HHS Public Access

Author manuscript

IEEE Trans Biomed Circuits Syst. Author manuscript; available in PMC 2015 November 03.

Published in final edited form as:

IEEE Trans Biomed Circuits Syst. 2014 February ; 8(1): 74–86. doi:10.1109/TBCAS.2013.2288035.

Time-Varying Causal Inference From Phosphoproteomic Measurements in Macrophage Cells

Maryam Masnadi-Shirazi,

Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093 USA

Mano Ram Maurya, and

San Diego Supercomputer Center and the Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093 USA

Shankar Subramaniam

Department of Bioengineering, Departments of Chemistry and Biochemistry, Cellular and Molecular Medicine and the Graduate Program in Bioinformatics, University of California, San Diego, La Jolla, CA 92093 USA

Maryam Masnadi-Shirazi: mmasnadi@ucsd.edu; Mano Ram Maurya: mano@sdsc.edu; Shankar Subramaniam: shankar@ucsd.edu

Abstract

Cellular signaling circuitry in eukaryotes can be studied by analyzing the regulation of protein phosphorylation and its impact on downstream mechanisms leading to a pheno-type. A primary role of phosphorylation is to act as a switch to turn “on” or “off” a protein activity or a cellular pathway. Specifically, protein phosphorylation is a major *leit motif* for transducing molecular signals inside the cell. Errors in transferring cellular information can alter the normal function and may lead to diseases such as cancer; an accurate reconstruction of the “true” signaling network is essential for understanding the molecular machinery involved in normal and pathological function. In this study, we have developed a novel framework for time-dependent reconstruction of signaling networks involved in the activation of macrophage cells leading to an inflammatory response. Several signaling pathways have been identified in macrophage cells, but the time-varying causal relationship that can produce a dynamic directed graph of these molecules has not been explored in detail. Here, we use the notion of Granger causality, and apply a vector autoregressive model to phosphoprotein time-course data in RAW 264.7 macrophage cells. Through the reconstruction of the phosphoprotein network, we were able to estimate the directionality and the dynamics of information flow. Significant interactions were selected through statistical hypothesis testing (t-test) of the coefficients of a linear model and were used to reconstruct the phosphoprotein signaling network. Our approach results in a three-stage phosphoprotein network that represents the evolution of the causal interactions in the intracellular signaling pathways.

M. Masnadi-Shirazi, M. R. Maurya (equal effort)

Index Terms

Granger causality; phosphoproteomics; signaling pathway; time-varying network reconstruction; vector autoregressive model

I. Introduction

The understanding of cellular function at the molecular level involves the study of intracellular signaling, metabolic pathways and gene regulatory networks, through “omics” measurements on biological systems. Protein phosphorylation is one of the main steps in intracellular signaling from the activated proteins located at the plasma membrane to the cytosolic space and nucleus. Phosphorylation is one of the most studied post-translational modification of proteins since is it vital for many protein interactions that regulate cellular processes such as cell growth, cell differentiation and development to cell cycle control and metabolism [1]. Phosphorylation is a key reversible modification with the combined involvement of protein kinases and phosphatases to activate and deactivate proteins [2]. Phosphorylation mainly occurs on serine, threonine and tyrosine residues that can regulate enzymatic activity, subcellular localization, complex formation and degradation of proteins. Activation of proteins through phosphorylation serves as the flux in the signaling pathways. Several signaling pathways such as the nuclear factor kappa B (NF-κB), mitogen-activated protein kinases (MAPK), and signal transducer and activator of transcription (STAT) play essential roles in transmitting signals that trigger the release of cytokines, which are central to the processes of inflammation and modulation of immune function [3]. The signaling pathways act as modules to regulate the transcription and release of various cytokines, some of which are involved in the pathogenesis of many diseases, e.g., chronic inflammatory diseases, autoimmunity and cancer. Thus, reconstructing protein networks from “omics” measurements can help us not only understand and model cellular signaling pathways but also assist in uncovering the mechanisms of disease progression. Since knowledge of protein-protein interaction is sparse, it is difficult to simultaneously analyze the dynamics of various proteins *in vitro* or *in vivo*. High-throughput technologies, such as nextgen sequencing, DNA microarray expression profiling, phosphoproteomics, metabolomics and high-content imaging, have made it possible to make concurrent quantitative measurements of various components of the cell, including mRNA levels, protein phosphorylation and metabolites, enabling the reconstruction of large-scale cellular networks. Complexities such as feedback and feed-forward loops and the cross-talk between different signaling pathways have hindered the problem of developing reliable mathematical approaches within an integrative framework, taking into account the dynamics of signaling networks [3].

During the last decade, the application of mathematical and statistical approaches to high-throughput biological data has been used extensively to decipher the relationship between different components in the cell to partially reconstruct intra-cellular networks. With the availability of large-scale omics data, computational systems biology has made substantial progress towards modeling and reconstruction of data-driven networks using (1) input/output-based models such as Partial Least Squares (PLS) [4] and Principal Component Regression (PCR) [5], (2) probabilistic graphical models such as Bayesian network-based

models [6]–[8], probabilistic Boolean network models [9], [10], and (3) information theory-based methods such as integrated correlation and transfer entropy based approach [11] and C3NET [12], [13]. Other approaches using differential equations [14], structural equation methods [15] and state-space models [16] have also been proposed during the past few years.

Biological systems evolve through time and it is important to study the dynamic behavior of the topology of the signaling pathways/networks themselves [16]. Thus, we allow the network topology (the set of connections/edges present in the network) to evolve with time. Our objective in this study is to derive a time-varying model for the phosphoprotein network to understand the dynamics of signaling pathways using the notion of Granger causality. Causality can be determined by prior biological information. However, in many cases, no *"a priori"* knowledge is available to provide causal relationships in network reconstruction. Furthermore, it is appealing to discover new causal relationships, rather than already known ones. In the present work, we have applied the notion of Granger causality and statistical hypothesis testing to estimate causal relationships between different phosphoproteins using time-series data. According to Granger's definition of causality, it is said that signal X(t) causes signal Y(t), if future values of Y(t) can be better predicted using the past values of X(t) and Y(t) than only using the past of itself [17].

Due to the fact that intracellular networks are not static, we use time series data in order to determine these dynamic changes in the network topology. In the present work, we use a vector autoregressive (VAR) model to infer relationships of Granger causality among phosphoproteins by analyzing the time-varying fold changes of phosphoproteins in response to single and double ligand stimuli. The quantitative levels of phosphoproteins were measured through western blot experiments by the Alliance for cellular Signaling (AfCS) [18] in RAW 264.7 macrophage cells. We infer the topology of the phosphoprotein networks in three distinct time intervals.

II. Approach

A. Granger Causality

Granger causality was first introduced by the Noble prize-winning economist, Clive Granger, and has proven useful for analyzing the relationships and influences among macroeconomic time series (e.g., income, exchange rate, etc.) [17]. We note that Granger causality is not meant to be equivalent to the true causality, but is intended to provide useful information regarding causation and the direction of information flow. Formally, a time series x_t is said to Granger-cause a time series y_t, if the future value of y_t can be predicted given the past values of y_t and x_t, (y_{t-1}, y_{t-2}, …, x_{t-1}, x_{t-2}, …), better than predicting the future of y_t given only the past values of y_t (y_{t-1}, y_{t-2}, …). Commonly, Granger causality is identified by VAR models [19]. A VAR model of p-order and k-dimensional time series is given by

$$y_t = v + A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} + \cdots + A_p y_{t-p} + \varepsilon_t \quad (1)$$

where $y_t = (y_{1t}, y_{2t}, \ldots, y_{kt})'$ is a $(k \times 1)$ random vector, y_{it} is the measurement at time t of the i^{th} random variable, A_l is a $(k \times k)$ autoregressive coefficient matrix, v is a $(k \times 1)$ vector of intercepts and $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{kt})'$ is a k-dimensional error vector of random variables with zero mean and covariance matrix Σ.

The optimal order of the VAR model can be found through approaches such as Minimum Description Length [20] which requires many samples in time. In the present work, since there are only three original samples in time, we consider the following first order VAR model:

$$y_t = v + A_1 y_{t-1} + \varepsilon_t. \quad (2)$$

VAR allows identification of Granger causality for linear relationships. In order to find causal relationships, we analyze the elements of matrix A_1. An important outcome of this approach is that the series y_{jt} is not the cause of y_{it} if and only if the ij^{th} entry of matrix A_1 is zero, $a_{ij} = 0$. Therefore, it is sufficient to estimate the autoregressive coefficient matrix of the VAR model in order to identify the direction of Granger causality.

This approach can be applied to the analysis of phosphoprotein time-course data to interpret functional connectivity between phosphoproteins to reconstruct their underlying network by testing the statistical significance of the estimated components of A_1. Considering the time series $(y_1, \ldots, y_T)$ for each of the k variables, the first-order VAR model in (2) can be written in the following matrix form [21]:

$$Y = XB + \varepsilon \quad (3)$$

where $Y = (y_1, \ldots, y_T)'$ is a $(T \times k)$ matrix whose columns are time series for each of the k random variables with sample size T, $B = (v, A_1)'$ is a $((k + 1) \times k)$ matrix, $X = (X_0, \ldots, X_{T-1})'$ is a $(T \times (k + 1))$ matrix with $X_t = (1; y_t)$, and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)'$ is a $(T \times k)$ matrix. For each of the k columns of matrices Y, B, and ε, we can write the following linear regression model:

$$Y_i = XB_i + \varepsilon_i; i = 1, \ldots, k \quad (4)$$

where vector Y_i represents the i^{th} column of matrix Y, vector B_i is the i^{th} column of matrix B and vector ε_i is the i^{th} column of matrix ε. In this linear model, we seek to estimate the unknown coefficients in matrix B. We can use least squares (LS) estimation method in order to compute the unknown parameters/coefficients. Therefore, each column of matrix B is estimated through the LS estimation shown below

$$\hat{B}_i = (X'X)^{-1} X' Y_i; i = 1, \ldots, k. \quad (5)$$

After estimating the coefficient vectors for each of the outputs, they can be concatenated to construct the estimated matrix $\hat{B}$, and therefore, the autoregressive coefficient matrix A_1 can be computed. The proposed VAR model analyzes causality between different variables in terms of how the future of a variable can be predicted using the past values of itself and other variables. According to this model, as stated earlier, variable j is said to Granger-cause

variable *i*, if the ij^{th} entry of matrix A_1 is nonzero. However, the least squares criteria favors solutions with many nonzero entries, which is contrary to the goal of finding purely zero entries to identify whether or not causations between pairs of variables exist. Hence, we need to apply statistical significance test to examine the significance of the estimated parameters. We know that LS estimation minimizes the root mean squared error (RMSE), and by computing the RMSE, we can perform a two-tailed t-test on the coefficients. The RMSE is computed as follows:

$$RMSE_{LS}=\sqrt{\frac{1}{T}\sum_{i=1}^{T}(Y_i-\hat{Y}_i)^2}=std(Y-\hat{Y})\times\sqrt{\frac{T-1}{T}} \quad (6)$$

where $\hat{Y}_i$ is the estimation of Y_i

$$\hat{Y}_i=X\hat{B}_i. \quad (7)$$

1) Significant Connections—The standard-deviation of the model parameters are estimated as

$$\sigma_{b,LS}\approx diag\left((X^TX)^{-1}\right)^{1/2}\times RMSE_{LS}\times\left(\frac{T}{\upsilon}\right)^{1/2};\upsilon=T-k-1 \quad (8)$$

where *T* is the length of the time series, *k* is the number of variables, and υ is defined as the degrees of freedom. Then the ratio $r_{ji}=\hat{B}_{ji}/\sigma_{b,LS}$ is computed for the j^{th} entry of the i^{th} column of the estimated matrix $\hat{B}$ and $|r_{ji}|$ is compared against $R = tinv(1-\alpha/2, \upsilon)$, where *tinv*(.) denotes the inverse of the cumulative t-distribution and $\alpha = 0.01$ (two-tailed) for a confidence interval of 99%. The estimated coefficients are considered statistically significant if their corresponding ratios are greater than *R* and insignificant otherwise (t-test on the model coefficients). We also computed the p-value and false-discovery rate (FDR) using the Benjamini-Hochberg (BH) method [22] for the connections retained. As presented in the Results section, the Benjamini-Hochberg FDR for the connections retained is less than 0.026.

2) Performance Metrics—Type I error, Type II error, and accuracy of the network is computed [23] as follows using the False Positives (FP), False Negatives (FN), True Positives (TP) and True Negatives (TN) in the network identified:

$$Type\ I\ Error=\frac{FP}{FP+TN} \quad (9)$$

$$Type\ II\ Error=\frac{FN}{FN+TP} \quad (10)$$

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN}. \quad (11)$$

B. Application of VAR Model to Phosphoproteomic Data

We applied this method to time-course data on the level of phosphorylation of proteins in RAW 264.7 macrophages in response to stimuli, provided by the Alliance for Cellular Signaling (AfCS) [18]. This data set consists of fold changes of 21 phosphoproteins at 4 time points; i.e., data at 1,3,10 and 30 minutes, in response to treatments with 22 single ligands and their double ligand combinations measured using the western blot method. The fold changes of the phosphoproteins are determined by dividing the volume of each phosphoprotein band for the ligand-treated samples by the average volume of the corresponding bands for the untreated samples (volume is the sum of the image pixel values within the area of the band). The replicates for the experiments with unique combination of ligand(s) for each phosphoprotein were averaged. Out of 327 unique ligand combinations, the number of combinations with 1, 2, 3, 4 and more than 4 replicates was 68, 68, 123, 37 and 31, respectively. Thus, most ligand combinations have three replicates, hence resulting in only a small bias due to the difference in the number of replicates.

Due to the fact that the time intervals are not equal, we interpolated the data using linear interpolation with steps of one minute. Other interpolation methods (e.g., cubic) may result in large deviations at the intermediate time points, and this may not be close to the real variation of the fold change of the phosphoproteins in the biological system. We excluded the last sample in the original data, since it was taken 20 minutes after the previous one, which is considered to be too large an interval for accurate interpolation. In these experiments, we had missing data for 4 of the 21 phosphoproteins, signal transducer and activator of transcription (STAT) 3, STAT5, c-Jun N-terminal kinases (JNK) long (JNKL) and JNK short (JNKS). Therefore, we excluded these variables from further analysis. We assumed that at a given time, the underlying phosphoprotein network that represents the structure or the topology of the biological system is the same across all experiments, i.e., the topology of the phosphoprotein network representing the behavior of the biological system remains unchanged regardless of which ligand(s) is stimulating the system. Thus, to deal with the problem of rank deficiency of matrix X in (4), we stacked the data from multiple experiments for both the output data in matrix Y (data related to present) and the input data in matrix X (data related to the past). This ensures that matrix X will have full column rank and there will be a unique solution to the least squares problem. Fig. 1 shows a schematic of how the input and output data from multiple experiments were stacked. Before implementing the VAR model, the data in matrix X was normalized and matrix Y was mean-centered for each variable.

In addition to implementing the VAR model, the correlation between the past and present values for each pair of variables was studied and the correlation matrix between the input and output variables was computed. Fig. 2 visualizes the correlation matrix as a heat-map, where the rows and columns of the heat-map are the input (at time $t - 1$) and the output variables (at time t) for the whole time-series data, respectively.

In order to investigate how the underlying topology of the network is changing, we partition the time series for all the variables into three segments and then apply the VAR model for each segment separately. Since we are considering the time-course data for 1 to 10 minutes,

and the granularity of the measurements is not fine, three overlapping segments, [1]–[4], [3]–[7] and [6]–[10] minutes were considered using interpolated data. Next, in order to investigate how the causal relationships are evolving with time, we estimate the causality coefficients and perform a statistical significance test (t-test) for each segment separately. We also compute the correlation matrix for each segment independently. It is expected that the results based on the interpolated data in the [3]–[7] minute interval are more affected by the actual experimental value at 3 minute, whereas those based on [6]–[10] minute interval are more affected by the actual experimental value at 10 minute. Among the statistically significant causal relationships that were estimated through the VAR model, only those with high correlation coefficients (≥ 0.4; p-value is quite significant since the number of rows in the matrices *X* and *Y*, 2943, is very large) were selected to reconstruct the final network for each time interval. Therefore, the network identified contains likely causal connections which also exhibit high correlation.

It can be noted that since we are considering three separate time intervals to study the temporal evolution of the network, we expect that the information provided in the time series data may differ from stage to stage. Therefore, a causal relationship A → B that exists at an earlier stage need not exist at the following stage, i.e., the past value of A may no longer contribute to predicting the future value of B at the following stage. Thus, according to Granger's definition of causality, there will be no causal relationship at the following stage. This implies that the weights of edges (resulting in fluxes through connections) change though time. For example, if the weight of a connection decreases and the corresponding p-value becomes more than the threshold of 0.01 (for a confidence interval of 99%), we no longer consider that connection to exist as a strong causal relationship even though we may observe the connection in the underlying network.

III. Results and Discussion

A. Graphical Network Reconstruction

We have reconstructed the phosphoprotein signaling network that represents the underlying network corresponding to the full time series data shown in Fig. 3. In this network, out of 17×17 possible connections, only 35 were significant, many of which have negative coefficients in matrix A_1. Connections with negative coefficients are considered as inhibitory relationships shown in Fig. 3. Important inhibitory edges include AKT → GSKα/β [24]–[26], ERK1/2 → RSK [27], [28]. Different edge-widths are used to indicate edges with low, medium or high correlation.

To test the robustness of our model to the choice of α and correlation threshold, we used different correlation thresholds and confidence intervals (for the two tailed t-test) to reconstruct the underlying network. To evaluate the performance of each trial, we compared the significant connections identified for the underlying network to the true connections from the literature.

Table I implies that by increasing α from 0.01 to 0.02 and 0.05, i.e., reducing the confidence interval from 99% to 98% and to 95%, the number of False Positives increase and thus, Type I error increases. We also tested the results for different correlation thresholds that

result in further trimming of the parameters. The optimal correlation threshold for which Type I and Type II errors are both minimized, is C = 0.4.

We also studied the effect of more fine time-intervals. If we interpolate with steps of half a minute instead of one minute, the accuracy of the model does not change significantly. With a sample time of one minute, accuracy is 0.86, and with that of half a minute, accuracy is 0.87. We found that by using the cubic interpolation rather than linear interpolation, Type II error increases, justifying the use of linear interpolation.

Many of the connections found using our approach (underlying network, Fig. 3) were also identified using a PLS-based approach [4]. There are some differences between our network and the network obtained using the PLS approach. The connections p38 ↔ p65, p65 → ERK1/2 and GSKα → RSK are found in our network (Fig. 3), but not in the PLS-based network. However, the connections PKCD → EZR, MOE/EZR → RSK and p38 → AKT are found using the PLS approach, but are absent in our network.

The correlation coefficients with their corresponding p-values, along with the Benjamini-Hochberg FDR and p-values based on the t-test on the model coefficients for the connections retained in the underlying network (Fig. 3) are listed in Table II. It can be noted, that the Benjamini-Hochberg FDR for all these connections/edges are less than 0.026. The distribution of the p-values (t-test on the model coefficients) from all 17 × 17 possible connections for the underlying network is shown in Fig. 4 (implicitly used to calculate FDR).

We also present the dynamic evolution of the network in three temporal stages shown in Fig. 5. The topology of the phosphoprotein network changes through time. Fig. 5(a) corresponds to the reconstructed network in the first stage of the network development. Fig. 5(b) and (c) correspond to the reconstructed phosphoprotein networks for the second and third stages of the network evolution, respectively. The inhibitory edges such as AKT → GSKα/β are shown in Fig. 5.

1) Effect of Single-Ligand Data versus Double-Ligand Data—To evaluate the consistency of the data across experiments involving different ligand combinations, we applied the VAR model to single ligand experiments (22 experiments). According to our results, the reconstructed network based on only single ligand experiments has higher Type I and Type II error. We also used only the double ligand experiments to model the network, and as we anticipated, the performance does not change significantly. It can be noted that the double ligand combinations result in activation of the signaling pathway in ways that are functionally distinct from single ligand experiments. Furthermore, as an estimate of the differences in the variability for different phosphoproteins across time and treatment, we computed the ratio of the standard deviation of the standard deviation (*std*) to the mean of the *std* of every phosphoprotein (*std* is computed at every time for every treatment, using the replicate data), and found that this measure is of the same order (about 1) for all phosphoproteins across experiments.

B. Temporal Evolution of the Phosphoprotein Network

In this subsection, we discuss the dynamic nature of the phosphoprotein network evolving in three successive temporal stages. For the sake of simplicity in our discussions, we treat each phosphoprotein as a node and each regulatory interaction as an edge in the network analysis.

Stage 1 [Fig. 5(a)] shows the initiation of interactions among phosphoproteins. Since this network captures the early phase of the response of the system to the ligands, there are very few interactions taking place in the network. Extracellular signal-regulated kinase (ERK) plays a crucial role in the regulation and phosphorylation of most of the proteins that are present in the first stage of the network including p38 MAP Kinase (p38), p90 ribosomal S6 kinase (RSK), glycogen synthase kinase-3 (GSK), and protein kinase C (PKC) D. Ribosomal protein S6 (S6) affects ERK1 and ERK2. There is also a regulatory interaction between Nuclear Factor Kappa B (NF-κB p65) and p38. In addition, it is evident that Moesin (MOE) and Ezrin/Radixin (EZR) are part of the same pathway since a bidirectional link exists between them. As the network progresses to stage 2, several other interactions emerge. Fig. 5(b) shows that protein kinase B (AKT) arises in stage 2 and regulates the phosphorylation of GSKα/β. The signal transducer and activator of transcription 1 A and B (STAT1A/B, also ST1A/B for short) pairs are variants of the same protein and are expected to be activating one another. Indeed, they show a bidirectional relationship. PKCD that was regulated by ERK2 in stage 1, now promotes the phosphorylation of EZR and mother against decapentaplegic homolog 2 (SMD2), as well as mutually regulating neutrophil cytosolic factor 4 (p40). In stage 2, PKCM also appears and plays role in the regulation of RSK, S6 and ERK1/2, while being activated by p38. Role of S6 almost stays unchanged; i.e., it continues to regulate ERK1/2, except that as a result of the network progression from stage 1 to stage 2, we also see its interaction with RSK. This progression also brings about the phosphorylation of GSKα/β by RSK. In stage 1, p38 was activated by p65 and ERK2, whereas in the second stage, p38 regulates ERK1/2 along PKCM and gets involved in a mutual regulatory relationship with p65. p65 also affects ERK1/2 as well as RSK.

The evolution of the network to stage 3 provides not only most of the links that existed in stage 2, but also includes some new interactions. For instance, AKT proceeds to phosphorylate GSKα/β, while other nodes such as p65, RSK and p38 start to have causal influences on the activation of GSKα/β. Furthermore, in this phase, PKCD is regulated with the activation of PKCM, p40, SMD2 and EZR. Another interesting change is that p65 takes part in the activation of PKCM and ERK1/2. Moreover, AKT, broadly known for the activation of GSK, gets involved in the activation of S6, while being activated by ERK2.

Upon careful investigation of the time-dependent cascade of the network, we realize that there are very few stableinteractions that exist in all three stages. Moreover, the well-known signaling pathways such as the MAPK, STAT1A/B, AKT/GSK and NF-κB pathways emerge only in the last two stages and not in stage 1. The few causal interactions that persist throughout the temporal progression of the network are S6 → ERK1/2, EZR → MOE, p38 → RSK and p65 → p38. The time-varying succession of the significant interactions along with the related literature which validates some of these connections is shown in Table III.

C. Summary of Results

We have used a linear-model structure, least-squares regression and statistical hypothesis testing (t-test) on the coefficients of the linear model to identify significant edges in the network. Two types of networks have been identified, (1) based on the entire (interpolated) time-course data during [1–10] min, referred to as the underlying network (Fig. 3), and (2) temporally evolving network, in three-stages, based on three overlapping temporal regimes (Fig. 5). There is considerable overlap between our networks and a network obtained by a PLS-based approach published in the literature [4]. The temporally-evolving network of Fig. 5(a) shows the initiation of interactions among the phosphoproteins in stage 1 (e.g., ERK → p38/RSK,/GSK/ PKCD and S6 → ERK1/2), and the addition (e.g., AKT → GSKα/*β* and PKCM → RSK,/S6/ ERK1/2 during stage 1 → stage 2) or deletion (ERK2 → PKCD during stage 1 → stage 2) of specific connections with progress to stages 2 and 3. Persistent connections throughout the temporal progression of the network are S6 → ERK1/2, EZR → MOE, p38 → RSK and p65 → p38. We also found that the reconstructed network based on only single ligand experiments has higher Type I and Type II error as compared to using both single- and double-ligand data.

D. Validation of Results and Discussion

The results shown above are acquired through data-driven reconstruction of the network with no *a priori* information about the behavior of the underlying biological system. Here, we inspect our results and compare them with the existing information in the biology literature. In Table III, every causal relationship between pairs of phosphoproteins is shown by a directed arrow, and each mutual interaction is shown by a bi-directed arrow.

1) Role of AKT/GSK—GSK mediates protein phosphorylation and is involved in various intracellular pathways, metabolism and cancer. In mammalian cells GSK is encoded by two genes GSKα and GSK*β*, with similar biochemical and substrate properties. GSK targets proteins that are involved in Alzheimer's disease and neurological disorders. AKT is broadly known for activation and inhibition of GSK phosphorylation in HEK293 (Human Embryonic Kidney 293) cells, zebrafish and xenopus embryo [24]–[26]. We can readily see that the relatronships AKT → GSKα and AKT → GSK/3, representing phosphorylation of GSKα and GSK*β* by AKT, are captured in our model. Our results also indicate that the bidirectional connection AKT ↔ GSK*β* exists in second and third stages. In addition to AKT, recent studies show that RSK plays a role in modulating the activity of GSK in cerebral granule neurons, xenopus development and intracellular neural signaling systems [29]–[31]. There is also indication that the activation of RSK is responsible for the phosphorylation of GSK*β* induced by epidermal growth factor (EGF) in human epidermoid A431 cells [32], and that GSK*β* expressed in HeLa cells (from human cervical cancer cell line) is phosphorylated on Ser-9 by activation of p90Rsk [33]. Our model suggests the connection RSK → GSK*α/β* in stages 2 and 3, and the reverse connection GSKα → RSK in the underlying network. In previous studies it has been discovered *in vitro* that GSK is differentially regulated by the stimulation of PKC in rabbit skeletal muscle cells, Sf9 cells and HEK293 cells [34]–[36].

Another phosphoprotein involved in the regulation of GSK is p38. Recent studies indicate that p38 induces GSK phosphorylation in brain, thymocytes and human breast cancer cells (MDA-MB-231 cells) [37], [38] which is detected in the last two stages in our network. Furthermore, ERK activates GSK through phosphorylation in Hep-G2 cells and myocardial tissue cells in mice [39], [40]. We detect this relationship in the first two stages. Moreover, the existing knowledge illustrates that GSK is involved in the activation of p65 in hepatocytes from mice and HeLa cells [41], [42] while our model captures the reverse connection p65 → GSKα/β in stage 3.

2) EZR and MOE—EZR and MOE are part of the same pathway, called Ezrin/radixin/moesin (ERM) protein pathway. The ERM proteins regulate actin cytoskeleton and are involved in signaling, transport, and structural functions of the cell [43], [44]. As we can see in Fig. 2, the heat-map shows high correlation between these variables. In addition, the pairs ERK1/2 and STAT1A/B are variants of the same protein and are expected to be regulated similarly. Thus, as expected, high correlations and bidirectional causal relationships are observed between the members of each pair in Figs. 3 and 5. Despite the fact that the heat-map in Fig. 2 shows very high correlation between GSKα and β in all stages, we observe the connection GSK$\beta \rightarrow$ GSKα only in stage 2. This is an interesting result confirming the fact that "correlation does not imply causality" in the sense that the two variables may be highly correlated but there is no information in the past of one of them that can be used to predict the future of the other. The same result was found for PKCD/M. The connection PKCM → PKCD was found only in stage 3.

3) S6 and RSK—Ribosomal protein S6, which is involved in cell growth and regulation of cellular translation, is phosphorylated at several serine residues with mitogen stimulation by activation of one or more protein kinase cascades. It is well known that in mammalian cells, phosphorylation of ribosomal protein S6 *in vitro* and *in vivo* is regulated by the activation of RSK [18], [28], while our results indicate the existence of a bidirectional connection S6 ↔ RSK. RSK is involved in receptor-mediated signal transduction. Phosphorylation of RSK, which promotes cell survival and proliferation, lies at the end of the signaling cascade mediated by ERK and is regulated through the activation of ERK subfamily of MAP kinases [27], [28]. We observed this relationship in the first and second stages. Furthermore, our network suggests that RSK can be activated by p38 through the connection p38 → RSK in stages 1 and 2 and p38 ↔ RSK in stage 3. In current literature there is some evidence confirming this interaction in HEK293 cells [45]. Protein kinase C (PKC) is a family of fatty acid-activated protein kinase enzymes that is involved in regulating cell growth, learning and memory, transcription and mediating immune response. PKC which exists in various isoforms, is known to be involved in the activation of ERK in FJEK293 cells [46], which then results in the activation of RSK through the MAP kinase pathway [28]. Therefore it is anticipated that RSK and PKC have a hidden indirect relationship that was captured in our model where the connection PKCM → RSK is found in stage 3 and the underlying network and the connection PKCM ↔ RSK is found in stage 2. Our model still captured this connection by considering a faster time step (half a minute) in the model. In addition, PKC mediates the phosphorylation of S6 *in vivo* in HEK 293 cells [47]. PKCM → S6 can be found in stage 2 and the underlying network and PKCD → S6 in stage 3.

4) ERK and P38 (MAPK)—There are three distinct subfamilies of MAPK pathway: ERK1/2, INK and p38 MAP kinases that have substantial impact on mediating various cellular signaling functions and physiological processes. These three enzymes are part of a phosphorylation system in which they regulate and phosphorylate one another [48]. In this study we do not analyze the role of JNK in the signaling pathway, and we focus on the role of ERK1/2 and p38 in regulation and phosphorylation of one another and other phosphoproteins. The activation or inhibition of p38 potentiates the activation of ERK [49]–[51]. Unlike other pathways that appear only in the last two stages in our results, the crosstalk between ERK and p38 is found in all three stages. The activation of NF-kappa B (p65) can be triggered by the phosphorylation of ERK1/2 and recent research affirms the existence of cross-talk between ERK and p65 and between p65 and p38 [52]–[54] that can be seen in Fig. 5. p38 MAPK plays a critical role as downstream effector of PKC enzymes in LNCaP human prostate cancer cells and SK-Hep-1 hepatocellular carcinoma cells [55], [56]. Our results indicate the connections p38 ↔ PKCM in stage 3 and the underlying network, and p38 → PKCM in stage 2. Furthermore, p38 modulates the phosphorylation of subfamilies of RSK such as 70 kDa ribosomal S6 kinase (p70S6K) and ribosomal S6 kinase 1 (S6K1) [37], [69]. We also know that RSK's target substrate is S6 [18], [28]. This implies that p38 may indirectly play a role in the phosphorylation of S6. Our findings indicate that the connection p38 → S6 exists in stage 3 and the underlying network. There is no evidence in the existing literature confirming this relationship. The correlation coefficients for these edges are close to the correlation threshold. With a faster time step in the model, this connection is no longer significant. Hence, this interaction can be considered as false positive in our results. Moreover, phosphorylation of ribosomal protein S6 is known to be dependent upon the activation of ERK in HeLa cells and in mouse dentate gyms [28], [65] whereas our model captured the reverse connection.

Recent evidence implies that stimulation of PKC activates ERK1 and ERK2 in myocardial cells of rabbit, glomeruli of diabetic rats and glomerular mesangial cell cultures under high glucose conditions and in human neutrophil cells [57]–[59]. In our results, this relationship arises in the last two stages.

5) p65—Nuclear Factor Kappa B (NF-κB) exists in almost all animal cell types and is involved in mRNA transcription, regulation of inflammation, apoptosis and immune responses. There is some evidence that p65 NF-κB exists in the cytoplasm of unstimulated cells in an inactive form, and that it can be activated by exposure to PKC in human YT cells [62]–[64], whereas our results captured the reverse connection p65 → PKCM. It is interesting that previous computational methods such as those in [4] also captured the same reverse connection. Furthermore, there is some evidence that activation of NF-κB requires RSK-dependent p65 phosphorylation in vascular smooth muscle cells [67], [68] but extended analysis is needed to thoroughly understand the role of p65 in the biological function of RSK [66]. Our model estimated the opposite relationship p65 → RSK in stages 2, 3 and in the underlying network. Interestingly, in our analysis, the coefficient for RSK → p65 is just below the threshold and hence is not included in the network.

6) Other Pathways—Recent studies show evidence that activation of AKT inhibits the activation of the ERK pathway in C2C12 mouse myoblast cells [71] and that specific drugs unravel the crosstalk between the AKT and ERK pathways in neural stem cells [70]. In fact, we found the connection ERK2 → AKT in stage 3. SMD2 relays extracellular signals from transforming growth factor beta (TGF-β) ligands to the nucleus [78], [79]. There is some evidence that activation of SMAD (SMAD2, also SMD2) is modulated by protein kinase C in NIH-3T3 cells [72], [73], while the connection PKCD → SMD2 in stage 2 and the underlying network and PKCD ↔ SMD2 in stage 3 is captured in our networks. Some evidence provide affirmation that phosphorylation of ezr/radixin/moesin (ERM) is dependent upon catalytic function of PKC in MCF-7 breast cancer cells and in endothelial cells [60], [61]. Our network reconstruction captures PKCD → EZR in stage 2, PKCD ↔ EZR and the reverse connection, MOE → PKCD, in stage 3.

The current knowledge confirms that p40 is phosphorylated *in vitro* by protein kinase C in HL-60 cells and human neutrophils [74]–[76]. The bidirectional connection PKCD ↔ p40 was found in stage 2 and 3 of our reconstructed network. Our model also captures the connection p40 → SMD2 in stage 2.

AKT → S6 appears in stage 3 of our networks. It is known that protein kinase B (AKT) plays a role in the phosphorylation of RSK in human 293 cells [77] and ribosomal protein S6 (S6) is a substrate of RSK [18], [28]. Thus, it can be anticipated that AKT is capable of having an indirect impact on the phosphorylation of S6. This connection is statistically significant even with a faster time step in the model. Another potential novel connection is the crosstalk between GSK and ezrin/radixin/moesin (ERM), GSK → ERM [4].

7) Relationship of Signaling Pathways With Diseases—Some of these pathways such as p38 and NF-κB regulate the transcription of the cytokine tumor necrosis factor α (TNFα) which is a target for rheumatoid arthritis [80]. NF-κB is involved in the regulation of pro-inflammatory chemokines and cytokines in meningitis [81]. Furthermore, deviations in the levels of MAPKs from their normal cellular levels have been implicated in the development of cancer [82].

IV. Conclusion

We have applied the notion of Granger causality through the vector autoregressive model to develop a novel framework for reconstructing dynamic networks from large-scale multi-experiment multivariate high-throughput data sets. We used an approach based on a linear-model template and statistical hypothesis testing (t-test) of the coefficients of the model to find significant or potentially causal connections. We have applied this methodology to phosphoprotein time-course data generated by the Alliance for Cellular Signaling (AfCS) in RAW 264.7 macrophage cells in single and double ligand experiments. We were able to predict connectivity, causality and dynamics of information flow in the progression of the phosphoprotein network. We also found that the reconstructed network based on only single ligand data has higher Type I and Type II error as compared to using both single- and double-ligand data.

Since the intracellular networks have a dynamic nature and their topology changes with time, in this work, our main goal was to investigate the temporal evolution of the phosphoprotein network. During the early stage, ERK plays an important role in regulating p38, RSK, PKCD and GSK, while ERK itself is regulated by S6. As the network evolves to the second and third stages, the well-known signaling pathways such as the MAPK, STAT1A/B, AKT/GSK and NF-κB pathways appear to play role in the network. These results have enhanced our knowledge about the important signaling pathways that activate macrophage cells and play an essential role in the secretion of cytokines during an inflammatory response, and may contribute to finding novel targets for inflammation-related diseases.

The method we have developed and applied here provides a strategy for reconstructing and analyzing dynamical networks in biological systems. In addition to providing networks in the temporal context, our method provides the directionality and potential causality of molecular interactions. We note that we built our methodology based on the notion of Granger causality, which is not meant to be equivalent to the true causality.

Acknowledgment

The authors would like to acknowledge the phosphoproteomic data made available by the Alliance for Cellular Signaling (AfCS) laboratories.

This research was supported by the National Science Foundation (NSF) collaborative grants DBI-0835541 and STC-0939370, and National Institutes of Health (NIH) collaborative grant U54GM69338, and NIH R01 grants HL087375-02, HL106579 and HL108735.

Appendix: Abbreviation for the Names of the Phosphoproteins

c-Jun N-terminal Kinases lg (JNK lg), c-Jun N-terminal Kinases sh (JNK sh); Extracellular-signal Regulated Kinases (ERK) 1 (ERK1) and ERK 2 (ERK2); Ezrin [Ezr]/Radixin [Rdx] (Ezr/Rdx); Glycogen Synthase Kinase 3 (GSK) α (GSKα) and β (GSKβ); Membrane-organizing Extension Spike Protein (Moesin or MSN); Nuclear Factor Kappa-light-chain-enhancer of activated B cells p65 (NF-κB p65); Protein kinase B (AKT); Protein Kinase C (PKC) δ (PKCD) and PKC μ (PKCM); Ribosomal Protein S6 (S6); Ribosomal S6 kinase (RSK); Signal Transducers and Activator of Transcription (STAT) 1α (STAT1A, ST1A), STAT1β (STAT1B, ST1B), STAT3 and STAT5, Sma and Mad related proteins (SMAD) 2 (SMAD2, SMD2).

References

1. Delom F, Chevet E. Phosphoprotein analysis: From proteins to proteomes. Proteome Sci. 2006; 4:15. [PubMed: 16854217]
2. Hunter T. Protein kinases and phosphatases: The yin and yang of protein phosphorylation and signaling. Cell. 1995 Jan.80:225–236. [PubMed: 7834742]
3. Papin JA, Hunter T, Palsson BO, Subramaniam S. Reconstruction of cellular signalling networks and analysis of their properties. Nat. Rev. Mol. Cell. Biol. 2005 Feb.6:99–111. [PubMed: 15654321]
4. Gupta S, Maurya MR, Subramaniam S. Identification of crosstalk between phosphoprotein signaling pathways in RAW 264.7 macrophage cells. PLoS Comput. Biol. 2010; 6:e1000654. [PubMed: 20126526]

5. Pradervand S, Maurya MR, Subramaniam S. Identification of signaling components required for the prediction of cytokine release in RAW 264.7 macrophages. Genome Biol. 2006; 7:R11. [PubMed: 16507166]
6. Dojer N, Gambin A, Mizera A, Wilczynski B, Tiuryn J. Applying dynamic Bayesian networks to perturbed gene expression data. BMC Bioinform. 2006; 7:249.
7. Friedman N, Linial M, Nachman I, Pe'er D. Using Bayesian networks to analyze expression data. J. Comput. Biol, J. Comput. Mol. Cell Biol. 2000; 7:601–620.
8. Sachs K, Perez O, Pe'er D, Lauffenburger DA, Nolan GP. Causal protein-signaling networks derived from multiparameter single-cell data. Science. 2005 Apr.308:523–529. [PubMed: 15845847]
9. Haider S, Pal R. Inference of a genetic regulatory network model from limited time series data. Proc. IEEE Int. Workshop Genomic Signal Proc. and Statistics. 2011:162–165.
10. Pal R, Datta A, Bittner ML, Dougherty ER. Intervention in context-sensitive probabilistic Boolean networks. Bioinform. 2005 Apr.21:1211–1218.
11. Damiani C, Lecca P. Model identification using correlation-based inference and transfer entropy estimation. Proc. 5th UKSim Eur. Symp. Comput. Modeling and Simulat. 2011:129–134.
12. Altay G. Empirically determining the sample size for large-scale gene network inference algorithms. IET Syst. Biol. 2012 Apr.6:35–U28. [PubMed: 22519356]
13. Altay G, Emmert-Streib F. Inferring the conservative causal core of gene regulatory networks. SMC Syst. Biol. 2010; 4:132.
14. Mestl T, Plahte E, Omholt SW. A mathematical framework for describing and analyzing gene regulatory networks. J. Theor. Biol. 1995 Sep.176:291–300. [PubMed: 7475117]
15. Xiong M, Li J, Fang X. Identification of genetic networks. Genetics. 2004 Feb.166:1037–1052. [PubMed: 15020486]
16. Chang YH, Tomlin C. Inference of temporally evolving network dynamics with applications in biological systems. Proc. 50th IEEE Conf. Decision and Control and Eur. Control Conf. 2011:3706–3711.
17. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. Econometrica. 1969; 37:414–417.
18. Sturgill TW, Wu J. Recent progress in characterization of protein kinase cascades for phosphorylation of ribosomal protein S6. Biochim Biophys. Acta. 1991 May.1092:350–357. [PubMed: 1646641]
19. Fujita A, Sato JR, Garay-Malpartida HM, Morettin PA, So-gayar MC, Ferreira CE. Time-varying modeling of gene expression regulatory networks using the wavelet dynamic vector autoregressive method. Bioinform. 2007 Jul.23:1623–1630.
20. Hansen MH, Yu B. Model selection and the principle of minimum description length. J. Amer. Stat. Assoc. 2001 Jun.96:746–774.
21. Lütkepohl, H. New Introduction to Multiple Time Series Analysis. New York, NY, USA: Springer; 2006.
22. Benjamini Y, Hochberg Y. Controlling the false discovery rate – A practical and powerful approach to multiple testing. J. Royal Stat. Soc. Ser. B-Methodolog. 1995; 57:289–300.
23. Asadi B, Maurya MR, Tartakovsky DM, Subramaniam S. Comparison of statistical and optimisation-based methods for data-driven network reconstruction of biochemical systems. IET Syst. Biol. 2012 Oct.6:155–163. [PubMed: 23101870]
24. Ali A, Hoeflich KP, Woodgett JR. Glycogen synthase kinase-3: Properties, functions, and regulation. Chem. Rev. 2001 Aug.101:2527–2540. [PubMed: 11749387]
25. Cross DA, Alessi DR, Cohen P, Andjelkovich M, Hemmings BA. Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B. Nature. 1995 Dec.378:785–789. [PubMed: 8524413]
26. Lee HC, et al. Glycogen synthase kinase 3 alpha and 3 beta have distinct functions during cardiogenesis of zebrafish embryo. BMC Dev. Biol. 2007; 7:93. [PubMed: 17683539]

27. Richards SA, Fu J, Romanelli A, Shimamura A, Blenis J. Ribosomal S6 kinase 1 (RSK1) activation requires signals dependent on and independent of the MAP kinase ERK. Curr. Biol. 1999 Jul-Aug;9:810–820. [PubMed: 10469565]
28. Roux PP, et al. RAS/ERK signaling promotes site-specific ribosomal protein S6 phosphorylation via RSK and stimulates cap-dependent translation. J. Biol. Chem. 2007 May.282:14056–14064. [PubMed: 17360704]
29. Grimes CA, Jope RS. The multifaceted roles of glycogen synthase kinase 3beta in cellular signaling. Prog. Neurobiol. 2001 Nov.65:391–426. [PubMed: 11527574]
30. Li M, Wang X, Meintzer MK, Laessig T, Birnbaum MJ, Heidenreich KA. Cyclic AMP promotes neuronal survival by phosphorylation of glycogen synthase kinase 3beta. Mol. Cell. Biol. 2000 Dec.20:9356–9363. [PubMed: 11094086]
31. Torres MA, Eldar-Finkelman H, Krebs EG, Moon RT. Regulation of ribosomal S6 protein kinase-p90(RSK), glycogen synthase kinase 3, and beta-catenin in early Xenopus development. Mol. Cell. Biol. 1999 Feb.19:1427–1437. [PubMed: 9891076]
32. Saito Y, Vandenheede JR, Cohen P. The mechanism by which epidermal growth factor inhibits glycogen synthase kinase 3 in A431 cells. Biochem. J. 1994 Oct.303(pt. 1):27–31. [PubMed: 7945252]
33. Stambolic V, Woodgett JR. Mitogen inactivation of glycogen synthase kinase-3 beta in intact cells via serine 9 phosphorylation. Biochem. J. 1994 Nov.303(pt. 3):701–704. [PubMed: 7980435]
34. Fang X, Yu S, Tanyi JL, Lu Y, Woodgett JR, Mills GB. Convergence of multiple signaling cascades at glycogen synthase kinase 3: Edg receptor-mediated phosphorylation and inactivation by lysophosphatidic acid through a protein kinase C-dependent intracellular pathway. Mol. Cell. Biol. 2002 Apr.22:2099–2110. [PubMed: 11884598]
35. Goode N, Hughes K, Woodgett JR, Parker PJ. Differential regulation of glycogen synthase kinase-3 beta by protein kinase C isotypes. J. Biol. Chem. 1992 Aug.267:16878–16882. [PubMed: 1324914]
36. Hughes K, Nikolakaki E, Plyte SE, Totty NF, Woodgett JR. Modulation of the glycogen synthase kinase-3 family by tyrosine phosphorylation. Eur. Mol. Biol. Org. J. 1993 Feb.12:803–808.
37. Choi CH, Lee BH, Aim SG, Oh SH. Proteasome inhibition-induced p38 MAPK/ERK signaling regulates autophagy and apoptosis through the dual phosphorylation of glycogen synthase kinase 3beta. Biochem. Biophys. Res. Commun. 2012 Feb.418:759–764. [PubMed: 22310719]
38. Thornton TM, et al. Phosphorylation by p38 MAPK as an alternative pathway for GSK3beta inactivation. Science. 2008 May.320:667–670. [PubMed: 18451303]
39. Das A, Salloum FN, Xi L, Rao YJ, Kukreja RC. ERK phosphorylation mediates sildenafil-induced myocardial protection against ischemia-reperfusion injury in mice. Amer. J. Physiol Heart Circ. Physiol. 2009 May.296:H1236–H1243. [PubMed: 19286961]
40. Ding Q, et al. Erk associates with and primes GSK-3beta for its inactivation resulting in upregulation of beta-catenin. Mol. Cell. 2005 Jul.19:159–170. [PubMed: 16039586]
41. Buss H, et al. Phosphorylation of serine 468 by GSK-3beta negatively regulates basal p65 NF-kappaB activity. J. Biol. Chem. 2004 Nov.279:49571–49574. [PubMed: 15465828]
42. Schwabe RF, Brenner DA. Role of glycogen synthase kinase-3 in TNF-alpha-induced NF-kappaB activation and apoptosis in hepatocytes. Amer. J. Physiol. Gastroint. Liver Physiol. 2002 Jul. 283:G204–G211.
43. Sala-Valdes M, et al. EWI-2 and EWI-F link the tetraspanin web to the actin cytoskeleton through their direct association with ezrinradixin-moesin proteins. J. Biol. Chem. 2006 Jul.281:19665–19675. [PubMed: 16690612]
44. Tepass U. FERM proteins in animal morphogenesis. Curr. Opin. Genet Dev. 2009 Aug.19:357–367. [PubMed: 19596566]
45. Pierrat B, Correia JS, Mary JL, Tomas-Zuber M, Lesslauer W. RSK-B, A novel ribosomal S6 kinase family member, is a CREB kinase under dominant control of p38alpha mitogen-activated protein kinase (p38alphaMAPK). J. Biol. Chem. 1998 Nov.273:29661–29671. [PubMed: 9792677]
46. Brändlin SHI. Protein Kinase C (PKC)η-mediated PKCμ activation Modulates ERK and JNK Signal Pathways. J. Biol. Chem. 2002; 277:6490–6496. [PubMed: 11741879]

47. Valovka Y, et al. Protein kinase C phosphorylates ribosomal protein S6 kinase betaII and regulates its subcellular localization. Mol. Cell. Biol. 2003 Feb.23:852–863. [PubMed: 12529391]
48. Johnson GL, Lapadat R. Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases. Science. 2002 Dec.298:1911–1912. [PubMed: 12471242]
49. Chiacchiera F, et al. Blocking p38/ERK crosstalk affects colorectal cancer growth by inducing apoptosis in vitro and in preclinical mouse models. Cancer Lett. 2012 Nov.324:98–108. [PubMed: 22579651]
50. Fey D, Croucher DR, Kolch W, Kholodenko BN. Crosstalk and signaling switches in mitogen-activated protein kinase cascades. Front Physiol. 2012; 3:355. [PubMed: 23060802]
51. Liu Q, Hofmann PA. Protein phosphatase 2A-mediated crosstalk between p38 MAPK and ERK in apoptosis of cardiac myocytes. Amer. J. Physiol. Heart Circ. Physiol. 2004 Jun.286:H2204–H2212. [PubMed: 14962831]
52. Bhattacharyya A, Pathak S, Datta S, Chattopadhyay S, Basu J, Kundu M. Mitogen-activated protein kinases and nuclear factor-kappaB regulate Helicobacter pylori-mediated interleukin-8 release from macrophages. Biochem. J. 2002 Nov.368:121–129. [PubMed: 12150710]
53. Hayden MS, Ghosh S. Shared principles in NF-kappaB signaling. Cell. 2008 Feb.132:344–362. [PubMed: 18267068]
54. Perkins ND. Post-translational modifications regulating the activity and function of the nuclear factor kappa B pathway. Oncogene. 2006 Oct.25:6717–6730. [PubMed: 17072324]
55. Tanaka Y, Gavrielides MV, Mitsuuchi Y, Fujii T, Kazanietz MG. Protein kinase C promotes apoptosis in LNCaP prostate cancer cells through activation of p38 MAPK and inhibition of the Akt survival pathway. J. Biol. Chem. 2003 Sep.278:33753–33762. [PubMed: 12824193]
56. Hsieh TTWYH, Huang C-Y, Hsieh Y-S, Hwang J-M, Liu J-Y. p38 mitogen-activated protein kinase pathway is involved in protein kinase Cα–regulated invasion in human hepatocellular carcinoma cells. Cancer Res. 2007 May.
57. Downey GP, et al. Chemotactic peptide-induced activation of MEK-2, the predominant isoform in human neutrophils. Inhibition by wortmannin. J. Biol. Chem. 1996 Aug.271:21005–21011. [PubMed: 8702863]
58. Haneda M, Araki S, Togawa M, Sugimoto T, Isono M, Kikkawa R. Mitogen-activated protein kinase cascade is activated in glomeruli of diabetic rats and glomerular mesangial cells cultured under high glucose conditions. Diabetes. 1997 May.46:847–853. [PubMed: 9133554]
59. Ping P, et al. PKC-dependent activation of p44/p42 MAPKs during myocardial ischemia-reperfusion in conscious rabbits. Amer. J. Physiol. 1999 May.276:H1468–H1481. [PubMed: 10330229]
60. Koss M, et al. Ezrin/radixin/moesin proteins are phosphorylated by TNF-alpha and modulate permeability increases in human pulmonary microvascular endothelial cells. J. Immunol. 2006 Jan. 176:1218–1227. [PubMed: 16394012]
61. Ng T, et al. Ezrin is a downstream effector of trafficking PKC-integrin complexes involved in the control of cell motility. Eur. Mol. Biol. Org. J. 2001 Jun.20:2723–2741.
62. Shirakawa F, Mizel SB. In vitro activation and nuclear translocation of NF-kappa B catalyzed by cyclic AMP-dependent protein kinase and protein kinase C. Mol. Cell. Biol. 1989 Jun.9:2424–2430. [PubMed: 2548081]
63. Silberman DM, Zorrilla-Zubilete M, Cremaschi GA, Genaro AM. Protein kinase C-dependent NF-kappaB activation is altered in T cells by chronic stress. Cell. Mol. Life Sci. 2005 Aug.62:1744–1754. [PubMed: 16003495]
64. Wooten MW. Function for NF-kB in neuronal survival: Regulation by atypical protein kinase C. J. Neurosci. Res. 1999 Dec.58:607–611. [PubMed: 10561688]
65. Gangarossa G, Valjent E. Regulation of the ERK pathway in the dentate gyrus by in vivo dopamine D1 receptor stimulation requires glutamatergic transmission. Neuropharmacol. 2012 Nov.63:1107–1117.
66. Carriere A, Ray H, Blenis J, Roux PP. The RSK factors of activating the Ras/MAPK signaling cascade. Front Biosci. 2008; 13:4258–4275. [PubMed: 18508509]

67. Zhang JCL, Ma Y, Thomas W, Zhang J, Du J. Dual pathways for nuclear factor κB activation by angiotensin II in vascular smooth muscle phosphorylation of p65 by IκB kinase and ribosomal kinase. Circ. Res. 2005 Nov.97:975–982. [PubMed: 16224066]
68. Zhang L, Ma Y, Zhang J, Cheng J, Du J. A new cellular signaling mechanism for angiotensin II activation of NF-kappaB: An IkappaB-independent, RSK-mediated phosphorylation of p65. Arterioscler Thromb Vase. Biol. 2005 Jun.25:1148–1153.
69. Zbinden-Foncea H, Deldicque L, Pierre N, Francaux M, Raymackers JM. TLR2 and TLR4 activation induces p38 MAPK-dependent phosphorylation of S6 kinase 1 in C2C12 myotubes. Cell. Biol. Int. 2012; 36:1107–1113. [PubMed: 22931089]
70. Huang W, et al. Fluoxetine upregulates phosphorylated-AKT and phosphorylated-ERK1/2 proteins in neural stem cells: Evidence for a crosstalk between AKT and ERK 1/2 pathways. J. Mol. Neurosci. 2013 Feb.49:244–249. [PubMed: 22674052]
71. Rommel C, et al. Differentiation stage-specific inhibition of the Raf-MEK-ERK pathway by Akt. Science. 1999 Nov.286:1738–1741. [PubMed: 10576741]
72. Itoh S, Itoh F, Goumans MJ, Ten Dijke P. Signaling of transforming growth factor-beta family members through Smad proteins. Eur. J. Biochem. 2000 Dec.267:6954–6967. [PubMed: 11106403]
73. Yakymovych I, Ten Dijke P, Heldin CH, Souchelnytskyi S. Regulation of Smad signaling by protein kinase C. Fed. Amer. Soc. Exp. Biol. J. 2001 Mar.15:553–555.
74. Fontayne A, Dang PM-C, Gougerot-Pocidalo M-A, Benna JE. Phosphorylation of p47phox sites by PKC α, βII, δ, and ζ: Effect on binding to p22phox and on NADPH Oxidase activation. Biochem. (Mosc). 2002; 41:7743–7750.
75. Bouin AP, Grandvaux N, Vignais PV, Fuchs A. p40(phox) is phosphorylated on threonine 154 and serine 315 during activation of the phagocyte NADPH oxidase. Implication of a protein kinase c-type kinase in the phosphorylation process. J. Biol. Chem. 1998 Nov.273:30097–30103. [PubMed: 9804763]
76. Olavarria VH, Figueroa JE, Mulero V. Prolactin-induced activation of phagocyte NADPH oxidase in the teleost fish gilthead seabream involves the phosphorylation of p47phox by protein kinase C. Dev. Comp. Immunol. 2012 Jan.36:216–221. [PubMed: 21884725]
77. Dufner A, Andjelkovic M, Burgering BM, Hemmings BA, Thomas G. Protein kinase B localization and activation differentially affect S6 kinase 1 activity and eukaryotic translation initiation factor 4E-binding protein 1 phosphorylation. Mol. Cell. Biol. 1999 Jun.19:4525–4534. [PubMed: 10330191]
78. Heldin CH, Miyazono K, ten Dijke P. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. Nature. 1997 Dec.390:465–471. [PubMed: 9393997]
79. Massague J. TGF-beta signal transduction. Annu. Rev. Biochem. 1998; 67:753–791. [PubMed: 9759503]
80. Foxwell B, et al. Efficient adenoviral infection with IkappaB alpha reveals that macrophage tumor necrosis factor alpha production in rheumatoid arthritis is NF-kappaB dependent. Proc. Nat. Acad. Sci. USA. 1998 Jul.95:8211–8215. [PubMed: 9653166]
81. Doran KS, Liu GY, Nizet V. Group B streptococcal beta-hemolysin/cytolysin activates neutrophil signaling pathways in brain endothelium and contributes to development of meningitis. J. Clin. Invest. 2003 Sep.112:736–744. [PubMed: 12952922]
82. Kim EK, Choi EJ. Pathological roles of MAPK signaling pathways in human diseases. Biochim. Biophys. Acta. 2010 Apr.1802:396–405. [PubMed: 20079433]

Biographies

Maryam Masnadi-Shlrazi was born in Albuquerque, NM, USA. She received the B.S. degree in electrical engineering from Shiraz University, Shiraz, Iran, and the M.S. degree in electrical and computer engineering from the University of California, San Diego (UCSD), La Jolla, CA USA, in 2010 and 2012, respectively.

Currently, she is working toward the Ph.D. degree in electrical and computer engineering at UCSD. She is a Research Assistant under supervision of Prof. Shankar Subramaniam in the Bioinformatics and Systems Biology Lab in the Bioengineering Department at UCSD. Her research interests include statistical machine learning, parameter estimation, information theory and their applications to bioinformatics and systems biology.

Mano Ram Maurya completed the B.Tech. degree in chemical engineering from the Indian Institute of Technology Bombay, Powai, Mumbai, India, the M.E. degree in chemical engineering from the City College of New York, NY, USA, and the Ph.D. degree in chemical engineering from Purdue University, West Lafayette, IN, USA, in 1998, 1999, and 2003, respectively.

From 2003 to 2006, he was a Postdoctoral Researcher in the Department of Bioengineering and the San Diego Supercomputer Center at the University of California, San Diego (UCSD), La Jolla, CA, USA. He worked in the Department of Bioengineering as an Assistant Scientist from October 2006 to November 2010. Since then, he is a Research Scientist at the San Diego Supercomputer Center and in the Department of Bioengineering at UCSD. He is the author of approximately 40 papers in peer-reviewed journals and conference proceedings. His current research interests include the study of complex biochemical processes and pathways using systems engineering/biology, machine learning, information theory and bioinformatics approaches and their applications to biomedicine.

Dr. Maurya received the Best Paper Award in 2005 jointly with his Ph.D. advisor, Dr. Venkat Venkatasubramanian, and co-advisor, Dr. Raghunathan Ren-gaswamy, for their paper published in the *Journal of Engineering Applications of Artificial Intelligence* in 2004. In August 2011, he joined the editorial board of *ISRN Biophysics*.

Shankar Subramaniam received the B.S. and M.S. degrees from Osmania University, Hyderabad, India, and the Ph.D. degree in chemistry from the Indian Institute of Technology Kanpur, Kanpur, India.

From 1991 to 1999, he was a Professor of Biophysics, Biochemistry, Molecular and Integrative Physiology, Chemical Engineering and Electrical and Computer Engineering at the University of Illinois, Urbana-Champaign, Urbana-Champaign, IL, USA, prior to moving to the University of California, San Diego (UCSD), La Jolla, CA, USA. Currently, he is the Joan and Irwin Jacobs Endowed Chair in Bioengineering and Systems Biology at UCSD and a Professor of Bioengineering, Bioinformatics, and Systems Biology, Cellular and Molecular Medicine, Chemistry and Biochemistry, and Nanoengineering. He was the Founding Director of the Bioinformatics and Systems Biology Program at UCSD. He has authored more than 200 articles. Research in his laboratory spans several areas of bioinformatics, systems biology, and medicine. He is on the editorial board of several journals.

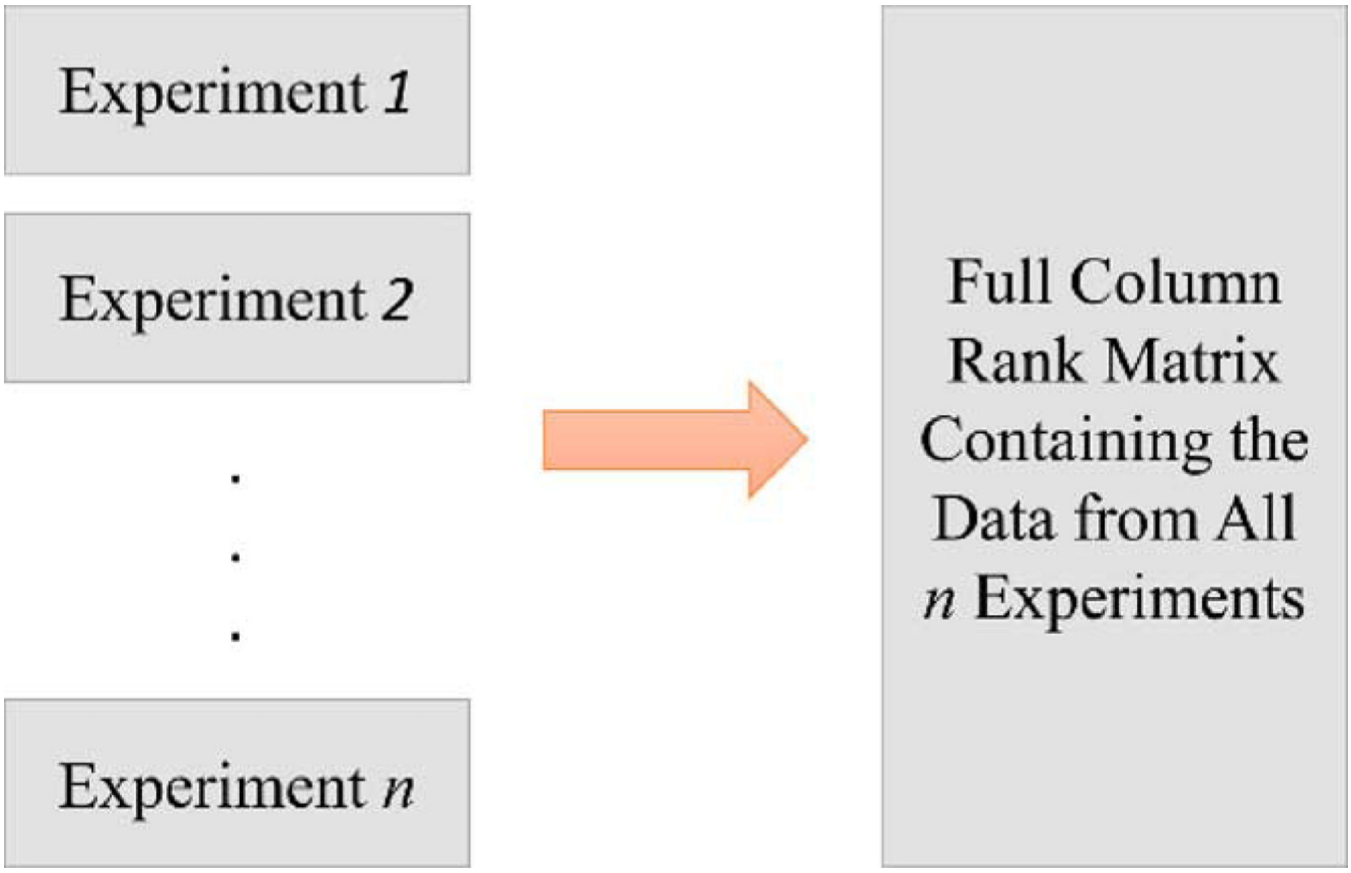

Fig. 1.
Schematic to show the stacking of the data matrices. Each column corresponds to the time series data of each of the *k* variables.

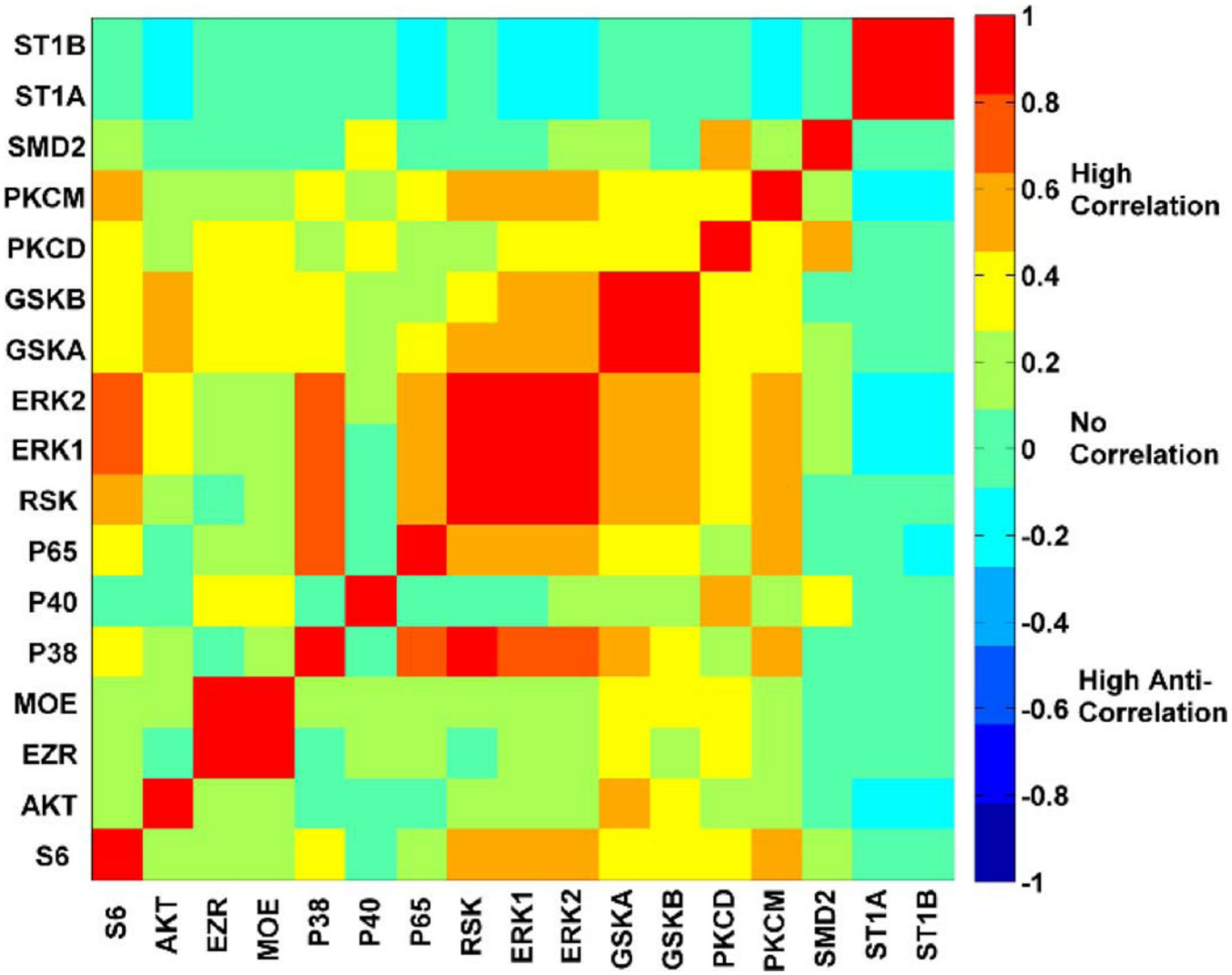

Fig. 2.

Heat-map of the correlation matrix between the input and output variables. This matrix contains the pairwise correlation coefficient between columns of matrix X and Y for the whole time series [1–10] minutes.

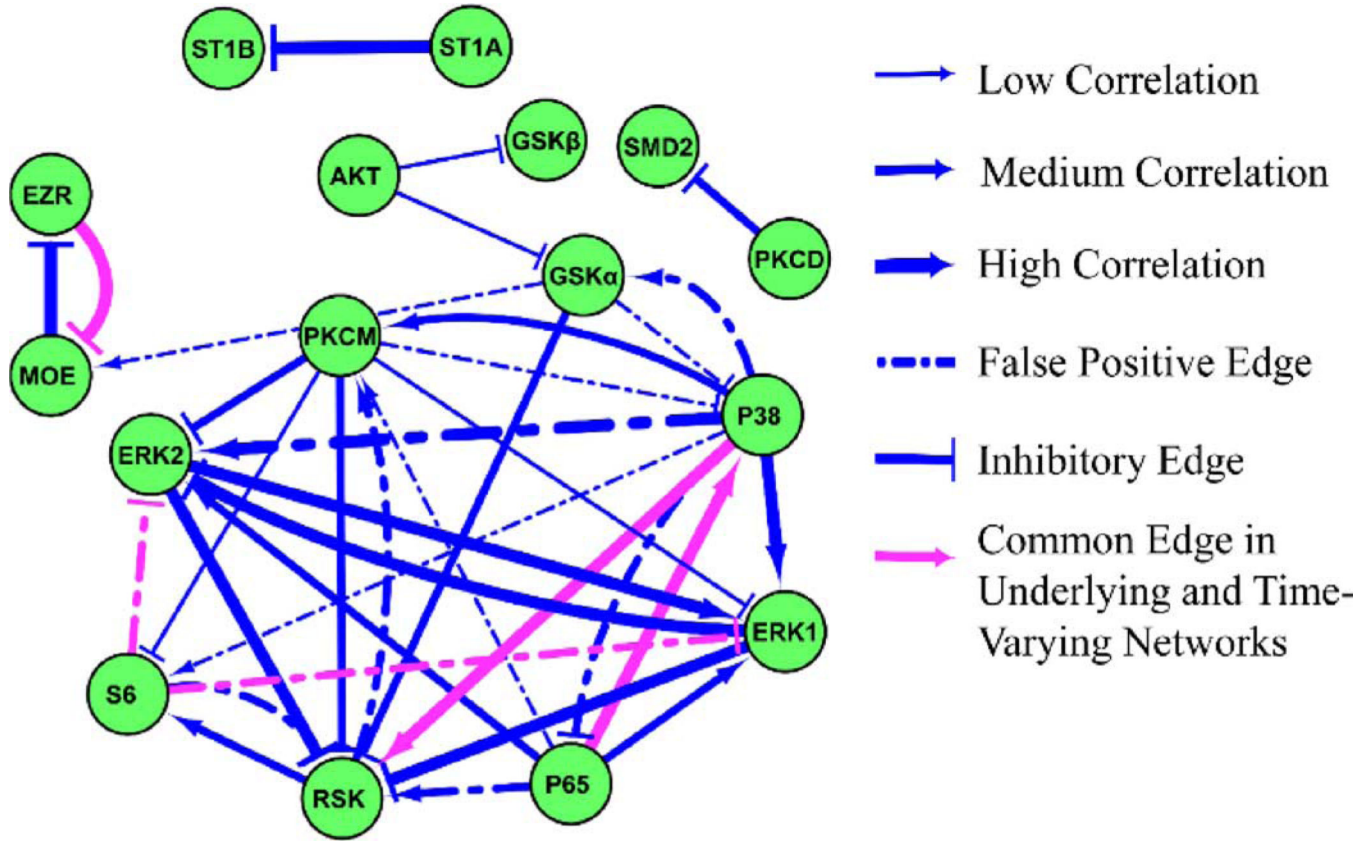

Fig. 3.

The reconstructed network for the underlying signaling network in RAW 264.7 macrophages. This network represents the cross-talk between phosphoproteins considering the whole time-series for [1]–[10] minute period. The pink connections are common edges in the underlying network and the timevarying network (Fig. 5). Different edge-widths are used to represent low ($0.4 \leq r < 0.5$), medium ($0.5 \leq r < 0.75$) and high ($r \geq 0.75$) correlation coefficients corresponding to the edges. Inhibitory connections are shown with a blunt end instead of an arrow.

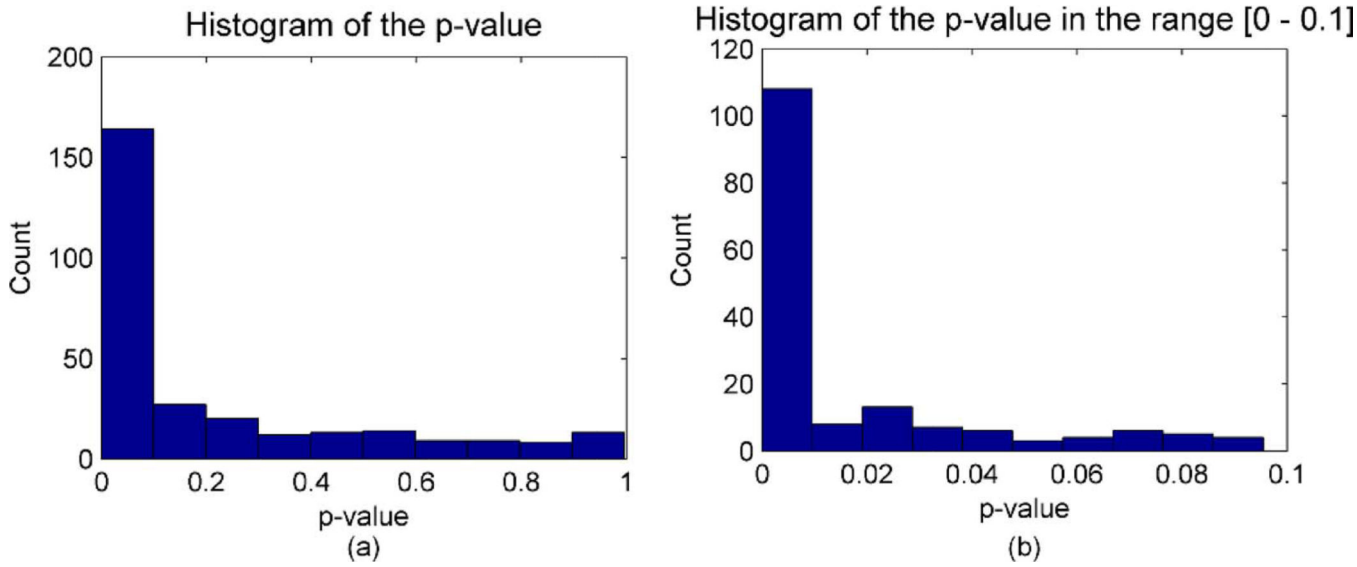

Fig. 4.

Histogram of the p-values (t-test on the model coefficients) for the underlying network generated from 17×17 p-value numbers.

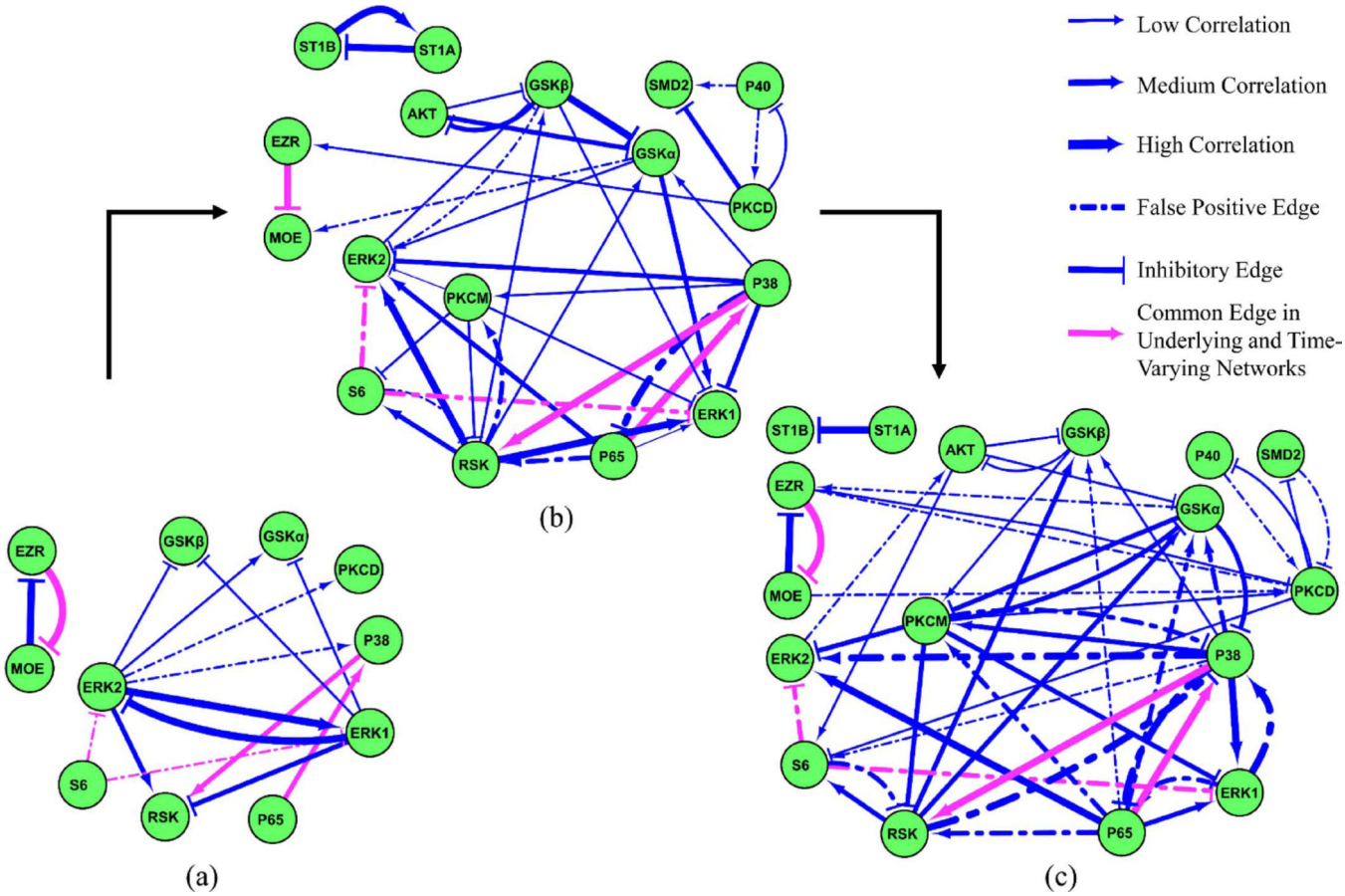

Fig. 5.

Time-dependent cascade of the phosphoprotein signaling network in RAW 264.7 macrophages in three stages, (a) Reconstructed network in stage 1 related to [1]–[4] minute interval. (b) Reconstructed network in stage 2 related to [3]–[7] minute interval, (c) Reconstructed network in stage 3 related to [6]–[10] minute interval. The pink connections are common to all the three networks as well as the underlying network (Fig. 3). Different edge-widths are used to represent low ($0.4 \leq r < 0.5$), medium ($0.5 \leq r < 0.75$) and high ($r \geq 0.75$) correlation coefficients corresponding to the edges. Inhibitory connections are shown with a blunt end instead of an arrow.

TABLE I

Robustness of Results of the Underlying Network to the Choice of Different Thresholds

Correlation Threshold	α	Type I Error	Type II Error	Accuracy
C=0.4	0.01	0.07	0.56	0.86
C=0.5	0.05	0.05	0.66	0.86
C=0.4	0.02	0.07	0.53	0.86
C=0.4	0.05	0.10	0.48	0.84
C=0.6	0.01	0.02	0.79	0.87
C=0.7	0.01	0.02	0.84	0.86

TABLE II

Correlation Coefficients and Statistical Significance of Edges Retained in the Underlying Network. Abbreviation: Benjamini-Hochberg (BH). False-Discovery Rate (FDR)

Edges		Correlation Based		Model Coefficient Based	
Source Node	**Target Node**	**Correlation Coefficient**	**p-value**	**FDR (BH)**	**p-value**
ERK2	**ERK1**	0.96	0	3.75E-10	3.89E-11
ST1A	**ST1B**	0.96	0	1.11E-02	3.74E-03
ERK1	**ERK2**	0.94	0	4.56E-09	5.52E-10
MOE	**EZR**	0.9	0	2.33E-08	2.98E-09
EZR	**MOE**	0.9	0	9.42E-03	3.00E-03
ERK2	**RSK**	0.85	0	1.29E-06	1.92E-07
P38	**RSK**	0.82	0	5.13E-03	1.56E-03
ERK1	**RSK**	0.82	0	1.57E-06	2.44E-07
P65	**P38**	0.81	0	0	0
P38	**ERK2**	0.8	0	1.04E-03	2.83E-04
P38	**ERK1**	0.77	0	1.89E-05	3.60E-06
P38	**P65**	0.72	0	1.50E-06	2.28E-07
P65	**ERK2**	0.62	0	0	0
RSK	**S6**	0.62	0	2.19E-04	4.99E-05
S6	**ERK1**	0.59	1.19E-279	0	0
S6	**ERK2**	0.59	1.21E-277	0	0
P65	**RSK**	0.59	2.50E-270	0	0
RSK	**PKCM**	0.58	6.52E-269	2.60E-02	9.80E-03
P65	**ERK1**	0.58	2.02E-262	0	0
P38	**PKCM**	0.54	6.74E-227	1.42E-11	1.43E-12
P38	**GSKA**	0.54	1.75E-218	4.38E-08	6.21E-09
S6	**RSK**	0.53	6.00E-214	0	0
PKCD	**SMD2**	0.51	1.29E-197	0	0
PKCM	**RSK**	0.51	2.73E-196	5.74E-06	9.92E-07
PKCM	**ERK2**	0.51	7.58E-193	7.67E-06	1.38E-06

IEEE Trans Biomed Circuits Syst. Author manuscript; available in PMC 2015 November 03.

Edges		Correlation Based		Model Coefficient Based	
Source Node	Target Node	Correlation Coefficient	p-value	FDR (BH)	p-value
GSKA	**RSK**	0.5	9.80E-188	2.12E-02	7.78E-03
PKCM	**ERK1**	0.5	5.13E-182	1.09E-03	3.09E-04
AKT	**GSKA**	0.48	4.45E-166	1.91E-04	4.16E-05
PKCM	**S6**	0.46	3.36E-156	2.04E-02	7.33E-03
P65	**PKCM**	0.45	7.45E-150	8.90E-04	2.34E-04
AKT	**GSKB**	0.45	5.72E-147	2.64E-06	4.39E-07
P38	**S6**	0.45	5.40E-144	1.03E-03	2.78E-04
GSKA	**MOE**	0.45	1.49E-143	3.00E-04	7.05E-05
PKCM	**P38**	0.44	9.25E-139	9.85E-03	3.20E-03
GSKA	**P38**	0.44	3.39E-136	2.66E-03	8.00E-04

TABLE III

Comparison of Our Results With the Current Literature

Correlated pairs	Stage 1	Stage 2	Stage 3	Underlying network	Current knowledge	References
(GSK, AKT)	—	AKT → GSKα/β	AKT → GSKα/β	AKT → GSKα/β	AKT → GSK	[24–26]
(GSK, RSK)	—	RSK → GSKα/β	RSK → GSKα/β	GSKα → RSK	RSK → GSK	[32, 33]
(GSK, P38)	—	P38 → GSKα	P38 → GSKα/ β	P38 ↔ GSKα	P38→ GSK	[37, 38]
(GSK, ERK)	ERK2/1 → GSKα/β	ERK2 ↔ GSKβ	—	—	ERK → GSK	[39,40]
(GSK, P65)	—	—	P65 → GSKα/β	—	GSK → P65	[41,42]
(RSK, S6)	—	S6 ↔ RSK	S6 ↔ RSK	S6 ↔ RSK	RSK → S6	[18,28]
(RSK, ERK)	ERK2 → RSK	RSK → ERK 1/2	—	ERK 1/2 → RSK	ERK → RSK	[27,28]
(RSK, P38)	p38 → RSK	P38 → RSK	P38 ↔ RSK	P38 → RSK	P38 → RSK	[45]
(PKC, S6)	—	PKCM → S6	PKCD → S6	PKCM → S6	PKC → S6	[47]
(PKC, P38)	—	P38 → PKCM	P38 ↔ PKCM	P38 ↔ PKCM	P38 → PKCM	[55, 56]
(PKC,ERK)	—	PKCM → ERK1/2	PKCM → ERKl/2	PKCM → ERKl/2	PKCM → ERK	[57–59]
(PKC, EZR)	—	PKCD → EZR	PKCD ↔ EZR	—	PKC → EZR	[60, 61]
(PKC, MOE)	—	—	MOE → PKCD	—	PKC → MOE	[60,61]
(PKC, P65)	—	—	P65 → PKCM	P65 → PKCM	PKC → P65	[62–64]
(PKC, RSK)	—	PKCM ↔ RSK	PKCM → RSK	PKCM ↔ RSK	PKC → ERK → RSK	[28,46]
(S6, **ERK)**	S6 → ERK1/2	S6 → ERK1/2	S6 → ERK1/2	S6 → ERK1/2	ERK → S6	[28, 65]
(P65, RSK)	—	P65 → RSK	P65 → RSK	P65 → RSK	RSK → P65	[66–68]
(P65, ERK)	—	P65 → ERK1/2	ERK1 → P65	P65 → ERK1/2	P65 → ERK	[52–54]
(P65, P38)	P65 → P38	P65 ↔P38	P65 ↔P38	P65 ↔P38	P65 → P38	[52–54]
(P38, ERK)	ERK2 → P38	P38 → ERK1/2	P38 → ERK2 P38 ↔ ERK1	P38 → ERK1/2	P38 → ERK	[49–51]
(P38, S6)	—	—	P38 → S6	P38 → S6	P38 → RSK → S6	[37, 69]
(AKT, ERK)	—	—	ERK2→ AKT	—	AKT → ERK	[70,71]
(SMD, PKC)	—	PKCD → SMD2	PKCD ↔ SMD2	PKCD → SMD2	PKC → SMD	[72, 73]
(SMD, P40)	—	P40 → SMD2	—	—	—	—
(P40, PKC)	—	PKCD ↔ P40	PKCD ↔ P40	—	PKC → P40	[74–76]
(AKT, S6)	—	—	AKT → S6	—	AKT → RSK → S6	[18,28,77]
(GSK, ERM)	—	GSKα → MOE	GSK → EZR	GSK → MOE	—	—