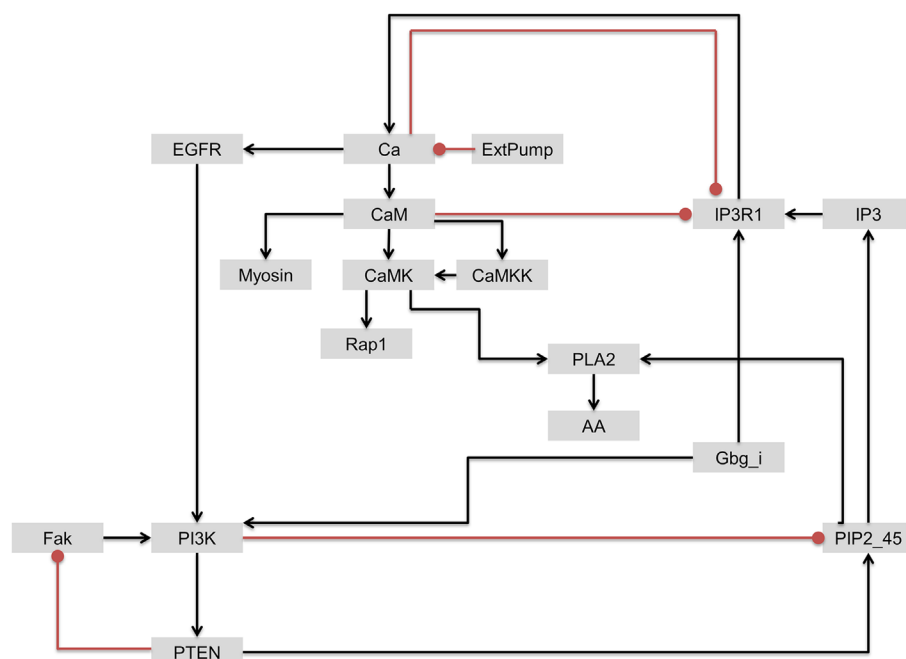
### Cotargeting IP3R1 with PI3K

As discussed earlier, although PI3K was identified as one of the most influential components, it has been also associated with drug resistance. Based on the interactions of upstream regulators of the most influential components discussed above, we further investigated the interactions involving PI3K and IP3R1 with the objective of identifying a secondary drug target that could be potentially used to address the issue of PI3K-associated drug resistance. In contrast to PI3K/Akt signaling, IP3R1 positively regulates apoptosis. We hypothesized that the rate of apoptosis will increase when IP3R1 is overactivated (activating perturbation) and PI3K is inactivated (inactivation perturbation). Despite the strong dynamical relationship between IP3R1 and PI3K, these two components are only connected indirectly through a sub-network. In this sub-network, Gbg\_i is upstream of and directly activates both components. IP3R1 regulates PI3K through a Ca → EGFR route, whereas PI3K regulates IP3R1 via a PTEN → PIP2\_45 → IP3 route (Figure 9).

The inactivating perturbation of PI3K resulted in the inactivation of 29 components across all four environmental conditions. To correlate PI3K inhibition results with laboratory experiments, we analyzed a gene expression dataset obtained from cells treated with PI3K inhibitors (Rosich et al., 2014). In two biological replicates, we found that the genes of components with affected AL had shown differential gene expression (at least in one experiment). As a result of the simulated constitutive inhibition of PI3K

in the model, the activity level of a total of 15 components (20 genes) increased more than twofold. Nine (60%) of these components were also significantly upregulated in the gene expression dataset (Table S4 in Supplementary Material). Out of these 20 genes, 9 genes (45%) were upregulated in biological replicate 1, whereas 12 (60%) genes were upregulated in biological replicate 2. Cumulatively, 18 genes (90%) were upregulated in both biological replicates. Two of these signal transduction components, Rap1 and PTPPEST, showed significant upregulation in both the biological replicates in gene expression data. Furthermore, the activity of a total of 26 components (41 genes) decreased more than twofold in our model. Genes of eight components (30%) were significantly downregulated in the obtained gene expression data (Table S4 in Supplementary Material). Out of these 41 genes, three genes (7%) were significantly downregulated in biological replicate 1, whereas eight genes (19.5%) were significantly downregulated in biological replicate 2. Cumulatively, 12 genes (29%) were upregulated in both biological replicates. Furthermore, we compared enriched biological processes within the components affected in the model with enriched biological processes in differentially expressed genes. We found that the “regulation of phosphorylation” biological process was enriched for the upregulated genes in both the model and the gene expression data. For downregulated components, “positive regulation of programmed cell death” was consistent for both the model and the gene expression data (biological replicate 1). Together, these results suggest that our simulation results are moderately correlated with the results of available gene expression data. In previous integrative studies of gene expression and biochemical models, at best moderate correlations were observed between gene expression and metabolic fluxes (Blazier and Papin, 2012). Post-transcriptional modifications and enzyme kinetics are possible reasons behind poor correlation between gene expression and protein abundance (Washburn et al., 2003; Blazier and Papin, 2012). As such, more laboratory experiments will be needed to further validate our results.

Under PI3K inactivation, the average activity of IP3R1 increased from 71.9% in WT to 85.18%. This perturbation also



**FIGURE 9 | The regulatory circuit connecting IP3R1 and PI3K and downstream components.** Edges with arrow = activation. Edges with oval end = inhibition.

**TABLE 6 | Activity of affected components under single (PI3K or IP3R1) and double perturbations (PI3K and IP3R1) under the cell growth environmental condition.**

| Affected components | PI3K inactivation (single perturbation) <sup>a</sup> (fold) | IP3R1 activation (single perturbation) <sup>a</sup> (fold) | Double perturbation <sup>a</sup> (fold) | Functional annotation <sup>b</sup>  |
|---------------------|---|--|---|---|
| Rap1                | 3.25  | 1.07   | 3.90                                    | Tumor-suppressor gene   |
| Ca                  | 1.17  | 1.41   | 1.43                                    | Calcium ion, apoptosis  |
| CaM                 | 1.17  | 1.41   | 1.43                                    | Cell death  |
| CaMKK               | 1.17  | 1.41   | 1.43                                    | Calcium ion binding, apoptosis  |
| Myosin              | 0.30  | 1.004  | 0.36                                    | Regulatory light chain of myosin  |
| CaMK                | 1.33  | 2.09   | 2.19                                    | May function in dendritic spine and synapse formation and neuronal plasticity |
| PLA2                | 0.32  | 1.24   | 0.63                                    | Tumor-suppressor gene, apoptosis  |
| AA                  | 0.32  | 1.24   | 0.63                                    | Apoptosis   |

<sup>a</sup>Compared to the activity of components in wild type.

<sup>b</sup>Functional annotations for proteins were obtained from UniProt database and literature.

led to downregulation of positive regulators of apoptosis phospholipase A2 (PLA2) and arachidonic acid (AA). AA released by PLA2 triggers Ca<sup>2+</sup>-dependent apoptosis through mitochondrial pathways (Penzo et al., 2004). The elevation in Ca<sup>2+</sup> is thought to be involved in apoptosis (Pinton et al., 2008). It was shown that blocking calcium channels can directly lead to tumor promotion (Mason, 1999). Thus, inactivation of PI3K can block cell proliferation; simultaneously, it can lower the rate of apoptosis. Interestingly, the positive regulation of the programmed cell death biological process was enriched in downregulated genes within the analyzed gene expression data.

Under the cell growth stimulating condition, the activating perturbation of IP3R1 increased the activity of apoptosis-associated

components: Ca, CaM, CaMK, CaMKK, and RGS in the range of +1.41- to +2.09-fold when compared to WT.

To simulate the cell death effect under the growth stimulating condition, we carried out a double perturbation of IP3R1 and PI3K, whereby IP3R1 was constitutively activated and PI3K was completely inactivated. Under this combinatorial perturbation, we found 27 proteins including proto-oncogenes such as Akt (which suppresses apoptosis) and Raf to be downregulated. Here, we found eight proteins with more than 19% increased activity than in the case of a single inactivating perturbation of PI3K. These proteins include Rap1 (+1.19-fold), Ca (+1.21-fold), CaM (+1.21-fold), CaMKK (+1.21-fold), Myosin (+1.22-fold), CaMK (+1.65-fold), PLA2 (+1.98-fold), and AA (+1.98-fold)

(**Table 6**; full list of all affected components is given in Table S5 in Supplementary Material). These components were downregulated when only PI3K was inactivated. Under the combinatorial perturbation (PI3K inactivated and IP3R1 activated), the increased activity of these components was achieved by constitutive expression of IP3R1 *via* the following routes: IP3R1 → Ca → CaM → CaMK → Rap1 and IP3R1 → Ca → CaM → CaMK → PLA2 → AA (**Figure 9**). It is noteworthy that these components have been found to positively regulate apoptosis or cell death. Therefore, under the aforementioned combinatorial perturbation, components involved in cell proliferation were downregulated through the inactivation of PI3K, and the activity of tumor-suppressor genes (PLA2) with arachidonic acid (AA) and other components, including Ca, CaM, and CaMK, was increased as a result of the IP3R1 overactivation.

Together, these results suggest a regulatory interaction between PI3K and IP3R1, and that cotargeting both of these components may serve as therapeutic strategy rather than targeting PI3K alone. Using this combination of targets, we simulated cell death behavior in cell proliferation inducing environmental condition. Thus, we predict that this novel target combination might increase the rate of apoptosis while blocking cell proliferation in tumor cells. However, additional experimental validation is needed to validate this computational result.

## DISCUSSION

We have presented a systemic perturbation analysis of a signal transduction network model to identify and characterize functionally important components. We used these components to explore novel therapeutic strategies against cancer. Specifically, we used a logical modeling approach to analyze the dynamics of a large-scale signal transduction model. Logical modeling approaches have been used, for example, to understand the dynamics of signal transduction and gene regulation networks to identify drug synergies in gastric cancers, and to identify potential drug combinations (Flobak et al., 2015). In biochemical networks, combined effect of topology and dynamical features has been shown to have the most significant impact on the dynamics of the network (Kochi et al., 2014). Computational approaches have become indispensable tools to understand biological pathways and disease phenotypes. Examples include computational methods such as molecular modeling, text mining, and network modeling to identify drug targets in a vast array of diseases from pathogens to complex disorders (Flórez et al., 2010; Yao et al., 2010; Folger et al., 2011; Madrahimov et al., 2013; Puniya et al., 2013).

In the present work, the identified most influential components were characterized for biological functions. The relevance of identified influential components was established with pathway analysis, mapping of housekeeping genes, essential proteins, and association with druggable proteins. Interestingly, we found enrichment of housekeeping genes in the most influential components that were independent of the extracellular environments. A notable agreement is obtained from literature surveys for the most influential components, which were unique to specific

environmental conditions. Because essential components are important from a disease perspective, the identified most influential components may serve as potential candidates and essential proteins under specific conditions. Under activating perturbations, we found that essential genes were enriched more within the most influential components than within the least influential components. The high association of dysregulated signal transduction proteins with different subtypes of cancers suggests that these components may be important candidates for drug targets. Notably, the most influential components are enriched with several already known drug targets. However, many of these drug targets (EGFR, ERK, Ras, PI3K, etc.) have been associated with drug resistance (West et al., 2002; Kobayashi et al., 2005; Linardou et al., 2008; Wheeler et al., 2010; Dienstmann et al., 2012). The mechanism of drug resistance includes mutation in the targeted protein or expression of other genes (altered expression) to bypass the effect caused by perturbation, deregulation in apoptosis, etc. (Gottesman, 2002; Holohan et al., 2013). Thus, to identify novel regulatory interactions, we explored components that are upstream to the most influential components associated with drug resistance. Interestingly, several upstream components (more than 90% in the case of inactivating perturbations) to the most influential components were also identified as most influential. Thus, the most influential components form a tightly connected sub-network of proteins interacting with each other. In yeast, it has previously shown that the essential proteins are hubs in the network and have more interconnections than non-essential proteins, and form a module or sub-network (Song and Singh, 2013).

The interaction between IP3R1 and PI3K was observed under all environmental conditions. This interaction was also observed as part of network motifs in the modeled signal transduction network. IP3R1 activation, when combined with PI3K inactivation, increases the activities of PLA2 and AA, which are decreased with a single PI3K knockdown. It was already shown that AA released by PLA2 helps to initiate apoptosis (Penzo et al., 2004). In a *Dictyostelium discoideum* chemotaxis experiment, it was also shown that cells with PI3K deficiency were more sensitive to PLA2 inhibition (Chen et al., 2007), which supports our predicted interaction between PI3K and PLA2. To this end, we hypothesized that the PI3K inactivation could be combined with the overactivation of IP3R1 to increase the activity of proteins involved in apoptosis. IP3R1 inactivation can lead to the downregulation of RGS, and reversibly, the overexpression of IP3R1 can lead to increased activity of RGS. Similar to IP3R1, RGS subtype RGS3T has been found to be involved in inducing cell death (Fisher, 2009), and it has also been found that RGS can suppress the PI3K activity downstream of the receptor (Liang et al., 2009). Therefore, the constitutive activation of IP3R1 might also negatively regulate the activity of PI3K. Systemic analysis of the most influential components and their upstream components has led us to identify novel combinations of drug targets. In various studies, combinatorial therapies have shown a decrease in drug resistance in pathogens. In combinatorial therapy, a protein associated with drug resistance can be targeted in combination with different protein of either the same or different pathway (Fischbach, 2011). Clinical trials have also suggested that the efficiency of cytotoxic drugs



increases when given in combinations (Al-Lazikani et al., 2012). If co-occurrence of two genetic events results in cell death, it can be termed as synthetic lethality (Nijman, 2011). The combinatorial perturbation of PI3K and IP3R1 could be considered as synthetically lethal. However, in this perturbation, the activation of IP3R1 is synergistic with the inactivation of PI3K. Upregulation of IP3R1 could be achieved using a targeted drug therapy, such as stress hormone dexamethasone, a synthetic glucocorticoid show to significantly upregulate the expression of IP3R1 in differentiating myoblasts (Chai et al., 2010).

As a validation of model's result, we used previously published gene expression data. Our model's results moderately correlate with this data. This agreement was based on only one dataset of PI3K inhibition with two biological replicates. Further addition of experimental data for other perturbations, including the combinatorial perturbation is required to validate the trends of perturbation analysis in model.

In conclusion, by combining IP3R1 (activation) and PI3K (inactivation), we were able to stimulate cell death under the cell growth stimulating condition. Based on this, one can hypothesize that it might be possible that the decrease in cell proliferation with increased apoptosis as a result of this combinatorial intervention could subsequently increase the rate of clearance of tumor cells, and serve as a novel strategy for important targets associated with drug resistance. However, more laboratory validations will be required to test this hypothesis.

## AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: BP, LA, CH, and TH. Performed experiments: LA and MM. Data analysis and interpretation: BP and CH. Wrote the manuscript: BP, LA, CH, and TH.

## REFERENCES

- Al-Lazikani, B., Banerji, U., and Workman, P. (2012). Combinatorial drug therapy for cancer in the post-genomic era. *Nat. Biotechnol.* 30, 679–692. doi:10.1038/nbt.2284
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi:10.1038/75556
- Barrett, T., Suzek, T. O., Troup, D. B., Wilhite, S. E., Ngau, W.-C., Ledoux, P., et al. (2005). NCBI GEO: mining millions of expression profiles – database and tools. *Nucleic Acids Res.* 33, D562–D566. doi:10.1093/nar/gki022
- Bastian, F., Parmentier, G., Roux, J., Moretti, S., Laudet, V., and Robinson-Rechavi, M. (2008). “Bgee: integrating and comparing heterogeneous transcriptome data among species,” in *Data Integration in the Life Sciences*, eds A. Bairoch, S. Cohen-Boulakia, and C. Froidevaux (Berlin; Heidelberg: Springer), 124–131.
- Berchtold, M. W., and Villalobo, A. (2014). The many faces of calmodulin in cell proliferation, programmed cell death, autophagy, and cancer. *Biochim. Biophys. Acta* 1843, 398–435. doi:10.1016/j.bbamcr.2013.10.021
- Birkeland, E., Wik, E., Mjøs, S., Hoivik, E., Trovik, J., Werner, H. M. J., et al. (2012). KRAS gene amplification and overexpression but not mutation associates with aggressive and metastatic endometrial cancer. *Br. J. Cancer* 107, 1997–2004. doi:10.1038/bjc.2012.477
- Blazier, A. S., and Papin, J. A. (2012). Integration of expression data in genome-scale metabolic network reconstructions. *Front. Physiol.* 3:299. doi:10.3389/fphys.2012.00299

## ACKNOWLEDGMENTS

We would like to thank Resa Helikar for providing feedback on the manuscript.

## FUNDING

Office of Sponsored Programs at the University of Nebraska at Omaha via Fund for Undergraduate Scholarly Experiences to LA and CH.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at <http://journal.frontiersin.org/article/10.3389/fbioe.2016.00010>

**TABLE S1 | Functional activity of signal transduction components in model.** Gene expression of signal transduction components were queried in four databases: Bgee, CleanEx, Expression Atlas, and Genevisible.

**TABLE S2 | Influential ranking of all the components of signal transduction model.**

**TABLE S3 | Network motifs that containing component pairs selected in Table 5.**

**TABLE S4 | Significant Z-ratios (cells treated with PI3K-inhibitor/untreated) of genes of signal transduction components showed activity level changes in PI3K inactivation in model.**

**TABLE S5 | Activity levels of all the components; on PI3K inactivation (single perturbation), on IP3R1 activation (single perturbation), and on PI3K-IP3R1 (double perturbations).**

**DATA SHEET 1 | Model file in SBML format.**

**FIGURES S1–S10 | Heatmaps of raw differences of activity levels between perturbed and WT conditions (for all the simulated environmental conditions under both types of perturbations).**

- Chai, J., Xiong, Q., Zhang, P., Zheng, R., Peng, J., and Jiang, S. (2010). Induction of Ca<sup>2+</sup> signal mediated apoptosis and alteration of IP3R1 and SERCA1 expression levels by stress hormone in differentiating C2C12 myoblasts. *Gen. Comp. Endocrinol.* 166, 241–249. doi:10.1016/j.ygcen.2009.08.011
- Cheadle, C., Vawter, M. P., Freed, W. J., and Becker, K. G. (2003). Analysis of microarray data using Z score transformation. *J. Mol. Diagn.* 5, 73–81. doi:10.1016/S1525-1578(10)60455-2
- Chen, L., Iijima, M., Tang, M., Landree, M. A., Huang, Y. E., Xiong, Y., et al. (2007). PLA 2 and PI3K/PTEN pathways act in parallel to mediate chemotaxis. *Dev. Cell* 12, 603–614. doi:10.1016/j.devcel.2007.03.005
- Chen, W.-H., Minguez, P., Lercher, M. J., and Bork, P. (2012). OGEE: an Online GEne Essentiality database. *Nucleic Acids Res.* 40, D901–D906. doi:10.1093/nar/gkr986
- Conroy, B. D., Herek, T. A., Shew, T. D., Latner, M., Larson, J. J., Allen, L., et al. (2014). Design, assessment, and in vivo evaluation of a computational model illustrating the role of CAV1 in CD4+ T-lymphocytes. *Front. Immunol.* 5:599. doi:10.3389/fimmu.2014.00599
- Consortium, U. (2011). Reorganizing the protein space at the universal protein resource (UniProt). *Nucleic Acids Res.* 40, D71–D75. doi:10.1093/nar/gkr981
- Dienstmann, R., De Dosso, S., Felip, E., and Tabernero, J. (2012). Drug development to overcome resistance to EGFR inhibitors in lung and colorectal cancer. *Mol. Oncol.* 6, 15–26. doi:10.1016/j.molonc.2011.11.009
- Eisenberg, E., and Levanon, E. Y. (2013). Human housekeeping genes, revisited. *Trends Genet.* 29, 569–574. doi:10.1016/j.tig.2013.05.010
- Fischbach, M. A. (2011). Combination therapies for combating antimicrobial resistance. *Curr. Opin. Microbiol.* 14, 519–523. doi:10.1016/j.mib.2011.08.003

- Fisher, R. A. (2009). *Molecular Biology of RGS Proteins*. London: Academic Press.
- Flobak, Å., Baudot, A., Remy, E., Thommesen, L., Thieffry, D., Kuiper, M., et al. (2015). Discovery of drug synergies in gastric cancer cells predicted by logical modeling. *PLoS Comput. Biol.* 11:e1004426. doi:10.1371/journal.pcbi.1004426
- Flórez, A. F., Park, D., Bhak, J., Kim, B.-C., Kuchinsky, A., Morris, J. H., et al. (2010). Protein network prediction and topological analysis in *Leishmania major* as a tool for drug target selection. *BMC Bioinformatics* 11:484. doi:10.1186/1471-2105-11-484
- Folger, O., Jerby, L., Frezza, C., Gottlieb, E., Rupp, E., and Shlomi, T. (2011). Predicting selective drug targets in cancer through metabolic networks. *Mol. Syst. Biol.* 7, 501. doi:10.1038/msb.2011.35
- Ghosh, S., Matsuoka, Y., Asai, Y., Hsin, K.-Y., and Kitano, H. (2011). Software for systems biology: from tools to integrated platforms. *Nat. Rev. Genet.* 12, 821–832. doi:10.1038/nrg3096
- Gottesman, M. M. (2002). Mechanisms of cancer drug resistance. *Annu. Rev. Med.* 53, 615–627. doi:10.1146/annurev.med.53.082901.103929
- Helikar, T., Konvalina, J., Heidel, J., and Rogers, J. A. (2008). Emergent decision-making in biological signal transduction networks. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1913–1918. doi:10.1073/pnas.0705088105
- Helikar, T., Kowal, B., McClenathan, S., Bruckner, M., Rowley, T., Madrahimov, A., et al. (2012). The cell collective: toward an open and collaborative approach to systems biology. *BMC Syst. Biol.* 6:96. doi:10.1186/1752-0509-6-96
- Helikar, T., Kowal, B., and Rogers, J. A. (2013). A cell simulator platform: the cell collective. *Clin. Pharmacol. Ther.* 93, 393–395. doi:10.1038/clpt.2013.41
- Helikar, T., and Rogers, J. A. (2009). ChemChains: a platform for simulation and analysis of biochemical networks aimed to laboratory scientists. *BMC Syst. Biol.* 3:58. doi:10.1186/1752-0509-3-58
- Holohan, C., Van Schaeybroeck, S., Longley, D. B., and Johnston, P. G. (2013). Cancer drug resistance: an evolving paradigm. *Nat. Rev. Cancer* 13, 714–726. doi:10.1038/nrc3599
- Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi:10.1038/nprot.2008.211
- Janssens, V., and Goris, J. (2001). Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem. J.* 353, 417–439. doi:10.1042/bj3530417
- Kanehisa, M. (2002). “The KEGG database,” in *In Silico Simulation of Biological Processes: Novartis Foundation Symposium*, Vol. 247. eds G. Bock and J. A. Goode (Chichester: John Wiley & Sons Ltd.), 91–103.
- Kayl, A. E., and Meyers, C. A. (2006). Side-effects of chemotherapy and quality of life in ovarian and breast cancer patients. *Curr. Opin. Obstet. Gynecol.* 18, 24–28. doi:10.1097/01.gco.0000192996.20040.24
- Kitano, H. (2002a). Computational systems biology. *Nature* 420, 206–210. doi:10.1038/nature01254
- Kitano, H. (2002b). Systems biology: a brief overview. *Science* 295, 1662–1664. doi:10.1126/science.1069492
- Kobayashi, S., Boggon, T. J., Dayaram, T., Janne, P. A., Kocher, O., Meyerson, M., et al. (2005). EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* 352, 786–792. doi:10.1056/NEJMoa044238
- Kochi, N., Helikar, T., Allen, L., Rogers, J. A., Wang, Z., and Matache, M. T. (2014). Sensitivity analysis of biological Boolean networks using information fusion based on nonadditive set functions. *BMC Syst. Biol.* 8:92. doi:10.1186/s12918-014-0092-4
- Le Novère, N. (2015). Quantitative and logic modelling of molecular and gene networks. *Nat. Rev. Genet.* 16, 146–158. doi:10.1038/nrg3885
- Liang, G., Bansal, G., Xie, Z., and Druey, K. M. (2009). RGS16 inhibits breast cancer cell growth by mitigating phosphatidylinositol 3-kinase signaling. *J. Biol. Chem.* 284, 21719–21727. doi:10.1074/jbc.M109.028407
- Linaridou, H., Dahabreh, I. J., Kanalioupi, D., Siannis, F., Bafaloukos, D., Kosmidis, P., et al. (2008). Assessment of somatic k-RAS mutations as a mechanism associated with resistance to EGFR-targeted agents: a systematic review and meta-analysis of studies in advanced non-small-cell lung cancer and metastatic colorectal cancer. *Lancet Oncol.* 9, 962–972. doi:10.1016/S1470-2045(08)70206-7
- Lito, P., Rosen, N., and Solit, D. B. (2013). Tumor adaptation and resistance to RAF inhibitors. *Nat. Med.* 19, 1401–1409. doi:10.1038/nm.3392
- Loscalzo, J., and Barabasi, A. L. (2011). Systems biology and the future of medicine. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 3, 619–627. doi:10.1002/wsbm.144
- Lotfi-Jam, K., Carey, M., Jefford, M., Schofield, P., Charleson, C., and Aranda, S. (2008). Nonpharmacologic strategies for managing common chemotherapy adverse effects: a systematic review. *J. Clin. Oncol.* 26, 5618–5629. doi:10.1200/JCO.2007.15.9053
- Madrahimov, A., Helikar, T., Kowal, B., Lu, G., and Rogers, J. (2013). Dynamics of influenza virus and human host interactions during infection and replication cycle. *Bull. Math. Biol.* 75, 988–1011. doi:10.1007/s11538-012-9777-2
- Mason, R. P. (1999). Effects of calcium channel blockers on cellular apoptosis. *Cancer* 85, 2093–2102. doi:10.1002(SICI)1097-0142(19990515)85:10<2093::AID-CNCR1>3.0.CO;2-E
- Mendelsohn, J. (2001). The epidermal growth factor receptor as a target for cancer therapy. *Endocr. Relat. Cancer* 8, 3–9. doi:10.1677/erc.0.0080003
- Molinelli, E. J., Korkut, A., Wang, W., Miller, M. L., Gauthier, N. P., Jing, X., et al. (2013). Perturbation biology: inferring signaling networks in cellular systems. *PLoS Comput. Biol.* 9:e1003290. doi:10.1371/journal.pcbi.1003290
- Naldi, A., Carneiro, J., Chauviya, C., and Thieffry, D. (2010). Diversity and plasticity of Th cell types predicted from regulatory network modelling. *PLoS Comput. Biol.* 6:e1000912. doi:10.1371/journal.pcbi.1000912
- Nijman, S. M. (2011). Synthetic lethality: general principles, utility and detection using genetic screens in human cells. *FEBS Lett.* 585, 1–6. doi:10.1016/j.febslet.2010.11.024
- Penzo, D., Petronilli, V., Angelin, A., Cusan, C., Colonna, R., Scorrano, L., et al. (2004). Arachidonic acid released by phospholipase A2 activation triggers Ca<sup>2+</sup>-dependent apoptosis through the mitochondrial pathway. *J. Biol. Chem.* 279, 25219–25225. doi:10.1074/jbc.M310381200
- Perna, D., Karreth, F. A., Rust, A. G., Perez-Mancera, P. A., Rashid, M., Iorio, F., et al. (2015). BRAF inhibitor resistance mediated by the AKT pathway in an oncogenic BRAF mouse melanoma model. *Proc. Natl. Acad. Sci. U.S.A.* 112, E536–E545. doi:10.1073/pnas.1418163112
- Petryszak, R., Burdett, T., Fiorelli, B., Fonseca, N. A., Gonzalez-Porta, M., Hastings, E., et al. (2014). Expression Atlas update – a database of gene and transcript expression from microarray-and sequencing-based functional genomics experiments. *Nucleic Acids Res.* 42, D926–D932. doi:10.1093/nar/gkt1270
- Pinton, P., Giorgi, C., Siviero, R., Zecchini, E., and Rizzuto, R. (2008). Calcium and apoptosis: ER-mitochondria Ca<sup>2+</sup> transfer in the control of apoptosis. *Oncogene* 27, 6407–6418. doi:10.1038/ncr.2008.308
- Praz, V., Jagannathan, V., and Bucher, P. (2004). CleanEx: a database of heterogeneous gene expression data based on a consistent gene nomenclature. *Nucleic Acids Res.* 32, D542–D547. doi:10.1093/nar/gkh107
- Pruitt, K. D., Tatusova, T., and Maglott, D. R. (2007). NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 35, D61–D65. doi:10.1093/nar/gkl842
- Puniya, B. L., Kulshreshtha, D., Verma, S. P., Kumar, S., and Ramachandran, S. (2013). Integrated gene co-expression network analysis in the growth phase of *Mycobacterium tuberculosis* reveals new potential drug targets. *Mol. Biosyst.* 9, 2798–2815. doi:10.1039/c3mb70278b
- Qu, Y., Chen, Q., Lai, X., Zhu, C., Chen, C., Zhao, X., et al. (2014). SUMOylation of Grb2 enhances the ERK activity by increasing its binding with Sos1. *Mol. Cancer* 13, 95. doi:10.1186/1476-4598-13-95
- Rachagani, S., Senapati, S., Chakraborty, S., Ponnusamy, M. P., Kumar, S., Smith, L., et al. (2011). Activated KrasG12D is associated with invasion and metastasis of pancreatic cancer cells through inhibition of E-cadherin. *Br. J. Cancer* 104, 1038–1048. doi:10.1038/bjc.2011.31
- Rocha, C., Mendonça, T., and Eduarda Silva, M. (2013). Modelling neuromuscular blockade: a stochastic approach based on clinical data. *Math. Comput. Model. Dyn. Syst.* 19, 540–556. doi:10.1080/13873954.2013.801865
- Rodon, J., Dienstmann, R., Serra, V., and Tabernero, J. (2013). Development of PI3K inhibitors: lessons learned from early clinical trials. *Nat. Rev. Clin. Oncol.* 10, 143–153. doi:10.1038/nrclinonc.2013.10
- Rosich, L., Montraveta, A., Xargay-Torrent, S., López-Guerra, M., Roldán, J., Aymeric, M., et al. (2014). Dual PI3K/mTOR inhibition is required to effectively impair microenvironment survival signals in mantle cell lymphoma. *Oncotarget* 5, 6788. doi:10.18632/oncotarget.2253
- Singh, P., and Singh, A. (2012). Ocular adverse effects of anti-cancer chemotherapy. *J. Cancer Ther. Res.* 1, 5. doi:10.7243/2049-7962-1-5
- Song, J., and Singh, M. (2013). From hub proteins to hub modules: the relationship between essentiality and centrality in the yeast interactome at different scales of organization. *PLoS Comput. Biol.* 9:e1002910. doi:10.1371/journal.pcbi.1002910

- Steinway, S. N., Zañudo, J. G., Ding, W., Rountree, C. B., Feith, D. J., Loughran, T. P., et al. (2014). Network modeling of TGF $\beta$  signaling in hepatocellular carcinoma epithelial-to-mesenchymal transition reveals joint Sonic Hedgehog and Wnt pathway activation. *Cancer Res.* 74, 5963–5977. doi:10.1158/0008-5472.CAN-14-0225
- Todd, R. G., and Helikar, T. (2012). Ergodic sets as cell phenotype of budding yeast cell cycle. *PLoS ONE* 7:e45780. doi:10.1371/journal.pone.0045780
- Vanneman, M., and Dranoff, G. (2012). Combining immunotherapy and targeted therapies in cancer treatment. *Nat. Rev. Cancer* 12, 237–251. doi:10.1038/nrc3237
- Wagle, N., Van Allen, E. M., Treacy, D. J., Frederick, D. T., Cooper, Z. A., Taylor-Weiner, A., et al. (2014). MAP kinase pathway alterations in BRAF-mutant melanoma patients with acquired resistance to combined RAF/MEK inhibition. *Cancer Discov.* 4, 61–68. doi:10.1158/2159-8290.CD-13-0631
- Wang, J., Tsang, W. W., and Marsaglia, G. (2003). Evaluating Kolmogorov's distribution. *J. Stat. Softw.* 8, doi:10.18637/jss.v008.i18
- Washburn, M. P., Koller, A., Oshiro, G., Ulaszek, R. R., Plouffe, D., Deciu, C., et al. (2003). Protein pathway and complex clustering of correlated mRNA and protein expression analyses in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 3107–3112. doi:10.1073/pnas.0634629100
- Wazir, U., Jiang, W. G., Sharma, A. K., and Mokbel, K. (2013). Guanine nucleotide binding protein  $\beta$  1: a novel transduction protein with a possible role in human breast cancer. *Cancer Genomics Proteomics* 10, 69–73. doi:10.1016/j.ejso.2012.07.172
- Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., Ellrott, K., et al. (2013). The Cancer Genome Atlas pan-cancer analysis project. *Nat. Genet.* 45, 1113–1120. doi:10.1038/ng.2764
- Wernicke, S., and Rasche, F. (2006). FANMOD: a tool for fast network motif detection. *Bioinformatics* 22, 1152–1153. doi:10.1093/bioinformatics/btl038
- West, K. A., Castillo, S. S., and Dennis, P. A. (2002). Activation of the PI3K/Akt pathway and chemotherapeutic resistance. *Drug. Resist. Updat.* 5, 234–248. doi:10.1016/S1368-7646(02)00120-6
- Wheeler, D. L., Dunn, E. F., and Harari, P. M. (2010). Understanding resistance to EGFR inhibitors – impact on future treatment strategies. *Nat. Rev. Clin. Oncol.* 7, 493–507. doi:10.1038/nrclinonc.2010.97
- Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., Stothard, P., et al. (2006). DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* 34, D668–D672. doi:10.1093/nar/gkj067
- Yao, L., Evans, J. A., and Rzhetsky, A. (2010). Novel opportunities for computational biology and sociology in drug discovery: corrected paper. *Trends Biotechnol.* 28, 161–170. doi:10.1016/j.tibtech.2010.01.004
- Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., and Gruissem, W. (2004). GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol.* 136, 2621–2632. doi:10.1104/pp.104.046367

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Tomáš Helikar has served as a scientific advisor and/or consultant to Discovery Collective.

Copyright © 2016 Puniya, Allen, Hochfelder, Majumder and Helikar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.