

Computational modeling with forward and reverse engineering links signaling network and genomic regulatory responses: NF- κ B signaling-induced gene expression responses in inflammation

Shih Chi Peng¹, David Shan Hill Wong², Kai Che Tung³, Yan Yu Chen², Chun Cheih Chao³, Chien Hua Peng⁴, Yung Jen Chuang^{*3} and Chuan Yi Tang^{*1}

Abstract

Background: Signal transduction is the major mechanism through which cells transmit external stimuli to evoke intracellular biochemical responses. Diverse cellular stimuli create a wide variety of transcription factor activities through signal transduction pathways, resulting in different gene expression patterns. Understanding the relationship between external stimuli and the corresponding cellular responses, as well as the subsequent effects on downstream genes, is a major challenge in systems biology. Thus, a systematic approach is needed to integrate experimental data and theoretical hypotheses to identify the physiological consequences of environmental stimuli.

Results: We proposed a systematic approach that combines forward and reverse engineering to link the signal transduction cascade with the gene responses. To demonstrate the feasibility of our strategy, we focused on linking the NF- κ B signaling pathway with the inflammatory gene regulatory responses because NF- κ B has long been recognized to play a crucial role in inflammation. We first utilized forward engineering (Hybrid Functional Petri Nets) to construct the NF- κ B signaling pathway and reverse engineering (Network Components Analysis) to build a gene regulatory network (GRN). Then, we demonstrated that the corresponding IKK profiles can be identified in the GRN and are consistent with the experimental validation of the IKK kinase assay. We found that the time-lapse gene expression of several cytokines and chemokines (TNF- α , IL-1, IL-6, CXCL1, CXCL2 and CCL3) is concordant with the NF- κ B activity profile, and these genes have stronger influence strength within the GRN. Such regulatory effects have highlighted the crucial roles of NF- κ B signaling in the acute inflammatory response and enhance our understanding of the systemic inflammatory response syndrome.

Conclusion: We successfully identified and distinguished the corresponding signaling profiles among three microarray datasets with different stimuli strengths. In our model, the crucial genes of the NF- κ B regulatory network were also identified to reflect the biological consequences of inflammation. With the experimental validation, our strategy is thus an effective solution to decipher cross-talk effects when attempting to integrate new kinetic parameters from other signal transduction pathways. The strategy also provides new insight for systems biology modeling to link any signal transduction pathways with the responses of downstream genes of interest.

Background

Signal transduction is a complex process in which a cell converts environmental signals to a series of intracellular

biochemical reactions. Diverse cellular stimuli can create a wide variety of transcription factor activities through signal transduction pathways, resulting in differential gene expression that dictates subsequent cellular behaviors. Although a great deal of effort has been made in modeling signal transduction pathways or gene regulatory networks independently, a strategy to link the signaling pathway with downstream gene expression responses

* Correspondence: yjchuang@life.nthu.edu.tw, cytang@cs.nthu.edu.tw

¹ Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, 30013, Taiwan, ROC

² Department of Computer Science, National Tsing Hua University, Hsinchu, 30013, Taiwan, ROC

Full list of author information is available at the end of the article

seems to be lacking. Thus, a systematic approach is needed to integrate experimental data and theoretical hypotheses, to identify the physiological consequences of environmental stimuli.

Over the past few years, a considerable number of studies have reported various systematic modeling protocols for the reconstruction of large-scale cellular signaling networks [1-8]. Besides the qualitative analysis of the signaling networks, mathematical approaches for the quantitative modeling and simulation of signal transduction pathways have also been developed [9-15]. Most of these quantitative signaling models are kinetic reactions represented by the assemblage of ordinary differential equations (ODE) [16]. The ODEs are designed to simulate dynamic changes throughout continuous but limited time points that define a function of rate changes for one independent variable and one or more of its derivatives with respect to that variable. Given the determined kinetic parameters, such ODE models provide a "forward-engineering" framework to simulate the spatiotemporal dynamics of the system. The time profiles of the target transcription factor activation in response to various stimuli can be obtained by this established approach.

In gene regulatory network modeling, some wet-bench experimental approaches have been used to detect the gene expression and transcriptional activities. Microarray technology is a powerful high-throughput technique enabling biologists to simultaneously measure the expression profile of tens of thousands of genes under prescribed conditions [17,18]. As for transcriptional activities, the electrophoretic mobility shift assay (EMSA) [19,20] is an affinity electrophoresis technique that can determine single protein or protein-DNA complex binding activity at a time. However, the ability to broadly assess the activities of transcription factors is still much limited. Therefore, most computational transcription activity and gene regulatory network are reconstructed from genome wide gene expression data.

Such a data-driven approach to gene expression analysis provides systematic information of underlying gene regulatory systems and offers the possibility to infer the dynamics and mechanisms of transcription control by reverse engineering [21-28]. There are two kinds of reverse engineering strategies for modeling gene regulatory networks based on DNA microarray data, namely the "influence" and "physical" approaches [29]. The "influence approach" produces the genetic network that illustrates regulatory influences between RNA transcripts. This strategy can integrate information pertaining to the relationships between regulated genes, protein-protein interactions, and enzyme catalysis to establish network based on transcript profiling data when the expression of certain transcripts is highly correlated. However, influence models are difficult to interpret in the context of

location and modification in the cell. The second strategy, known as the "physical approach", seeks to construct a physical interaction model between transcription factors and gene promoters. The transcriptional activities can be often predicted from gene expression data by a sigmoid function [30]. Moreover, factor analysis is another methodology to construct a physical regulation model which is represented as bipartite graph with transcription factors in the first layer and regulated genes in the second layer for reducing the dimensionality of the reverse engineering problem. In previous studies, principal component analysis (PCA) [31], independent component analysis (ICA) [32], and network component analysis (NCA) [33] have been applied to reconstruct transcription factor activities using gene expression profiles. PCA and ICA are traditional dimensionality reduction technologies. The transcription factor activity reconstructed by PCA and ICA is constrained, respectively, to be mutually orthogonal and statistical independent. These statistical assumptions do not match the real biological systems. However, NCA contrasts with traditional PCA and ICA in that it does not make any aforementioned statistical assumptions.

NF- κ B is a transcription factor that has long been recognized as the "master switch" in regulating the expression of various cytokines and host response effectors, as well as a wide array of genes to control inflammation, cell survival, apoptosis, and immune defense responses [34]. NF- κ B signaling can be initiated from membrane receptors, such as Toll-like receptors, tumor necrosis factor alpha (TNF- α) receptors, and interleukin-1 (IL-1) receptors, either individually or synergistically. In the past few decades, many studies have tried to resolve the complex NF- κ B dependent protein-protein and DNA-protein interactions, and significant progress has since been made on modeling signal transduction pathways and gene regulatory networks of the inflammatory response based on both biochemical and microarray data [35-40]. However, a systemic and dynamic view of how external stimuli evoke NF- κ B-dependent signal transduction activities to the downstream gene expression still remained unclear.

To overcome this challenge, we proposed a new computational modeling approach by connecting transcription activities derived from the reverse engineering of gene expression profiling to the transcription activities simulated from a forward engineering signaling model, using the NF- κ B signaling pathway with the corresponding gene responses as the case study. In this work, the NCA model was applied to reconstruct the regulatory activity of NF- κ B using gene expression profile data obtained in response to specific external stimuli [41]. A kinetic model was used to simulate the IKK-NF- κ B signaling pathway [42]. By mapping the NF- κ B profiles gen-

erated from the reverse engineering gene expression profiling data and the ones from the forward simulation, the bridging IKK activities induced by external stimuli were inferred. Features of this inferred signaling process were then confirmed by independent experiments using similar stimuli. This strategy successfully linked the initial signaling pathway with the relevant gene regulatory network. It also successfully inferred and distinguished the corresponding stimuli from gene responses under different inflammation conditions. Taken together, the strategy discussed in the present study can help enhance our understanding of inflammatory responses during the infection process; it is also applicable to other cellular processes.

Results and Discussion

The novel approach presented in this study combines forward engineering and reverse engineering strategies to infer the relationship between an external signal and its resulting genomic responses within a cell. This novel strategy can be divided into three parts, as illustrated in Figure 1. The first two parts are used to produce the regulatory profile of the transcription factor from the signal transduction simulation and the subsequent gene response, respectively. Using temporal gene expression profiles and known transcriptional regulation relationships, a reverse engineering approach is used to reconstruct the activity profiles of transcription factors with the relative influence strength of the regulatory relationships, such as the binding affinity, between transcription factors and their target genes in the regulatory network. Then, the forward simulation model, which describes the signal transduction from the external stimulus to the specific transcription factor, is constructed for simulating the regulatory profiles of transcription factors that are associated with different stimuli. Finally, the forward-simulated transcription factor profiles are used to match with the dynamic activity pattern of transcription factors that are obtained from the reverse engineering. In the following subsections, we demonstrated an example that links the NF- κ B signaling pathway with its consequential gene responses under different inflammatory stimuli. Thus, we demonstrated the feasibility and strength of this integrative approach to reconcile genetic responses induced by different stimuli into a gene regulatory model.

The reconstruction of the transcriptional activity of NF- κ B by NCA

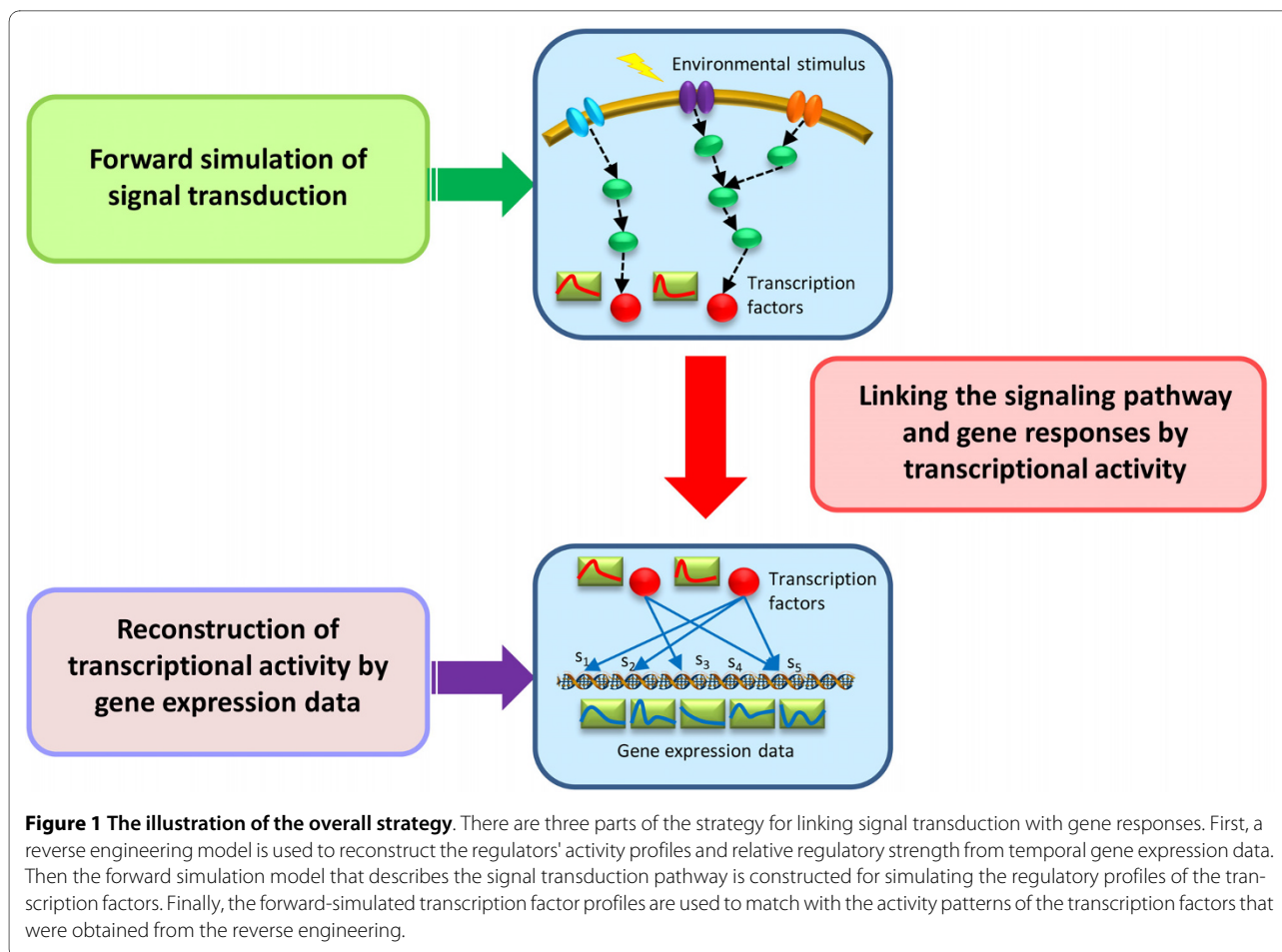
The Network Component Analysis [33] is a natural model with a reverse-engineering approach that has been used to deduce transcription factor activities and regulatory control strengths using temporal gene expression data and transcriptional regulatory relationships. The NCA framework regards the gene regulatory network as

a bipartite network model involving transcription factors and their target genes. The details of the NCA model are described in the Methods.

Lipopolysaccharide (LPS) is a kind of bacterial endotoxin widely used to stimulate inflammatory responses *in vitro*. The gene expression datasets in which cell cultures were treated with different doses of LPS from Boldrick et al. [41] were used. Briefly, Boldrick et al. added LPS to human peripheral blood mononuclear cells (PBMCs) at concentrations of 0.01 μ g/ml, 0.1 μ g/ml, and 1 μ g/ml. The transcript profiles of the PBMCs were determined by microarray and evaluated before the infection and at 0.5, 2, 4, 6, and 12 hours after infection.

The transcriptional regulatory module of NF- κ B was obtained from the TRANSFAC database [43] and MetaCore™ analytical suite (GeneGo Inc., St. Joseph, MI, USA). Both the TRANSFAC and MetaCore™ provide knowledge-based information, and all of the deposited data were collected from manually curated peer-reviewed literature. The TRANSFAC database contains expertly curated data on transcription factors, their experimentally proven binding sites, and regulated genes. MetaCore™ is an integrated database and software suite for pathway analyses of experimental data and gene annotation. Because this study was specifically aimed at simulating NF- κ B activity, which represents a combination of dimerized proteins and cannot easily be quantitatively determined by experimental approaches, we hence focused on the primary transcriptional mediator, NF- κ B/RelA-p50, of the canonical NF- κ B signaling pathway only. RelA and p50 are the major and most common constituents of the NF- κ B dimer complex in most cell types [44]. Either RelA or p50 can also form homodimers to regulate gene expression. Therefore, based on the TRANSFAC and MetaCore™ data, a transcription regulatory network of NF- κ B/RelA-p50 consisting of a total of 87 target genes with 125 direct transcription regulatory relationships was generated. However, the target genes of the NF- κ B/RelA-p50 transcription factor that were not expressed significantly in the LPS-induced gene profiles were filtered. Finally, as shown in Figure 2, the regulatory network was reduced to 54 target genes, with 77 transcription-regulatory interactions. As expected, most of these genes, including IL-1 β , IL6, IL8, TNF, CXCL1, and CXCL2, are related to pro-inflammatory functions.

From this processed dataset, the NCA approach decomposes the gene expression matrix [E] into an influence-strength matrix [S] and a transcription factor activities matrix [A]. In order to reconstruct the dose-dependant transcription factor activity profiles with the relative activation strength among different doses of LPS stimuli, we merged the three gene expression matrices with the LPS doses of 0.01 μ g/ml, 0.1 μ g/ml, and 1 μ g/ml into an expanded gene expressions matrix. Figure 3



shows the NCA-derived transcriptional activity profiles of NF- κ B under different dose of LPS (The activity and strength matrices is available as Additional file 1). All three of the NF- κ B profiles have an initial peak at two hours, but each NF- κ B activity pattern has a different activation intensity. The difference in activation intensities is due to the activation intensity depending on the strength of the different LPS stimuli. This activity pattern coincides with the experimentally validated result demonstrating that the NF- κ B activity peaks at approximately one to two hours after the induction by LPS [45].

It should be pointed out that even though the LPS dose changes over two orders of magnitude, the heights of the initial peaks of the reconstructed NF- κ B activities differ by less than 50%. On the other hand, the increasing LPS dose produced profiles with different features, as described below during the later stage after two hours. Figure 3 illustrates that the activation intensities of all the three NF- κ B profiles decrease after the peak at two hours. However, after four hours, the NF- κ B activity decreases more slowly (Figure 3A and 3B) in the conditions of both of the lower doses of LPS (0.01 μ g/ml, 0.1 μ g/ml), whereas NF- κ B maintains its activation intensity and results in a

shoulder shape (Figure 3C) at six hours under the high dose of LPS (1 μ g/ml). These patterns of reconstructed NF- κ B activity were used in identification of corresponding IKK activity profiles.

IKK-NF- κ B signaling simulation

The nuclear transcription factor NF- κ B is a protein complex that regulates numerous physiological processes [44,46] and plays a pivotal role in regulating the innate and adaptive immune responses [47-49]. Inactive NF- κ B is retained in the cytoplasm while bound to an inhibitory factor, I κ B. Upon stimulation by TNF- α , bacterial LPS or IL-1, NF- κ B can be released from I κ B and activated through a series of signal transduction pathways. Among these pathways, IKK proteins are the converging key regulators of NF- κ B signaling. For example, when bacterial pathogens present virulent factors into host cells, I κ B is phosphorylated by activated IKK and subsequently degraded by the ubiquitin-proteasome complex [50]. Upon I κ B degradation, the activated NF- κ B is then translocated into the nucleus where it regulates transcription from a large number of genes encoding components of the immune system, including a wide range of pro-

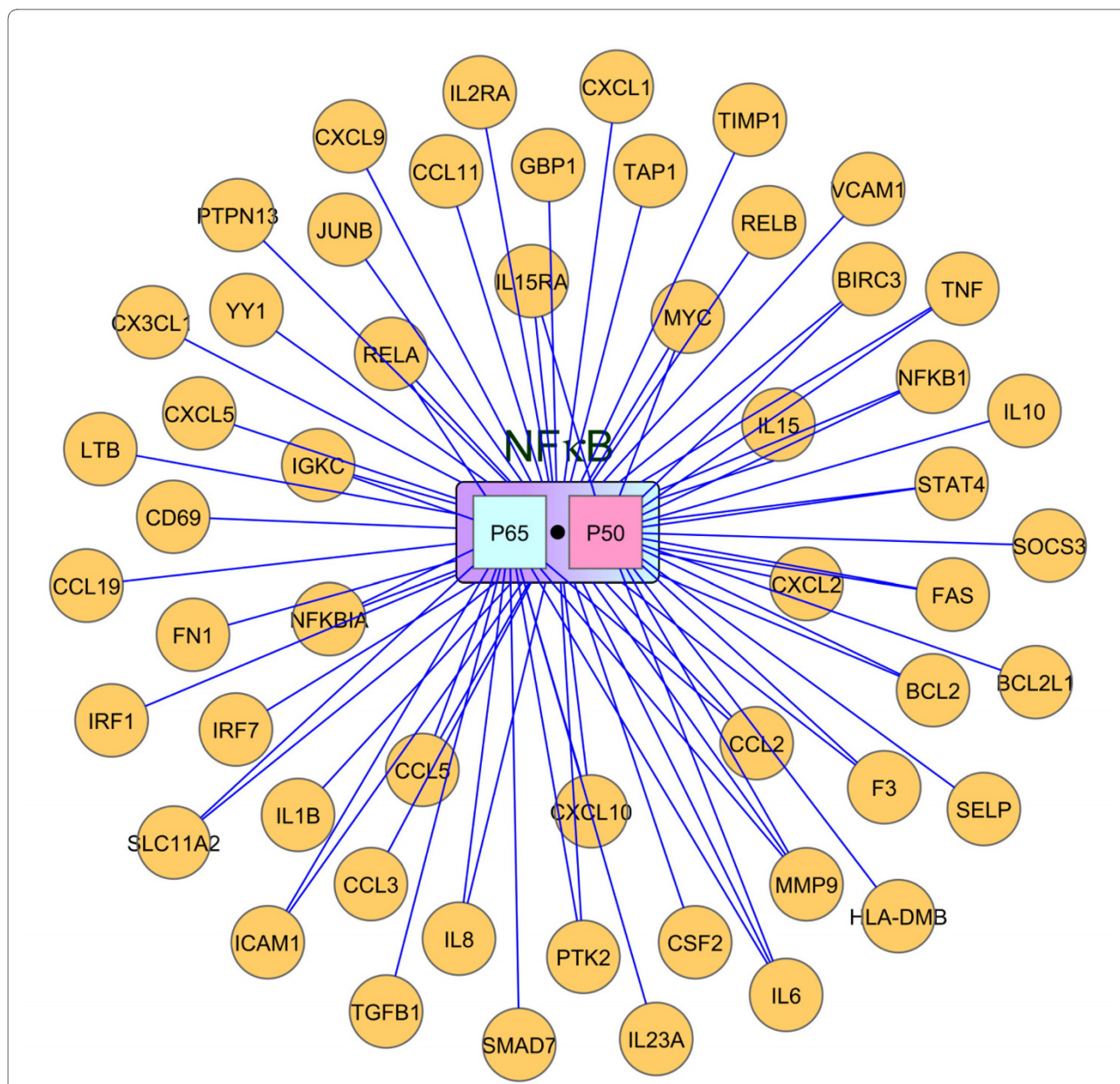
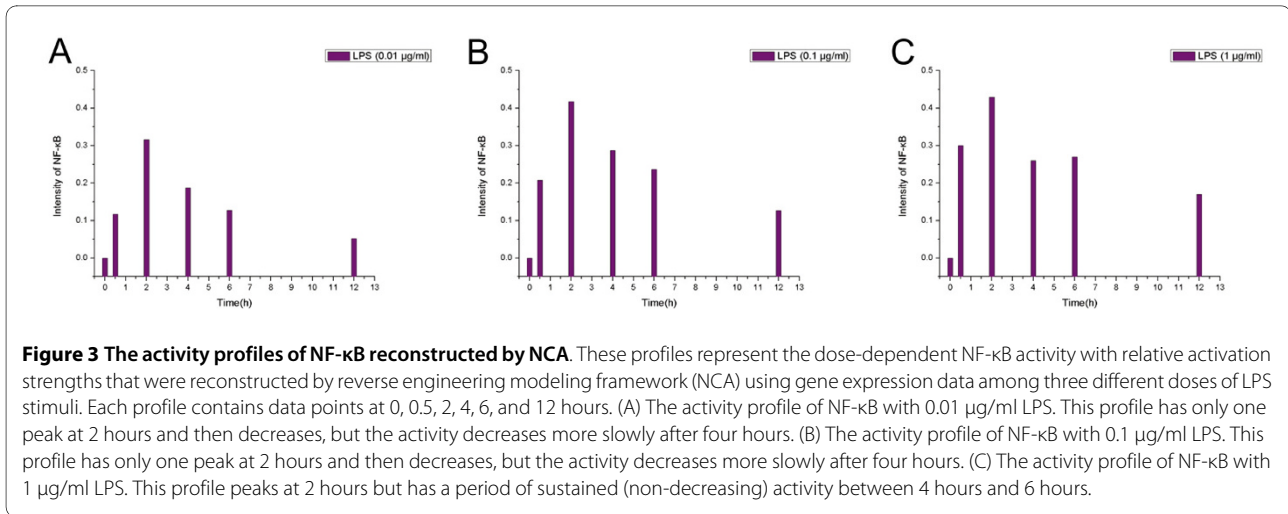


Figure 2 The target gene regulatory network, NF- κ B/RelA:P50. The most common NF- κ B in the canonical NF- κ B signaling pathway is a heterodimer that is composed of RelA and P50. RelA and P50 also can form homodimers and regulate gene expression. The regulatory relationships of NF- κ B/RelA:P50 were obtained from the TRAFAC database and the MetaCore™ analytical suite, which both are knowledge bases, and the data were collected from manually curated peer-reviewed literature. Based on these databases, we filtered out some target genes of NF- κ B/RelA:P50 that expressed insignificantly in the LPS-induced gene profiles. The regulatory network of NF- κ B/RelA:P50 finally consisted of 54 total target genes with 77 transcription regulatory interactions.

inflammatory cytokines, chemokines, adhesion molecules, and inducible enzymes [44]. Accordingly, a multitude of signals transmitted from various receptors in the cell results in IKK activation, thus initiating the IKK-I κ B-NF- κ B signaling cascade that regulates NF- κ B activity.

Previously, a computational model describing the inflammatory signal transduction dynamics in mammals

from IKK to NF- κ B has been proposed and experimentally validated by Hoffmann and colleagues [42,45,51,52]. Their model allows temporal simulations of the dynamic profile of NF- κ B in the nucleus upon in the input of different profiles of IKK activity. In this study, we adapted the Hoffmann model [45,51,52] and expressed it in ODEs representing the mass balances of 24 components involv-



ing 72 reactions. The Hybrid Functional Petri-Net (HFPN) [53,54] was then used to carry out simulations of the kinetic model. HFPN uses uncomplicated graph annotations to define kinetic functions. The architecture of the model are shown in the supplementary data, and a detailed explanation is provided in the Materials and Methods section.

The final outputs of the forward simulation model are the specific downstream transcription factor activity profiles, such as the NF-κB activity profile in this model. Figures 4A to 4C show three examples of NF-κB profiles generated from IKK profiles with different intensities and

activation durations. In all cases, the corresponding NF-κB profiles demonstrated high correlations with the IKK profiles, with small time delay. In Figure 4A, the intensity of the NF-κB activity also peaked with an acute pulse of IKK at approximately two hours. Moreover, as the IKK activity increased again after ten hours, the intensity of NF-κB activity also increased. Figure 4B shows that a smaller increase of IKK activity will produce weaker NF-κB activity. Figure 4C represents another situation in which the IKK activation intensity persists for several hours and then drops off. The corresponding NF-κB activity also lasted for almost the same duration.

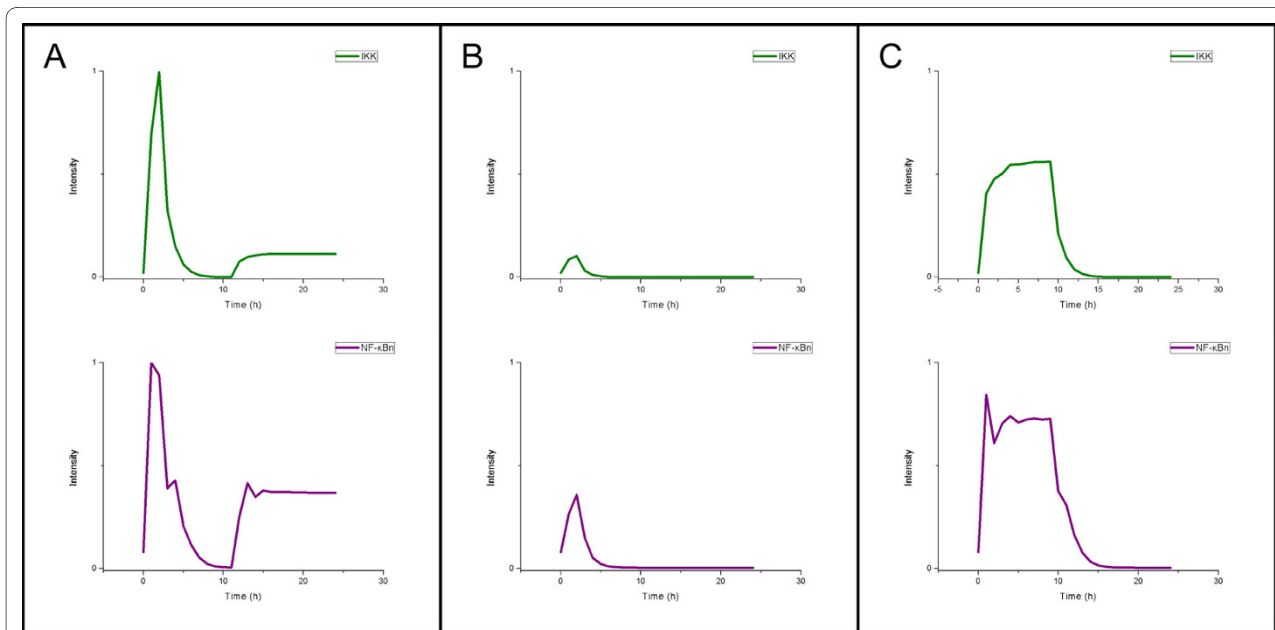


Figure 4 Three examples of NF-κB signaling pathway simulation. The NF-κB profile can be obtained by using various IKK profiles with different intensities and stimulation durations in the IKK-NF-κB signaling simulator. Three examples of IKK stimuli and the corresponding NF-κB profiles are shown. (A) The acute stimulus of IKK that has a peak at around two hours and a positive feedback control after ten hours. (B) A weak stimulus of IKK. (C) The IKK stimulus is not an impulse stimulation. It rises steeply during the first two hours and then increases slowly until ten hours.

Based on this simulation model of IKK-NF- κ B signaling, we can infer that IKK activity arises from the LPS stimuli by mapping the simulated NF- κ B profiles to those derived from the gene expression datasets. Moreover, because the simulated IKK profiles have consistent activity with the derived NF- κ B profiles, we would expect that the inferred IKK profiles would have similar features to the reconstructed NF- κ B profiles from the gene expression data.

Linking the NF- κ B signaling pathway to the corresponding changes in gene expression

The computational model of the forward simulation for the NF- κ B signaling pathway provides NF- κ B profiles under different IKK stimuli. Given the reconstructed NF- κ B profiles from the gene expression data and the computational model of the forward simulation under different IKK stimuli, we can generate IKK activity curves by matching the NCA-derived NF- κ B profiles and the forward simulated NF- κ B profiles initially. Then, the IKK activity curve that corresponds to the best-matched forward-simulated NF- κ B profile will be selected.

Figure 5 shows, at the specified time points (0, 0.5, 2, 4, 6, and 12 hours), the IKK profiles that can produce the simulated NF- κ B activity profiles best matched with those reconstructed from the gene expression by NCA. The patterns of two matched NF- κ B profiles are highly

correlated, with Pearson correlation coefficients of approximately 0.99. As expected, the IKK profile retains the key features of the NF- κ B profiles: initial peaks were observed within half an hour of the stimuli; the amplitude of the initial peak increased slightly with order-of-magnitude changes in the extracellular stimuli. It is worth noting that a second peak and signals are found for the highest dose at six hours. To demonstrate the predictive power of our integrative model, these candidates of the IKK activity profile were further validated by an experimental IKK Kinase Assay carried out under the same experimental conditions of the reported gene expression datasets, with stimuli of three different LPS doses.

Validations and model interpretation

In order to validate the computational results, the IKK Kinase Assay was performed individually with LPS concentrations of 0.01 μ g/ml, 0.1 μ g/ml, and 1 μ g/ml. Figure 6 indicates that the results of the IKK Kinase Assay are consistent with our IKK activity patterns derived from the computational prediction. Under the different stimuli strengths, all of the IKK profiles reached their peaks at 0.5 hours. Although there are ten-fold or hundred-fold differences in the LPS dose, the activity intensity of the IKK response did not show an increase of the same magnitude. The disparity between the stimulation intensity and the response activity is probably due to the limited num-

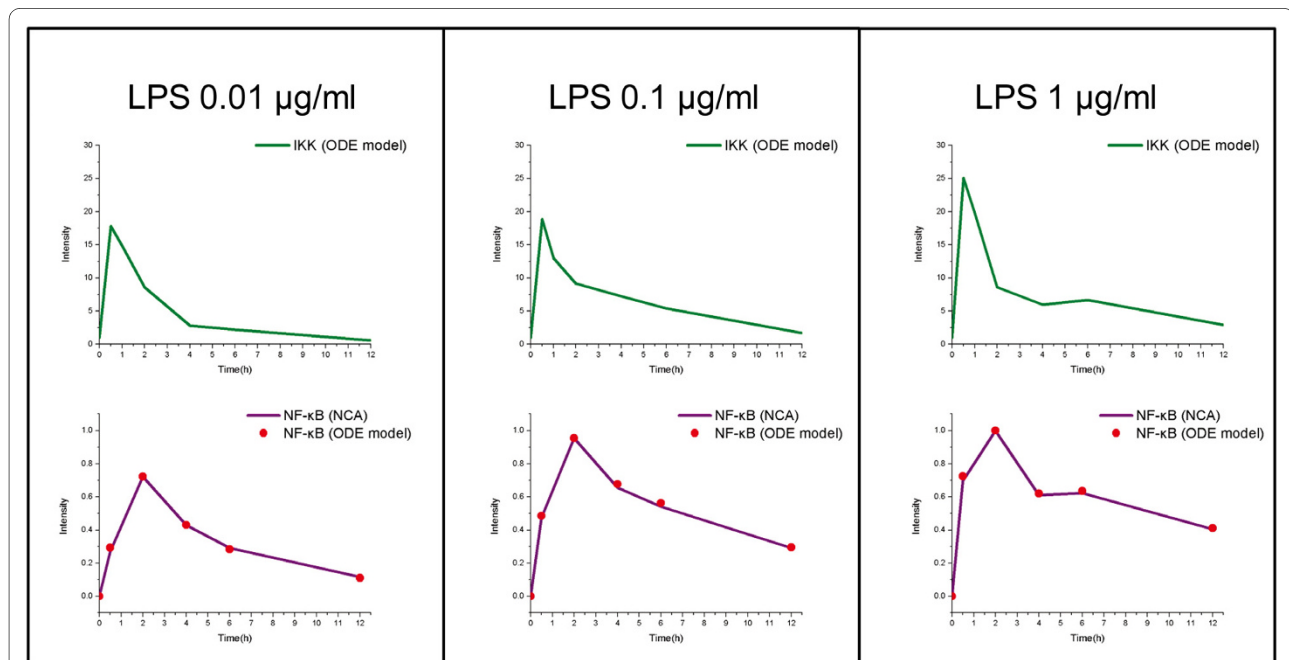


Figure 5 Linking the NF- κ B signaling pathway and gene responses by mapping the pattern of NF- κ B profiles. By simulating the NF- κ B signaling pathway, one of the NF- κ B profiles (pink data points) that is induced by a specified IKK profile (green line) matches the pattern of the NF- κ B profile (purple line) reconstructed from the gene expression data produced by NCA among the three different doses of LPS stimulus. In each conditions of LPS stimulus strength, the pattern of the two NF- κ B profiles can be mapped at the time points of 0, 0.5, 2, 4, 6, and 12 hours with a Pearson correlation of \sim 0.99.

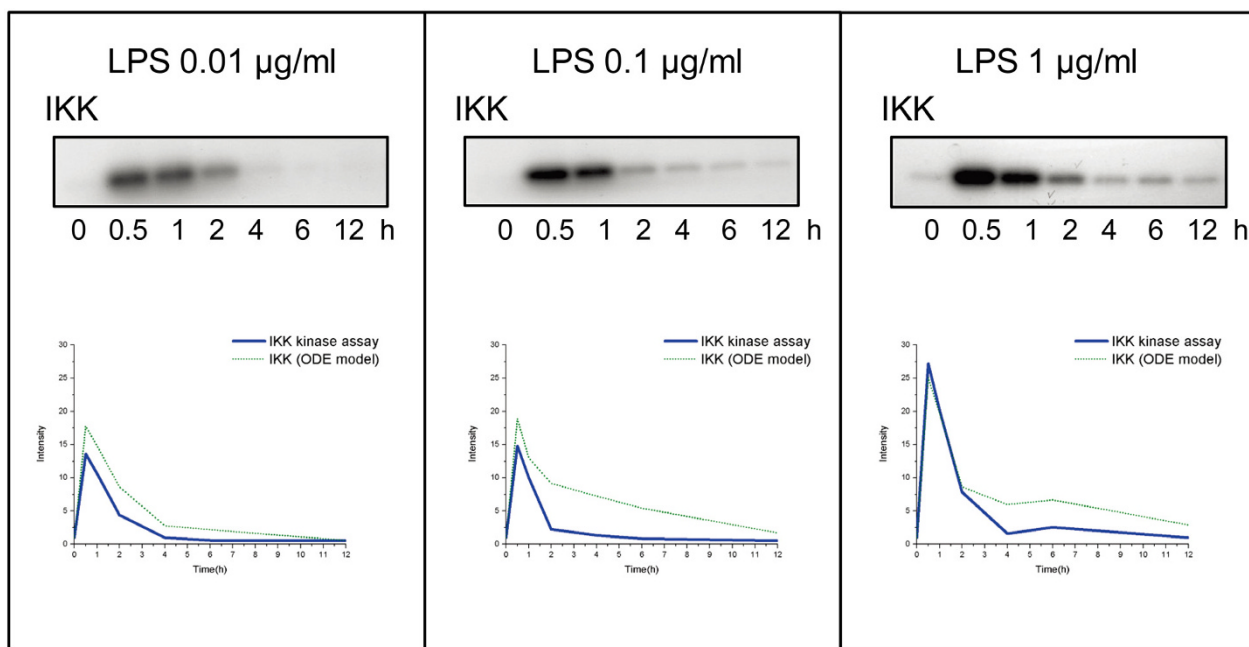


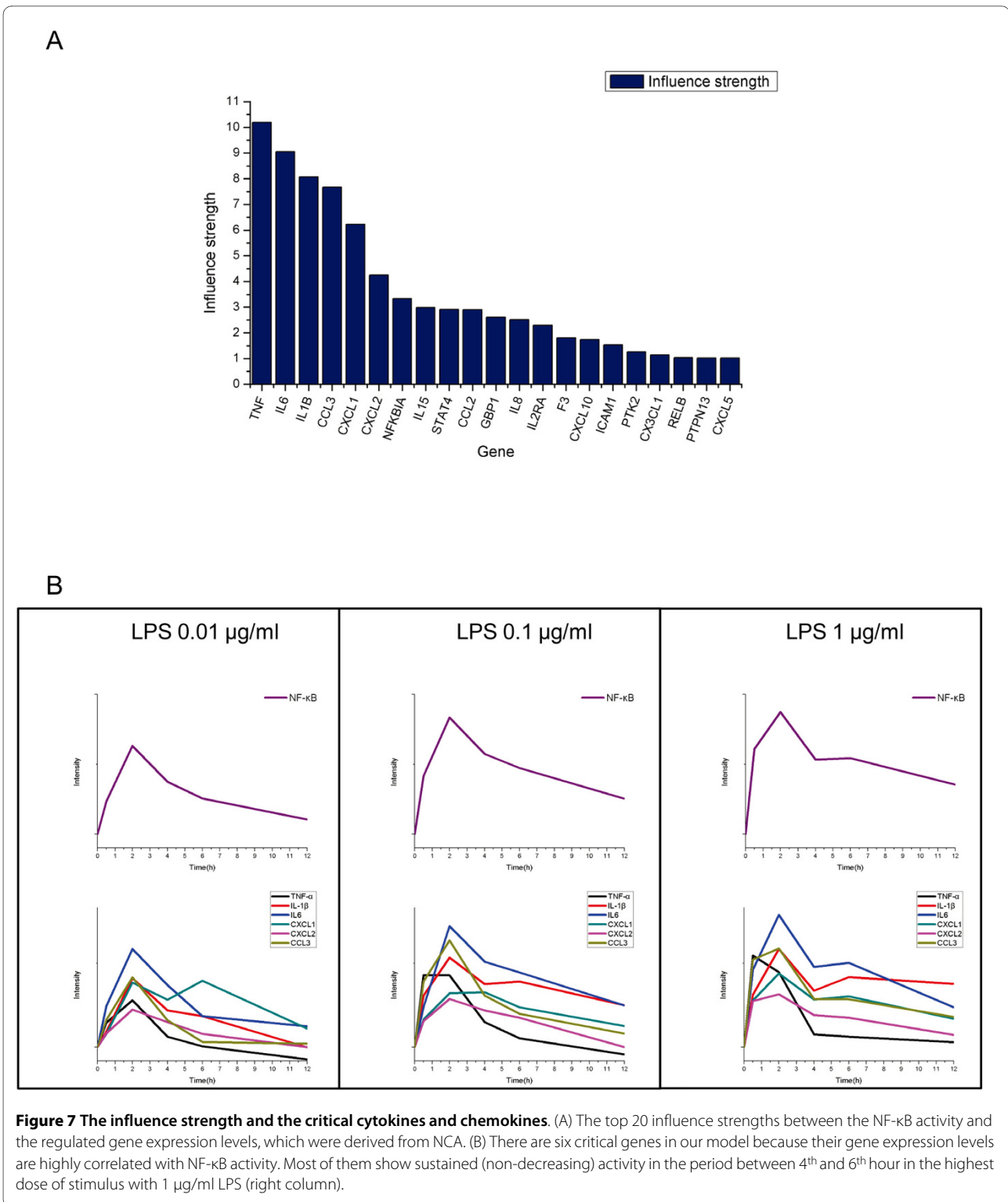
Figure 6 The experimental validation of the IKK Kinase Assay. The experimental validation of the IKK Kinase Assay under three different doses of LPS at the time points of 0, 0.5, 1, 2, 4, 6, and 12 hours. The blue lines are obtained by Image Q (GE Healthcare) that represents the western blot results of IKK activity. The green dotted lines represent the simulation results from forward engineering modeling framework (ODE model). The patterns of IKK activity from both the IKK Kinase Assay and the computational results are consistency. The ten-fold and hundred-fold dose of LPS appears not to trigger an extremely intense cellular response but induce periods of slow decay and non-decreasing inflammatory response between 4 to 6 hours respectively.

ber of receptors on the cell membrane or some damping mechanism of the signal transduction pathway. However, we observed that with the 1 µg/ml LPS stimulus, the IKK profile had a second wave at six hours. It is possible that large amounts of LPS cannot be completely consumed in the first wave of cellular reactions. Moreover, this disparity is also likely caused by positive feedback control the TNF-α and IL-1 autocrine signaling, to regulate cell proliferation and augment NF-κB activation [55,56]. Therefore, according to our model, the high dose of LPS appears not to trigger an extremely intense cellular response, but rather to induce a secondary reaction of an inflammatory response. This result is consistent with the hypothesis [57] that the innate response is not sufficient for a large number of invading microorganisms in an acute severe infection, which then results in a rapid multiplication of the microorganisms and induces the second response of the host. The circulating cytokines, such as TNF-α and IL-1, are thus produced overwhelmingly after some delay.

In addition, Figure 7A illustrates the influence strengths of the NF-κB target genes. Strong influence strength means that the activity of the transcription factor and the target gene expression are highly correlated. The top six genes include, besides TNF-α and IL-1, other cytokines and chemokines, namely IL-6, CXCL1, CXCL2,

and CCL3. All of these genes exhibit high correlations with the NF-κB profiling (Figure 7B). Interestingly, most of them have a second peak in the highest dose of stimulus with 1 µg/ml LPS. The Pearson correlation coefficients between them and NF-κB ranges from 0.78 to 0.96. Thus, the influence strength can be seen as the degree of the efficacy of a transcription factor to activate or inhibit one gene.

TNF-α and IL-1 are the main pro-inflammatory mediators of inflammation and tissue damage that are induced by NF-κB and associated with septic shock in acute severe infection. They play a synergistic role in orchestrating the inflammatory response and enhance the NF-κB activities by positive feedback loops. This positive effect of increasing TNFα and IL-1 on the NF-κB activity might explain the second wave of the NF-κB profile predicted by our integrative model in the presence of the highest LPS dose. Large amounts of TNF-α and IL-1 would result in a systemic inflammatory response syndrome (SIRS). This phenomenon leads to increased vascular permeability and the generation of thrombi in many small vessels; it also consumes massive amounts of blood-clotting proteins, thus deregulating hemostasis in the patient. The phenomenon also frequently results in the failure of several organs in the body; thus, septic shock often results in a high mortality rate [58,59].



IL-6 is also one of the most important mediators of the acute inflammatory response that is released in response to IL-1. IL-6 is responsible for the induction of fever; it stimulates energy mobilization, which then leads to an increased body temperature in the muscle and fatty tissue

[60,61]. Moreover, in acute injury, high-dose bacterial invasion would induce acute inflammatory responses. The aberrant secretion of pro-inflammatory cytokines alters fibrin deposition and degradation and results in the abnormal cell-cell adhesion formation. CXCL1, a

cytokine-induced neutrophil chemoattractant, has been revealed to be involved in cell-cell adhesion formation in sepsis [62], which is a serious challenge in acute infections and leads to high morbidity.

CXCL2 is also a critical chemokine for neutrophils. CXCL2 has been implicated in responses to both TNF- α and LPS via the NF- κ B-dependent pathway. During infection or sepsis, increased CXCL2 would recruit mucosal neutrophils and amplify the inflammatory response. This might give rise to intestinal injury [63]. Besides the cytokines and chemokines described above, the expression of CCL3 also correlates with NF- κ B activity, and NF- κ B also strongly influences CCL3 expression. CCL3, a β -chemokine, regulates the migration and recruitment of various effector cells, such as monocytes, neutrophils, and natural killer cells, into infected sites. Previous studies provide evidence that CCL3 is involved in the degranulation and recruitment of mast cells [64]. The key roles of mast cells in the innate immunity against acute inflammation include the release of TNF- α -histamine and other inflammatory mediators, as well as bacterial phagocytosis. CCL3 can activate neutrophils and enhance the phagocytic antibacterial functions of macrophages. CCL3, therefore, is considered to play a crucial role in mast cell-associated host resistance against sepsis or other acute inflammation.

Conclusions

In summary, we have linked a signaling pathway with the resulting gene responses by mapping the transcription factor activity (e.g., NF- κ B) inferred from forward and reverse engineering for three different stimuli strengths. The corresponding signal profiles (e.g., IKK) were successfully reconstructed by our proposed method. Furthermore, the genes that highly correlated with transcription factor activity and had relatively strong influence strengths were also identified to reflect the biological consequences. This method is a practical and effective solution for linking exogenous cellular stimuli with signal transduction dynamics and gene expression profiling. It provides an overall systematic view of the inflammation process. In the near future, this strategy may be applied to more systems, as genome-wide regulatory data (such as ChIP-chip, ChIP-PET) involving transcription factors in mammals become available. For example, some inflammatory and autoimmune diseases are caused by the aberrant regulation of NF- κ B, e.g., rheumatoid arthritis, inflammatory bowel disease, multiple sclerosis, atherosclerosis, and asthma [65,66]. Therefore, the application of our approach, which utilizes the dynamic simulation of a signaling pathway and reverse engineering, will be beneficial in the development of an in silico model for drug design and effective therapeutic strategies.

Methods

Network component analysis

In this work, a physical approach, the NCA proposed by Liao et al. [33], was used to reconstruct a model of the activity of transcription factors and the regulation strength from gene expression profiles [41]. The relationship between the activity of transcription factors and gene expression can be modeled by a log-linear combination. The gene expression level at each time point can be described as:

$$e_{ij} = \prod_{k=1}^L (A_{kj})^{s_{ik}}$$

where e_{ij} is the expression level of gene i at the j^{th} time point, A_{kj} is the activity intensity of transcription factor k at the j^{th} time point and s_{ik} is the influence strength of transcription factor k and gene i . Then we took the log of this equation, and the gene regulatory network is modeled as matrix representation:

$$[E]_{I \times J} = [S]_{I \times K} \bullet [A]_{K \times J}$$

The matrix $[E]_{I \times J}$ is the logarithm of the signals of the microarray data containing I genes and J time points. $[S]_{I \times K}$ represents the regulatory relationship and influence strength between K transcription factors and I regulated genes. $[A]_{K \times J}$ represents the activities of the K transcription factors at the J time points.

Given $[E]_{I \times J}$, $[S]_{I \times K}$ and $[A]_{K \times J}$ can be obtained using a bipartite decomposition technique that requires knowledge of the position of the nonzero elements of $[S]_{I \times K}$, i.e., the prior knowledge of the regulatory relationships between the transcription factors and genes. The regulatory information was obtained from the TRANSFAC database (version 11.2) and MetaCore™ analytical suite, as shown in Figure 2. All of the NF- κ B/RelA:P50-regulated genes in our network were expected to be significantly upregulated or downregulated.

Forward simulation by HFPN

HFPN is an extension of Petri Net that can handle continuous events of real numbers [53]. It has the advantage of an intuitive representation for modeling a dynamic biological system. There are two kinds of places and transition components in HFPN, discrete/continuous places and discrete/continuous transitions. Discrete places and transitions hold nonnegative integer number tokens as their content. A continuous place can contain nonnegative real numbers. A continuous transition fires continuously according to a given speed function and values in the places. In addition, three kinds of arcs are defined. A

normal arc makes the tokens in a place increase or decrease and triggers the next transition. An inhibitory arc is used to stop the action of transitions, and a test arc does not decrease tokens but triggers a connected transition.

In this study, the NF-κB signaling model was implemented by HFPN using the revised ODE model proposed by Hoffmann and colleagues [45,51,52]. This ODE model was modified from an earlier version [67] to include IKK time-course generator, nuclear-cytoplasmic volume ratio and other important factors as suggested by Lipniacki et al [68].

We created building blocks of HFPN according to a reaction categorized by the NF-κB signaling model, as shown in Figure S1a (Additional file 2). They include association, dissociation, nuclear import/export, and protein/RNA synthesis and degradation. The proteins are represented by continuous places, and the reactions are represented by continuous transitions. Based on these building blocks, the NF-κB signaling model was implemented in Cell Illustrator 3.0 [69] as shown in Figure S1b (Additional file 2).

We obtained NF-κB profiles by using different IKK profiles in this simulation model. To map the NF-κB profile from gene expression data, we developed an IKK generator, which generated the possible IKK profiles that had different peak times and intensities by adjusting the slope between each time point. The NF-κB profile was mapped in the order of time points at 0.5 h, 2 h, 4 h, 6 h, and 12 h.

Normalizing and mapping the pattern of transcription factor profiles

Based on the simulation model of the IKK-NF-κB signaling pathway, we generated IKK activity curves by tuning the various slopes between each time point as the simulation input. To map the pattern of transcription factor profiles obtained from the signaling simulation to those obtained using NCA, we normalized these profiles to the same scale of intensity. The transcription factor profiles that were derived from the NCA represented the relative intensity among all time points. Thus, these profiles were normalized to have a maximum intensity of 1.0 by the following equation:

$$\tilde{a}_{NF\kappa B,j}^{NCA} = \left(\frac{a_{NF\kappa B,j}^{NCA}}{a_{NF\kappa B,max}^{NCA}} \right)$$

$a_{NF\kappa B,j}^{NCA}$ is transcription factor activity obtained by the NCA at the j^{th} time point and $\tilde{a}_{NF\kappa B,j}^{NCA}$ is the relative transcription factor activity. On the other hand, we also normalized the transcription factor concentration profiles

$c_{NF\kappa B,j}^{STS}$ from the signal transduction simulation using the same method. However, the log of the value obtained was used as the transcriptional activity because gene expression data used in the NCA were recorded as a log ratio. The relative transcriptional activity is given by:

$$\tilde{a}_{NF\kappa B,j}^{STS} = \frac{\log_2 \left(\frac{c_{NF\kappa B,j}^{STS}}{c_{NF\kappa B,0}^{STS}} \right)}{\log_2 \left(\frac{c_{NF\kappa B,max}^{ST}}{c_{NF\kappa B,0}^{ST}} \right)}$$

After the transcription factor profiles from the NCA and signal simulation were normalized, we mapped them by calculating the Pearson correlation between the \tilde{a}_{kj}^{STS} obtained using different IKK stimuli profiles and the \tilde{a}_{kj}^{NCA} reconstructed from the gene expression to determine which IKK stimulus profile could induce the gene response. Further, the actual activity profile of a transcription factor can be obtained from the signaling simulation, and the connectivity strength can be derived from the NCA equation using the actual activity profile of the transcription factor.

IKK Kinase Assay

The stimulation was performed with the application of LPS at 1, 0.1, and 0.01 μg/ml to the THP-1 human monocytic cell line (American Type Culture Collection), which is used as a cell model system for PBMC [70]. At the indicated time points, before infection and at 0.5, 2, 4, 6, and 12 hours after infection, cells were harvested from the dish, washed twice by PBS, pelleted at 500 × g for 3 min, and snap-frozen in liquid nitrogen. The cell pellet was resuspended in lysis buffer containing 50 mM Tris-HCl, pH 7.6, 250 mM NaCl, 3 mM EDTA, 3 mM EGTA, 1% (v/v) Triton X-100, 0.5% (v/v) Igepal CA-630, 10% glycerol, 20 mM NaF, 40 mM β-glycero-3-phosphate, 2 mM DTT, 1 mM PMSE, 2 mM pNPP, 1 mM Na₃VO₄, and 10 μg/mL each of aprotinin, bestatin, leupeptin, and pepstatin.

The immuno-precipitation of the cell extract was performed using an anti-IKKγ Ab (BD). Protein A agarose beads were added to pull down the immuno-precipitated IKK complex. The immuno-pellets were washed and resuspended in 20 μl of kinase buffer containing 200 mM HEPES, pH 7.5, 100 mM MgCl₂, 5 μCi of [γ-32P]-ATP and 1 μg of GST-IκBα(1-54) substrate and incubated for 30 min at 30°C. The reactions were stopped by the addition of SDS-PAGE loading buffer, heated at 95°C for 5 minutes, and resolved on 10% SDS-PAGE gels by standard procedures. The bottom section of the gel contain-

ing the phosphorylated GST-I κ B α (1-54) as mounted, and the isotope intensity was detected by film. Proteins in upper section of the gel were transferred onto PVDF membrane (Millipore) and subjected to immunoblotting techniques to detect IKK α (Santa Cruz). The data were analysis by Image Q (GE Healthcare).

Additional material

Additional file 1 Activity and strength matrices. This file contains the activity matrices and strength matrices determined by NCA.

Additional file 2 The simulation model of the NF- κ B signaling pathway constructed by Cell Illustrator. This file contains Supplementary Figure S1 showing the simulation model of the NF- κ B signaling pathway constructed by Cell Illustrator. The simulation model of the NF- κ B signaling pathway was proposed by Hoffman et al. [45,51,52]. This model contains 24 components and 72 reactions and was rebuilt based on the Hybrid Functional Petri Net (HFPN) in Cell Illustrator 3.0. (A) The basic building blocks of HFPN according to the reaction category of the NF- κ B signaling model. (B) The full model of the NF- κ B signaling pathway that can yield the NF- κ B profile using the input IKK profile.

Abbreviations

NCA: network component analysis; TNF- α : tumor necrosis factor alpha; IL-1: interleukin-1; LPS: lipopolysaccharide; HFPN: Hybrid Functional Petri-Net; PCC: Pearson correlation coefficient.

Authors' contributions

SCP developed the method, performed the analyses and wrote the manuscript. DSHW advised on method design and wrote the manuscript. KCT and CCC performed the experiments, YYC and CHP performed the data analyses. YJC designed the experiments, interpreted the data and wrote the manuscript. CYT investigated the principle.

Acknowledgements

This study is supported by National Science Council, ROC, under grant NSC96-2627-B-007-007.

Author Details

¹Department of Computer Science, National Tsing Hua University, Hsinchu, 30013, Taiwan, ROC, ²Department of Chemical Engineering, National Tsing Hua University, Hsinchu, 30013, Taiwan, ROC, ³Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, 30013, Taiwan, ROC and ⁴Department of Resource Center for Clinical Research, Chang Gung Memorial Hospital, Taoyuan, 333, Taiwan, ROC

Received: 23 December 2009 Accepted: 8 June 2010

Published: 8 June 2010

References

- Bhalla US, Iyengar R: **Emergent properties of networks of biological signaling pathways.** *Science* 1999, **283**(5400):381-387.
- Cho KH, Wolkenhauer O: **Analysis and modelling of signal transduction pathways in systems biology.** *Biochem Soc Trans* 2003, **31**(Pt 6):1503-1509.
- Janes KA, Yaffe MB: **Data-driven modelling of signal-transduction networks.** *Nat Rev Mol Cell Biol* 2006, **7**(11):820-828.
- Levchenko A: **Dynamical and integrative cell signaling: challenges for the new biology.** *Biotechnol Bioeng* 2003, **84**(7):773-782.
- Liu Y, Zhao H: **A computational approach for ordering signal transduction pathway components from genomics and proteomics Data.** *BMC Bioinformatics* 2004, **5**:158.
- Neves SR, Iyengar R: **Modeling of signaling networks.** *Bioessays* 2002, **24**(12):1110-1117.
- Steffen M, Petti A, Aach J, D'Haeseleer P, Church G: **Automated modelling of signal transduction networks.** *BMC Bioinformatics* 2002, **3**:34.
- Suresh Babu CV, Joo Song E, Yoo YS: **Modeling and simulation in signal transduction pathways: a systems biology approach.** *Biochimie* 2006, **88**(3-4):277-283.
- Forsten-Williams K, Chua CC, Nugent MA: **The kinetics of FGF-2 binding to heparan sulfate proteoglycans and MAP kinase signaling.** *J Theor Biol* 2005, **233**(4):483-499.
- Goldstein B, Faeder JR, Hlavacek WS: **Mathematical and computational models of immune-receptor signalling.** *Nat Rev Immunol* 2004, **4**(6):445-456.
- Kitano H: **International alliances for quantitative modeling in systems biology.** *Mol Syst Biol* 2005, **1**:2005 0007.
- Meng TC, Somani S, Dhar P: **Modeling and simulation of biological systems with stochasticity.** *Silico Biol* 2004, **4**(3):293-309.
- Snoep JL: **The Silicon Cell initiative: working towards a detailed kinetic description at the cellular level.** *Curr Opin Biotechnol* 2005, **16**(3):336-343.
- Stucki JW, Simon HU: **Mathematical modeling of the regulation of caspase-3 activation and degradation.** *J Theor Biol* 2005, **234**(1):123-131.
- Yang CR, Shapiro BE, Hung SP, Mjolsness ED, Hatfield GW: **A mathematical model for the branched chain amino acid biosynthetic pathways of Escherichia coli K12.** *J Biol Chem* 2005, **280**(12):11224-11232.
- Cateau H, Tanaka S: **Kinetic analysis of multisite phosphorylation using analytical solutions to Michaelis-Menten equations.** *J Theor Biol* 2002, **217**(1):1-14.
- Chee M, Yang R, Hubbell E, Berno A, Huang XC, Stern D, Winkler J, Lockhart DJ, Morris MS, Fodor SP: **Accessing genetic information with high-density DNA arrays.** *Science* 1996, **274**(5287):610-614.
- Schena M, Shalon D, Davis RW, Brown PO: **Quantitative monitoring of gene expression patterns with a complementary DNA microarray.** *Science* 1995, **270**(5235):467-470.
- Garner MM, Revzin A: **A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system.** *Nucleic Acids Res* 1981, **9**(13):3047-3060.
- Fried M, Crothers DM: **Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis.** *Nucleic Acids Res* 1981, **9**(23):6505-6525.
- D'Haeseleer P, Liang S, Somogyi R: **Genetic network inference: from co-expression clustering to reverse engineering.** *Bioinformatics* 2000, **16**(8):707-726.
- Husmeier D: **Sensitivity and specificity of inferring genetic regulatory interactions from microarray experiments with dynamic Bayesian networks.** *Bioinformatics* 2003, **19**(17):2271-2282.
- Liang S, Fuhrman S, Somogyi R: **Reveal, a general reverse engineering algorithm for inference of genetic network architectures.** *Pac Symp Biocomput* 1998:18-29.
- Repsilber D, Liljenstrom H, Andersson SG: **Reverse engineering of regulatory networks: simulation studies on a genetic algorithm approach for ranking hypotheses.** *Biosystems* 2002, **66**(1-2):31-41.
- Soranzo N, Bianconi G, Altafini C: **Comparing association network algorithms for reverse engineering of large-scale gene regulatory networks: synthetic versus real data.** *Bioinformatics* 2007, **23**(13):1640-1647.
- Tegner J, Yeung MK, Hasty J, Collins JJ: **Reverse engineering gene networks: integrating genetic perturbations with dynamical modeling.** *Proc Natl Acad Sci USA* 2003, **100**(10):5944-5949.
- Wang Y, Joshi T, Zhang XS, Xu D, Chen L: **Inferring gene regulatory networks from multiple microarray datasets.** *Bioinformatics* 2006, **22**(19):2413-2420.
- Zak DE, Gonye GE, Schwaber JS, Doyle FJ: **Importance of input perturbations and stochastic gene expression in the reverse engineering of genetic regulatory networks: insights from an identifiability analysis of an in silico network.** *Genome Res* 2003, **13**(11):2396-2405.
- Gardner TS, Faith JJ: **Reverse-engineering transcription control networks.** *Phys Life Rev* 2005, **2**:65-88.
- Chen HC, Lee HC, Lin TY, Li WH, Chen BS: **Quantitative characterization of the transcriptional regulatory network in the yeast cell cycle.** *Bioinformatics* 2004, **20**(12):1914-1927.
- Raychaudhuri S, Stuart JM, Altman RB: **Principal components analysis to summarize microarray experiments: application to sporulation time series.** *Pac Symp Biocomput* 2000:455-466.

32. Liebermeister W: **Linear modes of gene expression determined by independent component analysis.** *Bioinformatics* 2002, **18**(1):51-60.
33. Liao JC, Boscolo R, Yang YL, Tran LM, Sabatti C, Roychowdhury VP: **Network component analysis: reconstruction of regulatory signals in biological systems.** *Proc Natl Acad Sci USA* 2003, **100**(26):15522-15527.
34. Hayden MS, West AP, Ghosh S: **NF-kappaB and the immune response.** *Oncogene* 2006, **25**(51):6758-6780.
35. Berardi M, Canonica GW: **The inflammatory network.** *Monaldi Arch Chest Dis* 2002, **57**(2):147.
36. Calvano SE, Xiao W, Richards DR, Felciano RM, Baker HV, Cho RJ, Chen RO, Brownstein BH, Cobb JP, Tschoeke SK, et al.: **A network-based analysis of systemic inflammation in humans.** *Nature* 2005, **437**(7061):1032-1037.
37. Hirsch E: **Signal transduction in inflammation. Perspective clues from the leukocyte-endothelium interface.** *Thromb Haemost* 2006, **95**(1):3-4.
38. Jayapal M, Tay HK, Reghunathan R, Zhi L, Chow KK, Rauff M, Melendez AJ: **Genome-wide gene expression profiling of human mast cells stimulated by IgE or FcepsilonRI-aggregation reveals a complex network of genes involved in inflammatory responses.** *BMC Genomics* 2006, **7**:210.
39. Kuwano K, Hara N: **Signal transduction pathways of apoptosis and inflammation induced by the tumor necrosis factor receptor family.** *Am J Respir Cell Mol Biol* 2000, **22**(2):147-149.
40. Saban R, D'Andrea MR, Andrade-Gordon P, Derian CK, Dozmozorov I, Ilnat MA, Hurst RE, Simpson C, Saban MR: **Regulatory network of inflammation downstream of proteinase-activated receptors.** *BMC Physiol* 2007, **7**:3.
41. Boldrick JC, Alizadeh AA, Diehn M, Dudoit S, Liu CL, Belcher CE, Botstein D, Staudt LM, Brown PO, Relman DA: **Stereotyped and specific gene expression programs in human innate immune responses to bacteria.** *Proc Natl Acad Sci USA* 2002, **99**(2):972-977.
42. Cheong R, Hoffmann A, Levchenko A: **Understanding NF-kappaB signaling via mathematical modeling.** *Mol Syst Biol* 2008, **4**:192.
43. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, et al.: **TRANSFAC and its module TRANSCOMP: transcriptional gene regulation in eukaryotes.** *Nucleic Acids Res* 2006:D108-110.
44. Moynagh PN: **The NF-kappaB pathway.** *J Cell Sci* 2005, **118**(Pt 20):4589-4592.
45. Werner SL, Barken D, Hoffmann A: **Stimulus specificity of gene expression programs determined by temporal control of IKK activity.** *Science* 2005, **309**(5742):1857-1861.
46. Gerondakis S, Grossmann M, Nakamura Y, Pohl T, Grumont R: **Genetic approaches in mice to understand Rel/NF-kappaB and IkappaB function: transgenics and knockouts.** *Oncogene* 1999, **18**(49):6888-6895.
47. Ghosh S, May MJ, Kopp EB: **NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses.** *Annu Rev Immunol* 1998, **16**:225-260.
48. Li Q, Verma IM: **NF-kappaB regulation in the immune system.** *Nat Rev Immunol* 2002, **2**(10):725-734.
49. Liou HC: **Regulation of the immune system by NF-kappaB and IkappaB.** *J Biochem Mol Biol* 2002, **35**(6):537-546.
50. Karin M, Delhase M: **The I kappa B kinase (IKK) and NF-kappa B: key elements of proinflammatory signalling.** *Semin Immunol* 2000, **12**(1):85-98.
51. Cheong R, Bergmann A, Werner SL, Regal J, Hoffmann A, Levchenko A: **Transient IkappaB kinase activity mediates temporal NF-kappaB dynamics in response to a wide range of tumor necrosis factor-alpha doses.** *J Biol Chem* 2006, **281**(5):2945-2950.
52. Kearns JD, Basak S, Werner SL, Huang CS, Hoffmann A: **Ikbepsilon provides negative feedback to control NF-kappaB oscillations, signaling dynamics, and inflammatory gene expression.** *J Cell Biol* 2006, **173**(5):659-664.
53. Matsuno H, Tanaka Y, Aoshima H, Doi A, Matsui M, Miyano S: **Biopathways representation and simulation on hybrid functional Petri net.** *Silico Biol* 2003, **3**(3):389-404.
54. Peng SC, Chang HM, Hsu DF, Tang CY: **Modeling Signal Transduction of Neural System by Hybrid Petri Net Representation.** In *Operations Research Proceedings: 2004*; Tilburg Springer; 2004:271-279.
55. Dunne A, O'Neill LA: **The interleukin-1 receptor/Toll-like receptor superfamily: signal transduction during inflammation and host defense.** *Sci STKE* 2003, **2003**(171):re3.
56. Guernon J, Chaussepied M, Sopp P, Lizundia R, Moreau MF, Blumen B, Werling D, Howard CJ, Langsley G: **A tumour necrosis factor alpha autocrine loop contributes to proliferation and nuclear factor-kappaB activation of Theileria parva-transformed B cells.** *Cell Microbiol* 2003, **5**(10):709-716.
57. Netea MG, van der Meer JW, van Deuren M, Kullberg BJ: **Proinflammatory cytokines and sepsis syndrome: not enough, or too much of a good thing?** *Trends Immunol* 2003, **24**(5):254-258.
58. Durocher A, Becq MC, Gosset P, Saulnier F, Lefebvre MC, Tonnel AB, Capron A, Wattel F: **[TNF and sepsis].** *Ann Med Interne (Paris)* 1991, **142**(2):91-94.
59. Dinarello CA: **Proinflammatory and anti-inflammatory cytokines as mediators in the pathogenesis of septic shock.** *Chest* 1997, **112**(6 Suppl):321S-329S.
60. Makino T, Noguchi Y, Yoshikawa T, Doi C, Nomura K: **Circulating interleukin 6 concentrations and insulin resistance in patients with cancer.** *Br J Surg* 1998, **85**(12):1658-1662.
61. Febbraio MA, Pedersen BK: **Contraction-induced myokine production and release: is skeletal muscle an endocrine organ?** *Exerc Sport Sci Rev* 2005, **33**(3):114-119.
62. Wagner TH, Drewry AM, Macmillan S, Dunne WM, Chang KC, Karl IE, Hotchkiss RS, Cobb JP: **Surviving sepsis: bcl-2 overexpression modulates splenocyte transcriptional responses in vivo.** *Am J Physiol Regul Integr Comp Physiol* 2007, **292**(4):R1751-1759.
63. De Plaen IG, Han XB, Liu X, Hsueh W, Ghosh S, May MJ: **Lipopolysaccharide induces CXCL2/macrophage inflammatory protein-2 gene expression in enterocytes via NF-kappaB activation: independence from endogenous TNF-alpha and platelet-activating factor.** *Immunology* 2006, **118**(2):153-163.
64. Takahashi H, Tashiro T, Miyazaki M, Kobayashi M, Pollard RB, Suzuki F: **An essential role of macrophage inflammatory protein 1alpha/CCL3 on the expression of host's innate immunities against infectious complications.** *J Leukoc Biol* 2002, **72**(6):1190-1197.
65. Tak PP, Firestein GS: **NF-kappaB: a key role in inflammatory diseases.** *J Clin Invest* 2001, **107**(1):7-11.
66. Yamamoto Y, Gaynor RB: **Therapeutic potential of inhibition of the NF-kappaB pathway in the treatment of inflammation and cancer.** *J Clin Invest* 2001, **107**(2):135-142.
67. Hoffmann A, Levchenko A, Scott ML, Baltimore D: **The IkappaB-NF-kappaB signaling module: temporal control and selective gene activation.** *Science* 2002, **298**(5596):1241-1245.
68. Lipniacki T, Paszek P, Brasier AR, Luxon B, Kimmel M: **Mathematical model of NF-kappaB regulatory module.** *J Theor Biol* 2004, **228**(2):195-215.
69. Nagasaki M, Doi A, Matsuno H, Miyano S: **Genomic Object Net: I. A platform for modelling and simulating biopathways.** *Appl Bioinformatics* 2003, **2**(3):181-184.
70. Sharif O, Bolshakov VN, Raines S, Newham P, Perkins ND: **Transcriptional profiling of the LPS induced NF-kappaB response in macrophages.** *BMC Immunol* 2007, **8**:1.

doi: 10.1186/1471-2105-11-308

Cite this article as: Peng et al., Computational modeling with forward and reverse engineering links signaling network and genomic regulatory responses: NF-?B signaling-induced gene expression responses in inflammation *BMC Bioinformatics* 2010, **11**:308