

## Innate Immunity Interactome Dynamics

Asmaa Elzawahry<sup>1,2</sup>, Ashwini Patil<sup>1</sup>, Yutaro Kumagai<sup>3,4</sup>, Yutaka Suzuki<sup>2</sup> and Kenta Nakai<sup>1,2</sup>

<sup>1</sup>Institute of Medical Science, The University of Tokyo, Minato-ku, Tokyo, Japan. <sup>2</sup>Graduate school of Frontier Sciences, The University of Tokyo, Kashiwa-shi, Chiba, Japan. <sup>3</sup>Laboratory of Host Defence, WPI Immunology Frontier Research Center, Osaka University, Suita, Osaka, Japan. <sup>4</sup>Global Center of Excellence Program, Frontier Biomedical Science Underlying Organelle Network Biology, Osaka University, Suita, Osaka, Japan.

**ABSTRACT:** Innate immune response involves protein–protein interactions, deoxyribonucleic acid (DNA)–protein interactions and signaling cascades. So far, thousands of protein–protein interactions have been curated as a static interaction map. However, protein–protein interactions involved in innate immune response are dynamic. We recorded the dynamics in the interactome during innate immune response by combining gene expression data of lipopolysaccharide (LPS)-stimulated dendritic cells with protein–protein interactions data. We identified the differences in interactome during innate immune response by constructing differential networks and identifying protein modules, which were up-/down-regulated at each stage during the innate immune response. For each protein complex, we identified enriched biological processes and pathways. In addition, we identified core interactions that are conserved throughout the innate immune response and their enriched gene ontology terms and pathways. We defined two novel measures to assess the differences between network maps at different time points. We found that the protein interaction network at 1 hour after LPS stimulation has the highest interactions protein ratio, which indicates a role for proteins with large number of interactions in innate immune response. A pairwise differential matrix allows for the global visualization of the differences between different networks. We investigated the toll-like receptor subnetwork and found that S100A8 is down-regulated in dendritic cells after LPS stimulation. Identified protein complexes have a crucial role not only in innate immunity, but also in circadian rhythms, pathways involved in cancer, and p53 pathways. The study confirmed previous work that reported a strong correlation between cancer and immunity.

**KEYWORDS:** innate immunity, protein–protein interactions, gene expression, differential networks, interactome dynamics

**CITATION:** Elzawahry et al. Innate Immunity Interactome Dynamics. *Gene Regulation and Systems Biology* 2014;8 1–15 doi: 10.4137/GRSB.S12850.

**TYPE:** Original Research

**FUNDING:** AE is supported by MEXT. This research is supported the Cabinet Office, Government of Japan, and the Japan Society for the Promotion of Science ( JSPS ) through the Funding Program for World-Leading Innovative R&D on science and Technology (FIRST program).

**COMPETING INTERESTS:** Author(s) declare no competing interests.

**COPYRIGHT:** © the authors, publisher and licensee Libertas Academica Limited. This is an open-access article distributed under the terms of the Creative Commons CC-BY-NC 3.0 License.

**CORRESPONDENCE:** [knakai@ims.u-tokyo.ac.jp](mailto:knakai@ims.u-tokyo.ac.jp)

## Background

The immune system of a host defends it against invading pathogens, and it has two main subsystems: innate immunity and acquired immunity.<sup>1</sup> To initiate immune responses, it is crucial that the immune system can recognize pathogens. Recognition of pathogens is mediated by pattern-recognition receptors (PRRs) that recognize pathogen-associated molecular patterns (PAMPs), which are molecular structures that are common in pathogens.<sup>2,3</sup> Once a PAMP has been recognized, PRRs initiate signaling cascades to achieve the first line of defense toward invading pathogens.<sup>3</sup> A group of major players in innate immunity response are toll-like receptors (TLRs), which are transmembrane proteins<sup>3</sup> that are specialized to recognize pathogens.<sup>4,5</sup> The specificity of TLRs is based on

PRRs, which recognize different PAMPs<sup>6</sup> such as nucleic acids, lipids, lipoproteins, and proteins from different pathogens such as viruses, fungi, bacteria, and parasites.<sup>6,7</sup> One of the TLR member proteins is TLR4, which can be stimulated by many PAMPs, including lipopolysaccharide (LPS),<sup>8</sup> an essential part of the structure of the outer membrane of gram-negative bacteria.

Once TLR4 recognizes LPS, it initiates signaling cascades of immune responses.<sup>8,9</sup> The TLR signaling pathways can be divided into MyD88-dependent pathways and TRIF-dependent pathways.<sup>7</sup> In TLR members, TLR4 is unique in<sup>7</sup> that it is the only TLR member that uses both MyD88- and TRIF-dependent pathways.



Many cell types play key roles in the immune response. Since dendritic cells (DCs) are among the key players involved in the immune response, we used them in this study. DCs bridge innate immune responses to acquired immune responses. They utilize various innate immune receptors such as TLRs.<sup>10</sup> TLR4 stimulation by LPS induces the production of various cytokines and these, in turn, stimulate the cells and induce downstream-signaling pathways.<sup>7,11</sup> Thus, understanding the signaling pathways invoked by LPS stimulation will improve our understanding of these pathways.

Advanced high-throughput technologies provided genome-wide information such as transcriptome data, which is essential for studying biological systems. Protein-protein interactions (PPIs) are crucial for every biological process.<sup>12</sup> All protein interactions that exist in a specific organism form an interactome. Studying biological process in the context of the interactome provides a system-wide understanding of a given process.<sup>12</sup> Protein interaction networks are represented by nodes and edges where nodes represent proteins, while edges represent interactions between proteins.<sup>13</sup> Though the innate immune response has been extensively studied, key questions remain. For example, it is not known how protein interaction networks differ during innate immune responses, what the important protein complexes at different stages during innate immune response are, and which protein interactions are conserved during innate immune responses.

In this study, we attempt to address these issues by combining PPIs with gene expression data. Combining PPIs and expression data has been done before.<sup>14,15</sup> A previous study examined the dynamic structure of a human protein interactome to determine if changes in an interactome organization could be used to predict patient outcomes.<sup>15</sup> Another study has applied a network-based approach to identify markers for breast cancer that are correlated with metastasis.<sup>14</sup> Using a network-based approach in which expression data was combined with PPIs data, breast cancer markers were identified as protein interaction subnetworks. Genes with known mutations in breast cancer were not detected by typical differential expression analysis, but they did connect many differentially expressed genes, indicating that they do play an important role.

Taken together, these studies indicated that using PPIs in combination with gene expression data leads to better results than using either expression or PPI data alone. In this study, we combined PPIs with time-series gene expression data to study interactome dynamics in order to identify interactome differences across different stages of the innate immune response.

## Methods

**Cell culture.** Bone marrow cells were prepared from C57BL6/J mice. The cells were seeded into a 24-well plate at a concentration of  $1 \times 10^6$  cells/mL, and they were cultured in Roswell Park Memorial Institute (RPMI) medium

containing 10% fetal bovine serum and 10 ng/mL of granulocyte-macrophage colony-stimulating factor purchased from PeproTech (PeproTech, Rocky Hill, NJ, USA) for 6 days, and the culture medium was changed every 2 days. Six days later, floating cells were major histocompatibility complex (MHC) class2+CD11c+, which is a molecular hallmark of DCs.<sup>16</sup> The cells were collected and suspended in RPMI containing 10% fetal calf serum. LPS (Re595, purchased from Sigma-Aldrich, St Louis, MO, USA) were added at the concentration of 100 ng/mL and at 0 hours, 0.5 hours, 1 hour, 2 hours, 3 hours, 4 hours, 6 hours, 8 hours, 16 hours, and 24 hours later, the cells were collected and lysed in TRIzol (Life Technologies, Carlsbad, CA, USA). Ribonucleic acid (RNA) extraction was performed according to the manufacturer's instruction. Cells were subjected to TSS-Seq analysis as previously described,<sup>17</sup> where the TSS-Seq tag library was constructed by combination of the oligo-capping method with massively parallel sequencing.<sup>17</sup>

**TSS-Seq data.** TSS-Seq data generated in this study were mapped to mouse genome mm9<sup>18</sup> using Bowtie.<sup>19</sup> We constructed a gene model that is slightly different from that used by Tsuchihara et al<sup>17</sup> in that for each Ensembl gene,<sup>17</sup> if the intergenic distance between a gene and its preceding gene is greater than or equal to 50 kb, we counted tags upstream of the TSS up to 50 kb. If the distance is less than 50 kb, the tags were counted upstream of TSS with distance equal to the intergenic region. Biomart<sup>20</sup> was used to import gene features. Tag count was done based on the gene model. Counts were smoothed by add one smoothing; then filtered, normalized, and differentially expressed genes were identified based on a *P*-value of likelihood ratio test after fitting to a generalized linear model using edgeR.<sup>21</sup> Only genes with a *P*-value <0.001 were regarded as significantly differentially expressed genes. We selected the top 100 differentially expressed genes and confirmed their validity using Gene Ontology (GO) enrichment analysis.<sup>22</sup> The genes were enriched for terms related to the innate immune response.

GO enrichment analysis for the top 100 differentially expressed genes showed high enrichment for the GO term "immune responses", which indicates the importance of differentially expressed genes in the immune response.

**Protein-protein interaction data.** We prepared a set of unique mouse interactions from publicly available protein interactions. PPIs were collected from the PPIs databases, InnateDB<sup>23</sup> and HitPredict.<sup>24</sup> Network visualization and analysis was done using Cytoscape.<sup>25</sup> Cytoscape plugins used for the network analysis included MCODE,<sup>26</sup> and Network Analyzer.<sup>27</sup> Network edges were weighted by Pearson's correlation coefficient of gene co-expression.

**Time-specific networks.** We defined an interactome map for each time point during the innate immune response based on gene expression values at that time point. We constructed networks of proteins that are encoded by genes, which showed significant up-/down-regulation after stimulation for

each time point. In this study, we considered proteins encoded by the genes that show up-regulation by fourfold or more to be up-regulated proteins, and those encoded by genes that are down-regulated by four fold or more were regarded as down-regulated proteins. We constructed nine time-specific networks.

**Differential networks.** From the time-specific networks, we constructed differential networks between two successive time points such that differential network between time t1 and t2 only had interactions that were unique to the t1 network but not to the t2 network, as shown in Figure 1. We identified differences in the PPIs networks between two consecutive time points by constructing differential maps.

**Module identification and scoring – feature-weighted score.** Protein modules were identified based on the density of the protein complex using Cytoscape plugin, MCODE.<sup>25</sup>

Identified modules were scored using a novel scoring method to assess the significance of an identified module. The feature-weighted score is based on the following factors:

1. Module's density, as expressed by a clustering coefficient (MCODE score).
2. GO enrichment of the module nodes and whether the enrichment is statistically significant.
3. Proportion of differentially expressed genes in the module.
4. The average co-expression correlation coefficient of interactions in the module that reflect the co-expression of any two interacting partners.
5. Weights and the score calculation for each criteria are shown in Table 1. We calculated the weighted mean and called it a feature-weighted score and scored each identified module by the corresponding score.

The **feature weighted score** was calculated as follows:

$$\bar{S} = \frac{\sum_{i=1}^n WS}{\sum_{i=1}^n W} \quad (1)$$

Where S is the score of each criterion in a protein complex; W is the weight of each criterion; and N = 4 (as we included four criteria in the score).

We minimized errors to avoid spurious scores as follows:

1. Co-expression scores of self-loops were ignored;
2. If the average PCC for co-expression was negative, it was replaced with zero; and
3. If the assigned GO term was significant, the score = 1; otherwise, the score = 0.

The weight of each criterion was assigned based on the significance of each criterion. For example, we assigned co-expression a weight of 0.5, as co-expression is the most important factor in ranking the significance of a module. GO enrichment was assigned a weight of 0.1, as it is the least significant criterion, and a module with no significant GO enrichment may still play an important role in immunity. Modules were classified into highly significant, significant, or less significant modules based on their respective scores. Identified modules were compared with protein complexes in the CORUM database, a database of protein complexes in mammals.<sup>28</sup>

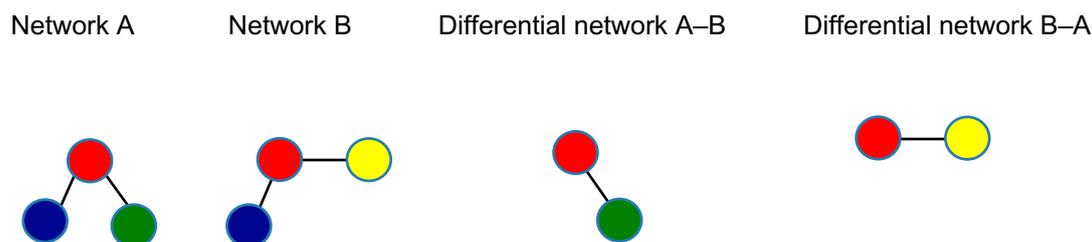
**Interactions protein ratio (IPR).** The IPR for a network is defined as the ratio of the total number of interactions to the total number of proteins in the network.

IPR can be determined by the following equation:

$$IPR = \frac{\sum_{i=1}^M J}{\sum_{j=1}^N x} \quad (2)$$

It may be also expressed as:

$$IPR = \frac{\bar{D}}{2} \quad (3)$$



**Figure 1.** Differential networks.

Diagram explaining differential maps. Differential network (A–B) has interactions specific to network (A) and not to network (B). Differential network (B–A) has interactions specific to network (B) and not to network (A).

**Table 1.** Criteria of feature weighted score and their weights.

CRITERION	WEIGHT
Coexpression	0.5
Differentially expressed genes	0.3
GO enrichment	0.1
Module's density	0.1

where

$$\bar{D} = \frac{\sum_{j=1}^N D}{N} \quad (4)$$

where  $J$  is an interaction in a network;  $x$  is a protein in a network;  $M$  is the total number of interactions in the subnetwork;  $N$  is the total number of proteins in the subnetwork;  $D$  is the degree of a protein in the network; and  $\bar{D}$  is the average degree of proteins in the network.

The node degree is defined as the number of interactions that the node has. The proof of Equation 3 is in Supplementary data.

To determine the significance of IPR, we constructed 1,000 random networks that were the same size as the time-specific networks, and we calculated the IPR for each randomized network. We compared the IPRs of the time-specific and randomized networks and calculated  $P$ -values as the ratio of the number of randomized networks that have same or higher IPR than that of a time-specific network over the total number of randomized networks.

**Pairwise differential matrix (PWDM).** PWDM is a matrix for normalized differences between each pair of networks specific for each time point. PWDM can be used to express differences in nodes or interactions between networks. Heatmaps and figures were made using R.<sup>29</sup>

**GO and KEGG enrichment.** All GO term and KEGG pathway<sup>30</sup> enrichment analyses were done using DAVID.<sup>31,32</sup>

## Results

**Interactome dynamics: global overview.** To reveal interactome dynamics during the innate immune response, we constructed a network of all unique PPIs from the databases InnateDB and HitPredict. The constructed network has 4,822 proteins with 10,549 interactions; we identified their expression patterns. In the interaction network, each protein was represented by a node where the node size represented the degree of interactions of the protein. To acquire a global view of interactome dynamics during the innate immune response, we identified the time point that had the largest number of interactions resulting from the smallest number of up-/down-regulated genes after stimulation with LPS. We considered

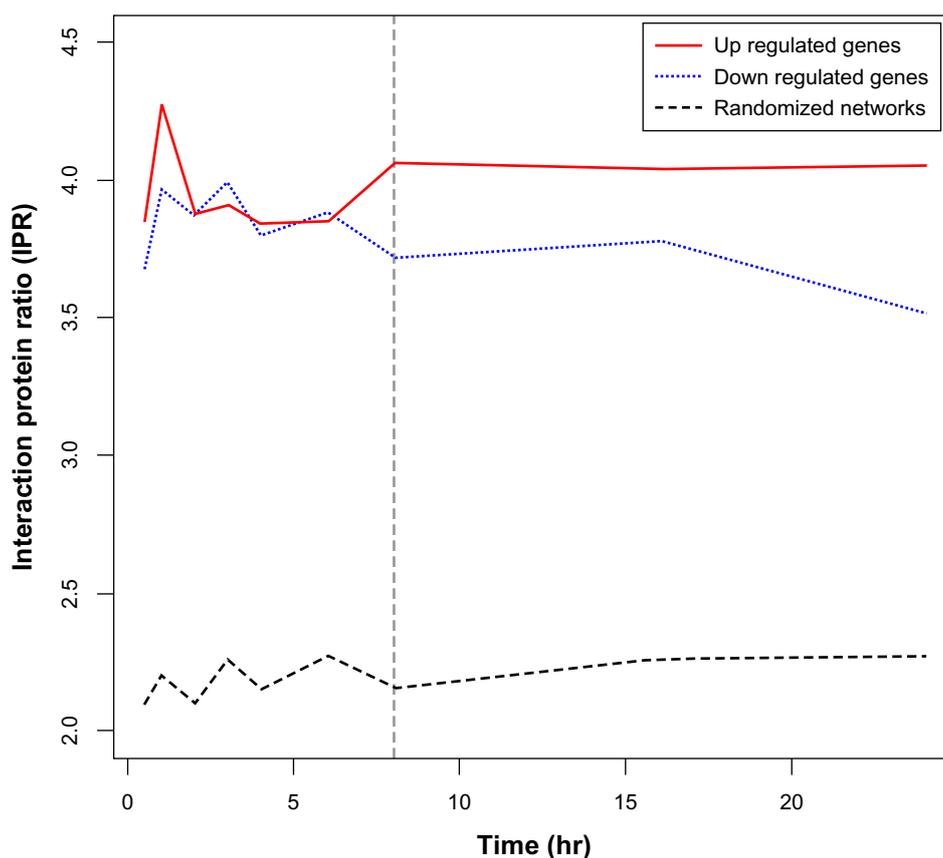
interactome dynamics from two perspectives: First, we identified protein complexes from a network of proteins that are encoded by genes that show differential expression during the whole time course of an innate immune response. These proteins and their interactions constitute a differentially expressed network. Second, we identified differences in the up-/down-regulation of an interactome between two consecutive time points during an innate immune response (differential networks), as well as core proteins/interactions that were conserved in an interactome during the innate immune responses. Protein modules were identified from each differential network and were scored as shown in Equation 1, as described in the Methods section. The modules were classified into highly significant, significant, and less significant modules based on the score of each module.

**Interaction protein ratio (IPR) during the innate immune response.** A previous study has reported that proteins with the highest number of interactions are the most essential proteins for survival.<sup>33</sup> We asked which time point after LPS stimulation is most essential for an innate immune response.

In order to determine the time point at which the largest number of up-/down-regulation occurs, we calculated the IPR of both up- and down-regulated proteins separately in each time-specific network, as shown in Figure 2. We found that the point with highest IPR was 1 hour after stimulation. This indicates that up-/down-regulated proteins at 1 hour may have the largest effect on the interactome during an innate immune response. The difference in IPR between 0.5 hours and 1 hour is small, but it is statistically significant ( $P = 0.004$ ). The  $P$ -value was calculated by considering the IPR of randomized networks. The IPR of randomized networks is significantly lower than that of time-specific networks, as shown in Figure 2. This IPR showed a sharp increase at 1 hour after stimulation, followed by a sharp decrease at 2 hours. This further decreased at 3 hours, followed by large increase at the transition from 6 hours to 8 hours, which is followed by stability in the IPR after 8 hours, which lasted until 24 hours after stimulation. This showed that the largest number of interactions per up-regulated protein were most affected at 1 hour after stimulation, followed by another appreciable fraction being stimulated at 6 hours. Down-regulated genes showed an increase in IPR after stimulation, though it was less than that of the up-regulated genes that were followed by a fluctuating decrease in IPR during the innate immune response.

**Hubs in innate immune response.** We found that networks that were constructed from up-/down-regulated genes had a high IPR, which indicated that these up-/down-regulated networks have a high number of hubs relative to the random networks, suggesting a role for those hubs in the innate immune response.

**Differentially expressed network.** We constructed a network of PPIs for differentially expressed proteins and their interactors. Differentially expressed proteins are proteins



**Figure 2.** Interactions protein ratio (IPR) during immune response.

The figure shows the behavior of IPR from 0.5 hours to 24 hours after LPS stimulation for up-regulated genes, down-regulated genes, and random networks. For up-regulated genes, the 1-hour time period has the highest IPR during the immune response.

whose genes showed statistically significant changes in expression throughout the whole time course of the innate immune response. The network of differentially expressed proteins consists of 3,379 proteins with 5,472 interactions. We identified protein modules from this network using MCODE<sup>25</sup> and scored them, as described in the Methods section.

Four of the functionally significant modules with their GO analysis and enriched KEGG pathways are listed in Table 2. All 47 modules identified from the differentially expressed network are noted in Table S1. Many identified modules are listed in the CORUM database, which confirms the quality of our data.

Many modules have functions related to immunity. Several modules are also enriched for pathways in cancer, such as leukemia. A link between innate immunity and cancer has been previously reported.<sup>34</sup> Among highly significant modules is the PRIM1–PRIM2–POLA1–POLA2 module, which has a score of 0.61. This module is listed in CORUM<sup>28</sup> database. Another example is the SAP30–SIN3B–IFRD1 module, which has a score of 0.28 and forms a part of a bigger complex in the CORUM database. The JAK–STAT module has a score of 0.7. In addition, the CCR3–CCR5–CCL4–CCBP complex has a score of 0.535. A few of the identified modules have high scores, but they do not have significant GO enrichment.

These modules have many differentially expressed genes and need further investigation to determine their exact role in innate immunity, such as NPHP1–NPHP4–RPGRIB1L (which has a score of 0.41) and PSTPIP1–FAS1–PTPN12, which has a score of 0.34.

**Core PPI network.** Conserved interactions during innate immune responses are those that exist in all interaction maps during the innate immune response. We identified proteins and interactions that are conserved in all time-specific interactome maps during the innate immune response and constructed a network of 390 proteins with 198 interactions, as shown in Figure 3. GO<sup>22</sup> and KEGG<sup>30</sup> pathway analysis results (Table S2) showed enrichment for biological processes such as immune system development and protein kinase cascade, while the KEGG pathway analysis showed enrichment across pathways involved in cancer and in the TLR signaling pathway. This further confirms a relationship between cancer and immunity, as reported previously.<sup>34</sup> We validated our pathway enrichment analysis by mapping genes in the KEGG pathways to the core interaction network. We found that the majority of core network members are involved in KEGG pathways, as illustrated in Figure S1.

We constructed the Trp53 network, a mouse homolog of p53, as shown in Figure S2, and found that many

**Table 2.** Four identified protein modules from network of differentially expressed proteins with their GO biological processes and KEGG pathways.

MODULE	GO FUNCTION	PATHWAY
	Synaptic transmission transmission of nerve impulse Cell-cell signaling Neurotransmitter secretion	SNARE interactions in vesicular transport
	Regulation of cell cycle Cell cycle Phosphorylation	P53 signaling pathway Cell cycle
	Blood vessel development Transforming growth factor beta receptor signaling pathway Transmembrane receptor protein serine/threonine kinase signaling pathway Phosphorylation response to cholesterol response to lipid Palate development	Cytokine-cytokine receptor interaction Pancreatic cancer pathway Chronic myeloid leukemia
	Circadian rhythm Rhythmic process Transcription Response to DNA damage stimulus Negative regulation of gene-specific transcription	Circadian rhythm

Trp53 interactors are differentially expressed due to LPS stimulation, which further confirms a relationship between the p53 pathway and LPS stimulation.

**Differential networks.** We constructed eight differential maps from nine PPI networks and identified protein complexes (modules) from each differential map, as described in the Methods section. Differential maps have been previously introduced in Bandyopadhyay et al.<sup>35</sup> Modules identified from the 0.5-hour–1-hour differential network with enriched GO biological process terms and KEGG pathways are noted in Table 3. Modules identified from other differential networks are noted in Table S3. For instance, the 0.5-hour network was characterized by down-regulation of the protein complex, which consisted of SYT1–CLSTN1–SYP–VAMP2–ATP60A1 subunits. This module is responsible for transmission of nerve impulse and cell–cell signaling. In addition, 0.5 hours after stimulation, interactome was characterized by up-regulation of the TICAM1/TBK1/TRAF6 subunits of a protein complex that consists of TICAM1–TBK1–TRAF6–TLR3. This protein module has an extremely important role in the innate immune

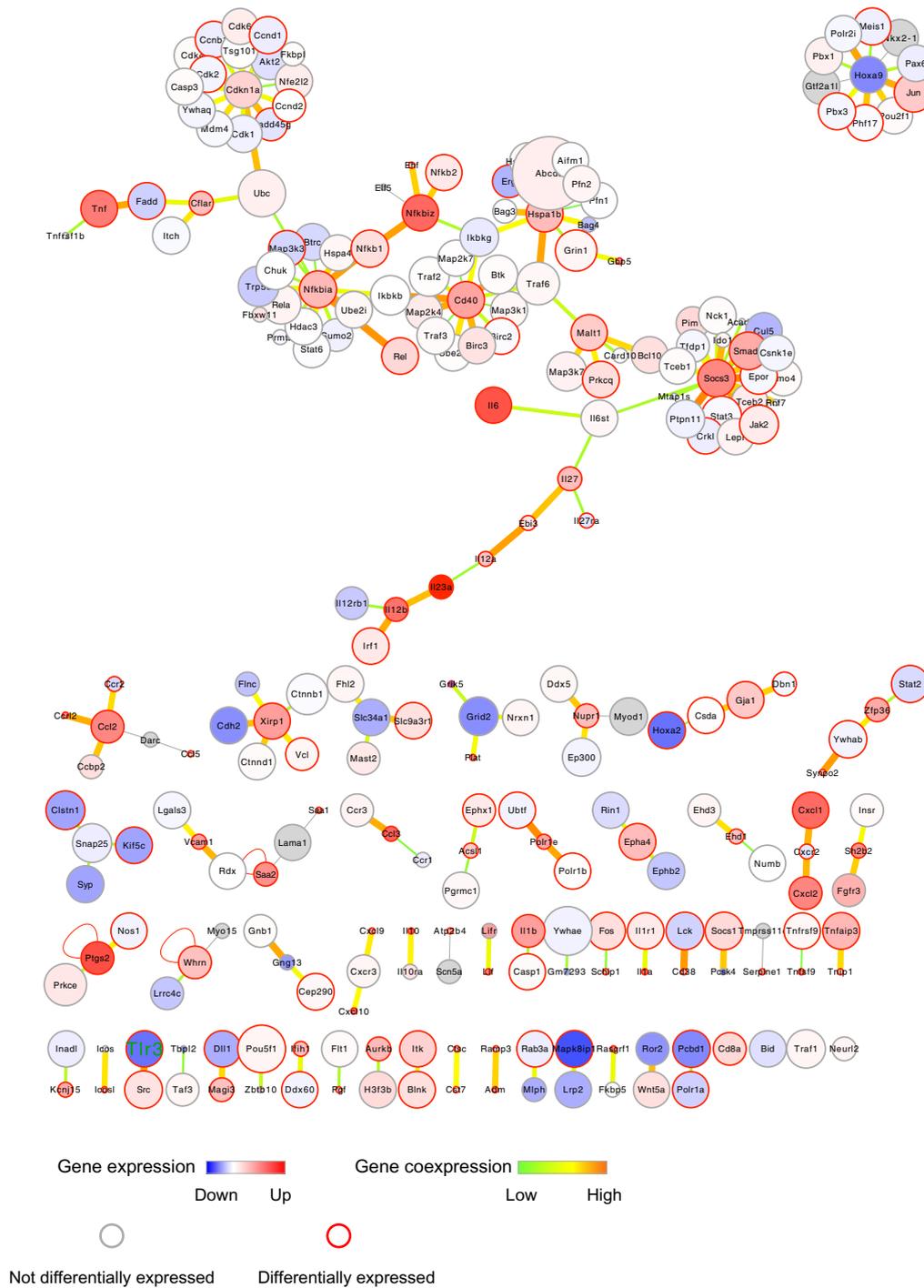
response. Moreover, the TICAM1–TBK1–TRAF6–TLR3 complex exists in the 16-hour–24-hour differential network, with all of its subunits up-regulated. Identified modules also showed changes in their expression behavior. For example, the TICAM1–TBK1–TRAF6–TLR3 module was identified at 0.5 hours, where TLR3 is significantly down-regulated. However, the same modules identified at 16 hours showed that Tlr3 was significantly up-regulated. TICAM1 was significantly up-regulated at both 0.5 hours and 16 hours. In the 6-hour–8-hour differential network, we identified a protein complex that is related to circadian rhythms and consists of the CRY2–PER3–PER2 subunits. This complex is down-regulated during innate immune responses and shows up-regulation beginning at 16 hours after stimulation. The CRY2–PER3–PER2 module has a score of 0.89, indicating that it is highly significant. From modules that were functionally identified during the innate immune response, we summarized their interactome dynamics during the immune response in terms of their biological processes (as per GO terms), as noted in Table 4.

**Pairwise differential matrix (PWDM).** For a more efficient visualization of the differences or similarities between the PPI maps at different time points throughout the immune response, we introduced the PWDM, which quantified the changes between networks at any two time points, allowing for a global view of these changes.

PWDM can be used to assess the differences between different networks. In this study, we used PWDM to assess the differences between time-specific networks throughout the experimental time course (differential network). We quantified these changes as changes in the number of nodes or interactions. The asymmetry of the heatmap can be explained by considering Figure 1. If we have two networks (A and B), the differential network A–B is different from the differential network B–A, as the interactions specific to network A are different from the interactions specific to network B. The PWDM heatmap showed that the half-hour network is the most different from other networks during the innate immune response. It also showed that the largest changes in the interactome between each of the two successive networks occurs at the transition from 0.5 hours to 1 hour. Figure 4 shows the heatmap for changes in interactions between time-specific networks. This is in line with the IPR values, which indicate that the networks show the greatest perturbation at 1 hour.

Twenty-four hours after the stimulation of DC with LPS, the differences between the time-specific networks for each consecutive time point were the least pronounced at successive time points, with the exception of the transition between 0.5 hours to 1 hour. The half-hour network showed a greater difference with respect to the other time-specific networks (Fig. 4). All of the differences were normalized and transformed to Z-scores.

**Toll-like receptor (TLR) interactome dynamics.** TLRs play a crucial role in innate immunity. Mice have 12 TLRs, which are TLR1–TLR9, and TLR11–TLR13,



**Figure 3.** Core interaction network for proteins and their interactions, which are common to all time points during the immune response. The node color represents gene expression, the edge color represents gene co-expression, the node size corresponds to the node degree, which is the number of interactions that the protein has. The node border represents whether a gene is differentially expressed or not during the immune response.

where TLR10 is not functional. However, humans have ten TLRs (TLR1–TLR10).<sup>7</sup> We constructed a TLR network from all of the available mouse PPIs in our dataset. The network included eight TLRs (TLR2, TLR3, TLR4, TLR6, TLR7, TLR8, TLR9, and TLR13) with their interactors. We constructed a network of 49 proteins and 66 interactions. The heatmap of the TLR network members is shown in Figure 5. The TLR subnetwork dynamics are shown in

Figure 6. The TLR subnetworks across the different time points during the innate immune response are noted in Figure S3.

We found that half an hour after stimulation, TLR3 was significantly down-regulated, while Ticam1 was significantly up-regulated. At 1 hour after stimulation, TLR2, SRC, and BIRC3 were significantly up-regulated. However, 2 hours after stimulation, TLR2, TLR3, TLR6, TLR7, SRC,



**Table 4.** Summary of interactome dynamics during the immune response in terms of GO terms for biological processes.

	0.5HR	1HR	2HR	3HR	4HR	6HR	8HR	16HR
UP	•Immune response	•Initiation of T cell	•Stress-activated protein	•Regulation of cell	•Protein	•Transport	•Jak-STAT cascade involved in growth	•Activation of innate immune response
	•Cytokine production	activation and differentiation	kinase signaling pathway	cycle	kinase	•Membrane/cytoskeleton organization	hormone signaling pathway	
Down	•Interferon-beta production	•Response to stimulus	•MAPKKK cascade	•rRNA transcription	cascade		•Proteolysis, ubiquitination	response
	•Type I interferon production	•Calcium mediated signaling	•Phosphorylation	•Negative regulation of cell growth			•NF-kappaB TF activity	•Defense response
	•Response to stimulus	•Peptidyl-tyrosine phosphorylation		•Apoptosis			JAK-Stat cascade	•Cytokine production
	•IL-6 production						•Cytokine mediated signaling pathway	•Positive regulation of NF-kappaB transcription
	•Positive regulation of NF-kappaB transcription factor activity							factor activity
	•Defense response							•Myeloid leukocyte activation
								•Cytokine mediated signaling pathway
Down	•Cell-cell signaling.		•DNA replication			•Negative regulation of gene specific transcription	Myeloid cell differentiation	
	•Secretion by cell		•Transcription			•Circadian rhythm		
	•Transmission of nerve impulse							

tumors.<sup>36</sup> In this study, we found the significant up-regulation of IL12 after LPS stimulation, which is in agreement with a previous study.<sup>37</sup> We further expanded the networks, as evident in Figure S4.

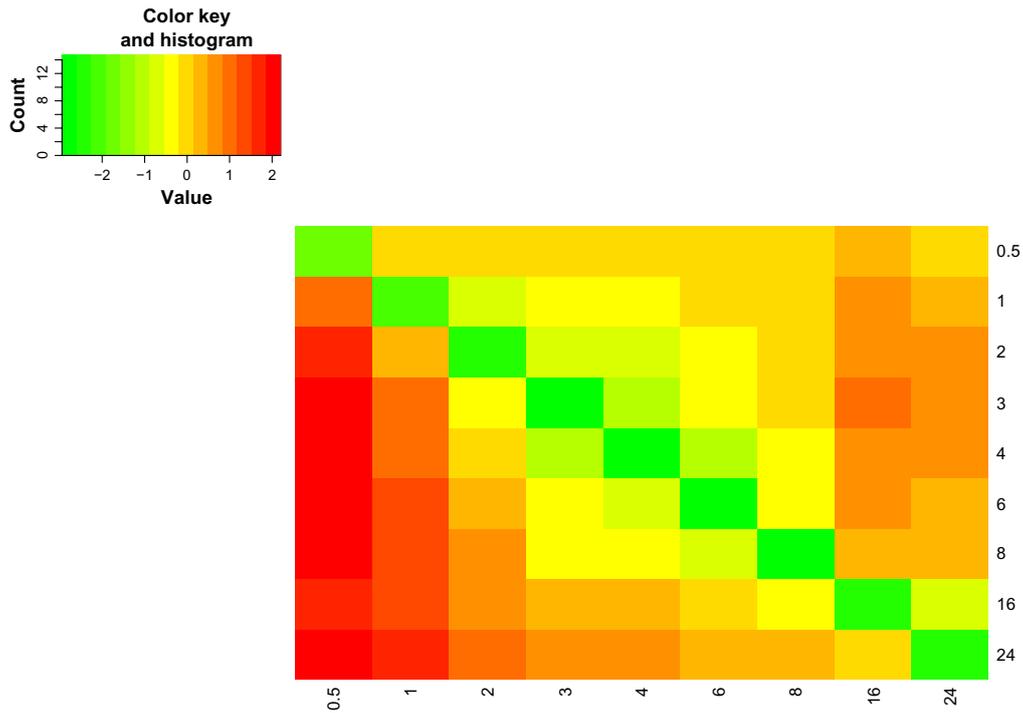
Network expansion is more informative, as it allows for the extraction of protein complexes from the network, which is not possible in smaller networks.

#### S100A8 down-regulation after LPS stimulation.

S100A8, which is called calgranulin A or migration inhibitor factor-related protein 8, and S100A9 (also called calgranulin B), are members of the S100 calcium-binding protein family. S100A8 and S100A9 showed up-regulation in many inflammatory and autoimmune cases.<sup>38,39</sup> S100A8 and S100A9 form a protein complex, which is generally known to play an inflammatory role. However, there are previous studies that suggested an anti-inflammatory role and immune-regulatory role for the S100A8/A9 complex.<sup>40–42</sup> The modulation of inflammatory processes by the S100A8/A9 complex is both context- and cell-type-specific, which suggests that there is a complex network of regulation at play. The S100A8/A9 complex has both intracellular and extracellular functions. Among the intracellular functions is calcium binding.<sup>38,39</sup> Released S100A8/A9 achieves extracellular functions, some of which are mediated

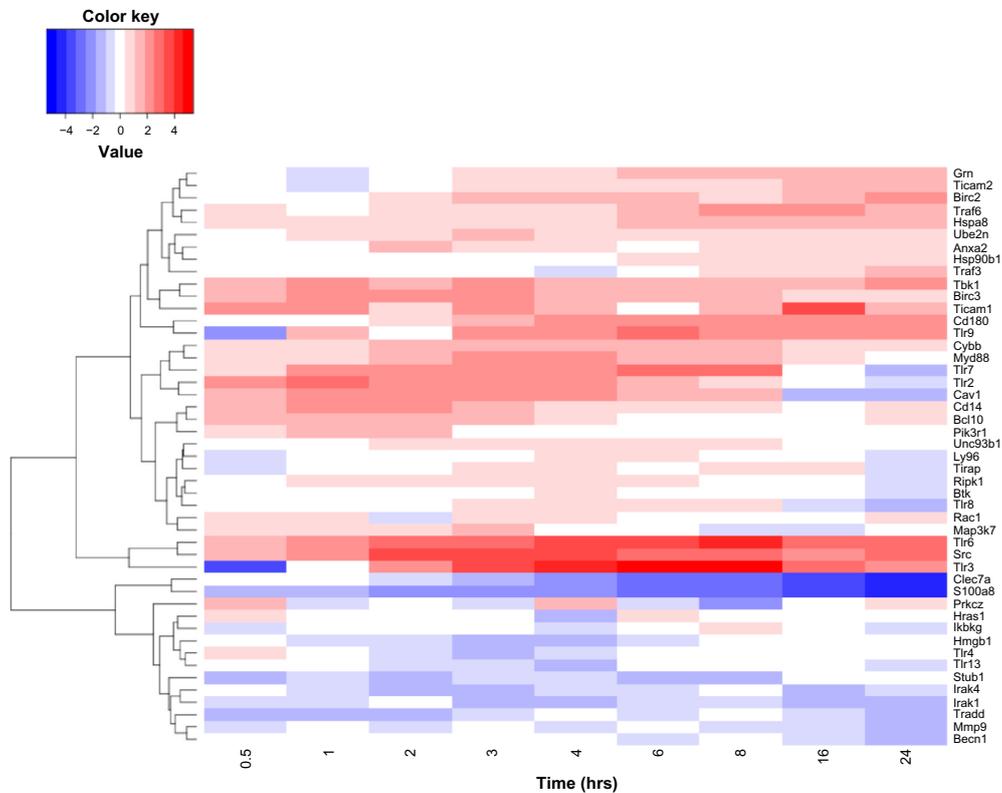
by TLR4.<sup>43</sup> A previous study reported that S100A9 over-expression blocks DC differentiation.<sup>44</sup> In addition, it showed that those DCs that over-express S100A8/A9 have less ability to stimulate allogeneic T-cell proliferation than control DCs.<sup>43</sup> S100A8 showed down-regulation during differentiation of DCs,<sup>44</sup> but in this study, we found down-regulation of S100A8 following LPS stimulation. The down-regulation is more than fourfold from 2 hours~24 hours after LPS stimulation, as shown in Figure 4.

We constructed a subnetwork of three members of the S100 protein family (S100A8, S100A9, and S100A10), as illustrated in Figure 8. We found that there is a high correlation in the co-expression of S100A8 and S100A9 proteins, which is consistent with the properties of the subunits of a complex. S100A9 only interacts with S100A8; however, S100A8 interacts with S100A9, LY96, and TLR4 (as is evident in the S100A8/A9 subnetwork). The S100A10 protein, which is another member of the S100 protein family, interacts with Irak1, an important player in the TLR signaling pathways. We checked the co-expression of Zc3h12a with S100a8. Zc3h12 is a ribonuclease, and has an important role in controlling innate immune responses.<sup>45</sup> We found that there is a high co-expression of these two genes ( $r = 0.66$ ),



**Figure 4.** Pairwise differential matrix (PWDM).

Global overview of the interactome dynamics during the immune response. The heatmap shows the differences between the networks at different time points during the innate immune response, while each cell represents the difference (unique interactions) between the PPI network in a given row and a corresponding network in a given column.



**Figure 5.** Heatmap of the TLR network genes during the time course after LPS stimulation.

The red color is for up-regulated genes, and the blue color is for down-regulated genes. The heatmap shows the gene expression of TLR network genes during the immune response. The X-axis represents the time in hours after the stimulation of dendritic cells.

which may indicate a possible functional association between them.

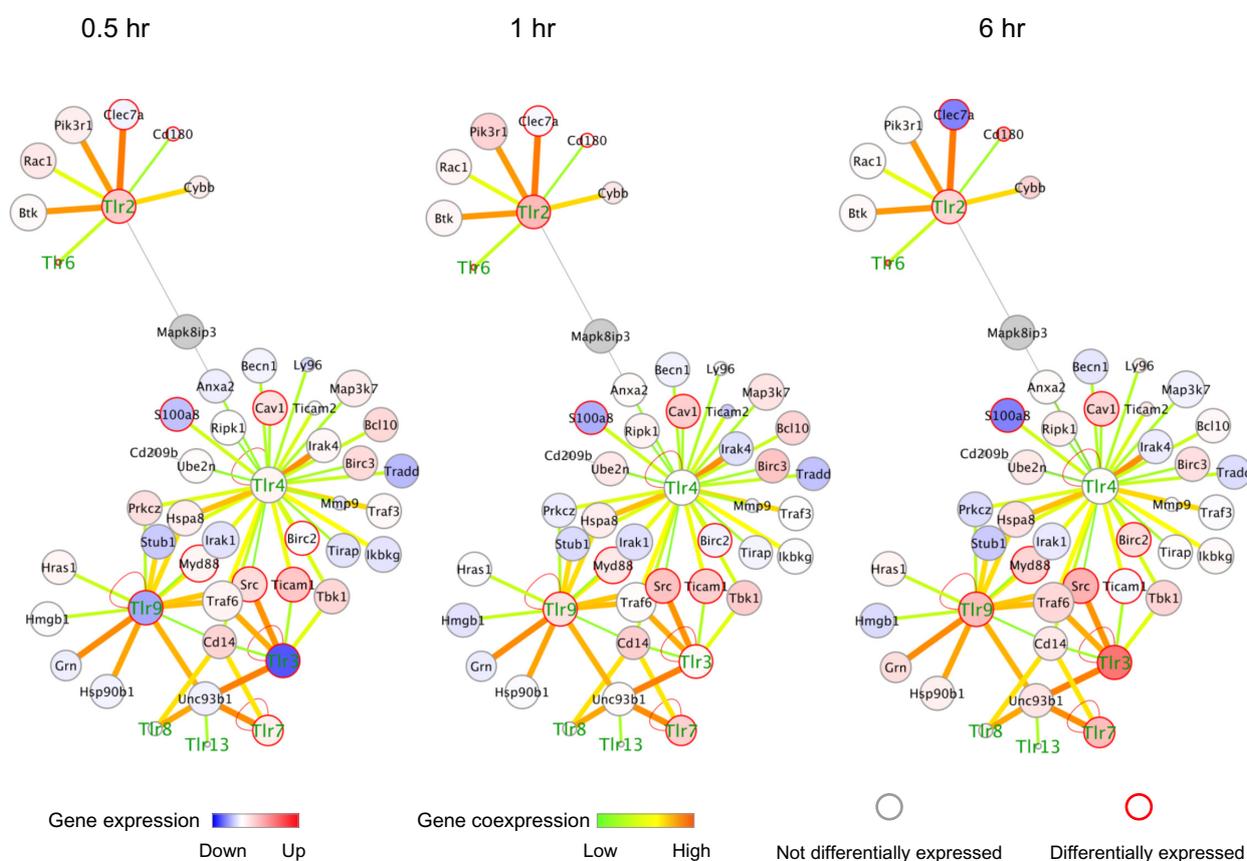
## Discussion

The goal of this study was to investigate interactome dynamics during the innate immune response. To achieve this goal, PPI data was combined with time series gene expression data. Combining PPIs with expression data is a well-established technique. One advantage of combining expression and PPI data is that it improves data reliability, as results supported by both expression and interaction data are more likely to be correct than results supported from either expression or interaction data alone.<sup>46</sup> To our best knowledge, this is the first study to infer innate immunity interactome dynamics by combining protein–protein interaction data with time series gene expression data followed by the identification of protein complexes.

Module identification is an essential step in network analysis, as protein subunits that form a protein complex have a higher probability of interacting with each other, and they have higher functional dependencies compared to other proteins in a network.<sup>47</sup> Understanding how protein complexes

are arranged in the protein interaction network will increase our understanding of the coordination of cellular and biological processes.<sup>47</sup> Therefore, identifying protein complexes/modules in a protein interaction network will highlight the proteins that have physical and functional dependencies in a given network. When identifying interactome dynamics during the innate immune response, we considered the dynamics from two points of view: First, we identified protein modules in a differentially expressed network. We found that some identified modules have a role in innate immunity, cell cycle, signaling, phosphorylation, transcription, deoxyribonucleic acid (DNA) repair, chromosome organization, transport, development, and differentiation. In addition, some modules showed enrichment of the p53 pathway, as well as the pathways involved in cancer (such as colorectal cancer and acute myeloid leukemia). A previous study suggested the existence of specific features of innate immunity mediated by TLRs in tumors; this needs further investigation.<sup>48</sup>

A direct interaction between cancer and the innate immune system has been reported.<sup>10,34</sup> Different immune cells (including DCs, which are antigen-presenting cells) target cancer cells directly, acting as the bridge between innate



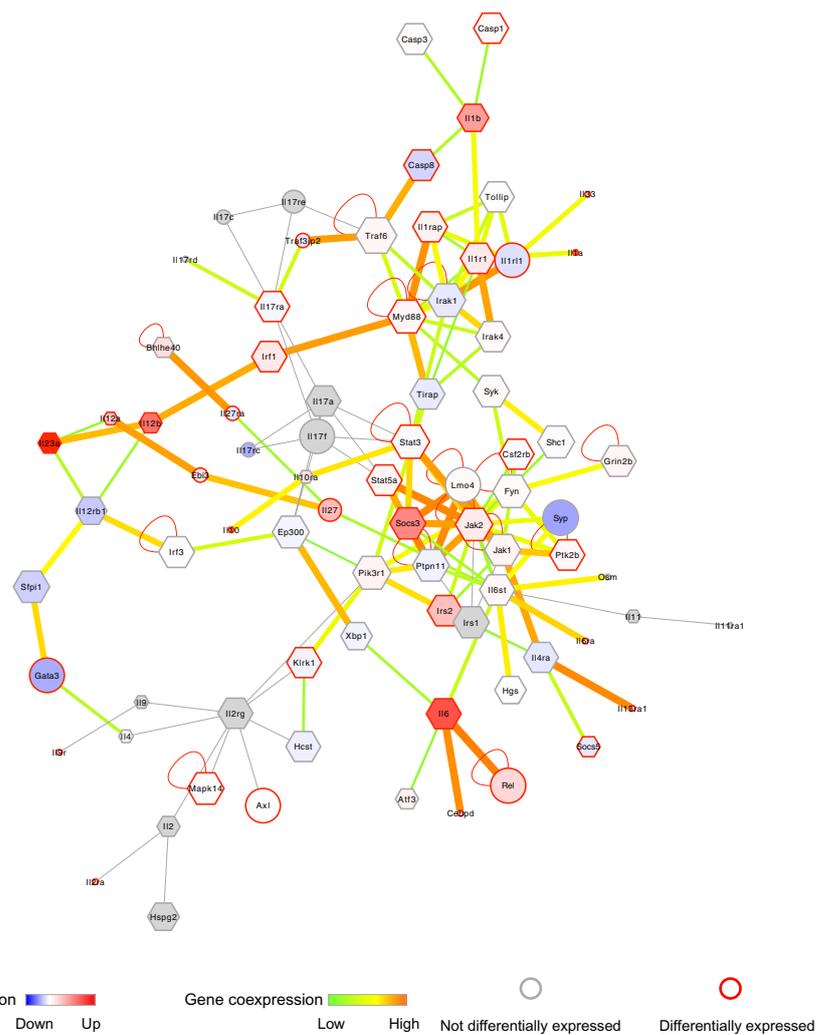
**Figure 6.** The TLR interactome dynamics.

The TLR subnetworks at different time points during the immune response. The node color represents gene expression, the edge color represents gene co-expression, node border represents whether a gene is differentially expressed or not during the immune response (as in the color key). The TLR genes have a green font, so they can be distinguished from other genes. Gene expression is different at different time points during the immune response. For example, TLR3 was down-regulated at 0.5 hours, followed by a slight up-regulation at 1 hour, and significant up-regulation at 6 hours.



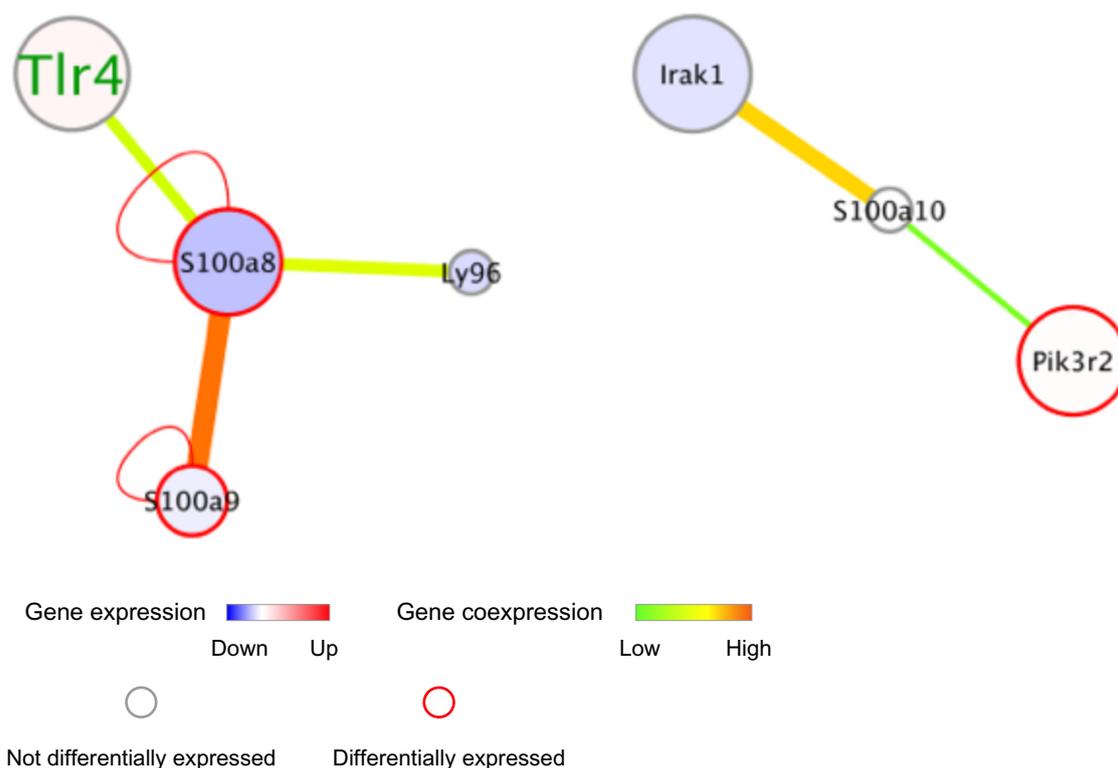
**Table 5.** Significantly up-/down-regulated proteins at each time point in the TLR interactome.

0.5HR	1HR	2HR	3HR	4HR	6HR	8HR	16HR	24HR
	<b>Tlr2</b>	<b>Tlr2</b>						
		<b>Tlr6</b>						
	<b>Birc3</b>	<b>Birc3</b>						
<b>Tlr3</b>		<b>Tlr3</b>	<b>Tlr3</b>	<b>Tlr3</b>	<b>Tlr3</b>	<b>Tlr3</b>	<b>Tlr3</b>	
		<b>Tlr7</b>	<b>Tlr7</b>	<b>Tlr7</b>	<b>Tlr7</b>	<b>Tlr7</b>		
		<b>Cav1</b>	<b>Cav1</b>	<b>Cav1</b>				
			<b>Tlr9</b>	<b>Tlr9</b>	<b>Tlr9</b>		<b>Tlr9</b>	
<b>Ticam1</b>							<b>Ticam1</b>	
	<b>Src</b>	<b>Src</b>	<b>Src</b>	<b>Src</b>	<b>Src</b>	<b>Src</b>	<b>Src</b>	<b>Src</b>
				<b>Cd180</b>	<b>Cd180</b>	<b>Cd180</b>		
				<b>Myd88</b>				
						<b>Traf6</b>		
					<b>Clec7a</b>	<b>Clec7a</b>	<b>Clec7a</b>	<b>Clec7a</b>
						<b>Prkcz</b>		
		<b>S100A8</b>						
								<b>Birc2</b>



**Figure 7.** Interleukin subnetwork.

The hexagonal nodes represent the elements of the interleukin subnetwork that can be mapped to genes involved in the KEGG pathways. Differentially expressed genes have a red border, while the edge color shows gene co-expression.



**Figure 8.** S100A8/A9 and S100A10 protein subnetworks.

The subnetworks show the interacting couplings of the S100A8, S100A9, and A100A10 proteins. Node color represents gene expression, edge color represents gene co-expression, and the node border represents whether a gene is differentially expressed or not during the immune response.

and adaptive immunity and phagocytose apoptotic cancer cells.<sup>10</sup> Furthermore, TLRs play a role in sensing the endogenous molecules released from apoptotic cells.<sup>7,10</sup> In addition, some identified modules from the differentially expressed network play a role in circadian rhythms (Table 2). A previous review had discussed a link between circadian rhythms and immunity.<sup>49</sup> Second, we identified protein modules specific to a certain time point. For instance, the half-hour time point following stimulation was characterized by up-regulation of the gene-encoding proteins that were responsible for the innate immune response, cytokine production, interferon-beta production, IL-6 production, positive regulation of nuclear factor-kappa B transcription factor activity, and down-regulation of cell-cell signaling and transmission of the nerve impulse.

This study also introduces protein complexes as candidates that play a role in the innate immune response. However, identified protein modules need further experimental validation. This study can be used to predict the immune stage of a patient. This study also introduced PWDMM, which is a novel way to assess the differences between different networks either in terms of the number of proteins or their interactions.

This study is based on an assumption that change in gene expression could reflect the change in protein abundance. A previous report<sup>50</sup> has proposed that the interactome can be affected due to changes in the abundance, sequence, or

posttranslational modification changes in an individual protein. There is a possibility that interactions in modules with small average Pearson Correlation Coefficient are transient, occurring only at a certain time point. In addition, another study has reported that changes in human transcriptome can result in significant changes in interactome.<sup>51</sup> Future studies should consider adding various parameters like post-translational modification data to increase data reliability.

In addition, this study investigated the TLR network in a comprehensive way in terms of an interaction network-based analysis for the TLR-subnetwork. One of the limitations for this kind of study is the quality of PPI data. Some of the data in manually curated databases include some protein-DNA interactions that are stored as PPIs. Cleaner PPI databases will help improve such analyses. In addition, PPI databases include many spurious interactions that come from high throughput experiments.

## Conclusion

This study has focused on an analysis of the dynamics of an innate immunity interactome, and it has identified protein modules that contain differentially expressed proteins. In addition, we identified protein modules that contain up-/down-regulated proteins specific to each time point in innate immune responses, as well as the biological roles and pathways involved in each module. We introduced IPR, which identifies the time point at which the least number of proteins with the largest number of interactions



are up-/down-regulated. We also introduced PWDMM, which allows for a global view of the differences between different networks. In this study, IPR identified 1 hour as the time point at which the least number of proteins with the greatest number of interactions were noted to be up-/down-regulated. We also found that the S100A8 protein is down-regulated during the innate immune response in DCs after LPS stimulation. Thus, we conclude that our analysis has helped in the understanding of the dynamic nature of the innate immune response.

### Authors Contributions

AE designed the analysis, performed Tss-Seq data analysis, did protein–protein interaction network analysis, and wrote the manuscript under the guidance of AP and KN. YK performed LPS stimulation experiments, YS performed TSS-sequencing. All authors reviewed and approved of the final manuscript.

### Acknowledgements

Most computations done using Human Genome Center supercomputer at University of Tokyo.

### Supplementary Data

**Supplementary Figures and Tables.pdf.** This file contains the following.

- Proof of Equation 3
- Supp Fig. 1: Core interaction network which mapped to KEGG pathway genes
- Supp Fig. 2: Trp53 (p53 mouse homolog) subnetwork
- Supp Fig. 3: TLRs interactome dynamics
- Supp Fig. 4: Interleukins expanded network
- Supp Table 1: Modules of differentially expressed network
- Supp Table 2: Gene Ontology and KEGG pathway enrichment analysis of Core Network
- Supp Table 3: Time specific modules (Modules of differential networks)

**Additional File 1.** This file contains Supp Table 4, the gene expression data and p-value for a differentially expressed network protein, which consists of differentially expressed proteins and their interactors. Gene expression values are in log. (Fold change where fold change is relative to 0hr.)

### REFERENCES

1. Takeda K, Akira S. Toll-like receptors in innate immunity. *Int Immunol.* 2005;17(1):1–14.
2. Janeway CA. Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harb Symp Quant Biol.* 1989;54 Pt 1:1–13.
3. Kawai T, Akira S. Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. *Immunity.* 2011;34(5):637–50.
4. Akira S, Takeda K, Kaisho T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nat Immunol.* 2001;2(8):675–80.
5. Janeway CA Jr, Medzhitov R. Innate immune recognition. *Annu Rev Immunol.* 2002;20:197–216.
6. Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. *Cell.* 2006;124(4):783–801.
7. Kawai T, Akira S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol.* 2010;11(5):373–84.
8. Lu YC, Yeh WC, Ohashi PS. LPS/TLR4 signal transduction pathway. *Cytokine.* 2008;42(2):145–51.
9. Poltorak A, He X, Smirnova I, et al. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science.* 1998;282(5396):2085–8.
10. Liu Y, Zeng G. Cancer and innate immune system interactions: translational potentials for cancer immunotherapy. *J Immunother.* 2012;35(4):299–308.
11. Takeuchi O, Akira S. Pattern recognition receptors and inflammation. *Cell.* 2010;140(6):805–20.
12. Cusick ME, Klitgord N, Vidal M, Hill DE. Interactome: gateway into systems biology. *Hum Mol Genet.* 2005;14 Spec No. 2:R171–81.
13. Pieroni E, de la Fuente van Bentem S, Mancosu G, Capobianco E, Hirt H, de la Fuente A. Protein networking: insights into global functional organization of proteomes. *Proteomics.* 2008;8(4):799–816.
14. Chuang HY, Lee E, Liu YT, Lee D, Ideker T. Network-based classification of breast cancer metastasis. *Mol Syst Biol.* 2007;3:140.
15. Taylor IW, Linding R, Warde-Farley D, et al. Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat Biotechnol.* 2009;27(2):199–204.
16. Inaba K, Inaba M, Romani N, et al. Generation of large numbers of dendritic cells from mouse bone marrow cultures supplemented with granulocyte/macrophage colony-stimulating factor. *J Exp Med.* 1992;176(6):1693–702.
17. Tsuchihara K, Suzuki Y, Wakaguri H, et al. Massive transcriptional start site analysis of human genes in hypoxia cells. *Nucleic Acids Res.* 2009;37(7):2249–63.
18. Flicek P, Amodè MR, Barrell D, et al. Ensembl 2011. *Nucleic Acids Res.* 2011;39(Database issue):D800–6.
19. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10(3):R25.
20. Smedley D, Haider S, Ballester B, et al. BioMart—biological queries made easy. *BMC Genomics.* 2009;10:22.
21. Robinson MD, Smyth GK. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics.* 2008;9(2):321–32.
22. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25(1):25–9.
23. Lynn DJ, Winsor GL, Chan C, et al. InnateDB: facilitating systems-level analyses of the mammalian innate immune response. *Mol Syst Biol.* 2008;4:218.
24. Patil A, Nakai K, Nakamura H. HitPredict: a database of quality assessed protein-protein interactions in nine species. *Nucleic Acids Res.* 2011;39(Database issue):D744–9.
25. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration network visualization. *Bioinformatics.* 2011;27(3):431–2.
26. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics.* 2003;4:2.
27. Assenov Y, Ramirez F, Schelhorn SE, Lengauer T, Albrecht M. Computing topological parameters of biological networks. *Bioinformatics.* 2008;24(2):282–4.
28. Ruepp A, Brauner B, Dunger-Kaltenbach I, et al. CORUM: the comprehensive resource of mammalian protein complexes. *Nucleic Acids Res.* 2008;36(Database issue):D646–50.
29. R Statistical Language [www.r-project.org].
30. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 2012;40(Database issue):D109–14.
31. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 2009;37(1):1–13.
32. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44–57.
33. Jeong H, Mason SP, Barabási AL, Oltvai ZN. Lethality and centrality in protein networks. *Nature.* 2001;411(6833):41–2.
34. Maruyama K, Selmani Z, Ishii H, Yamaguchi K. Innate immunity and cancer therapy. *Int Immunopharmacol.* 2011;11(3):350–7.
35. Bandyopadhyay S, Mehta M, Kuo D, et al. Rewiring of genetic networks in response to DNA damage. *Science.* 2010;330(6009):1385–9.
36. Trinchieri G. Interleukin-12 and the regulation of innate resistance and adaptive immunity. *Nat Rev Immunol.* 2003;3(2):133–46.
37. Langenkamp A, Messi M, Lanzavecchia A, Sallusto F. Kinetics of dendritic cell activation: impact on priming of TH1, TH2 and nonpolarized T cells. *Nat Immunol.* 2000;1(4):311–6.
38. Averill MM, Kerkhoff C, Bornfeldt KE. S100A8 and S100A9 in cardiovascular biology and disease. *Arterioscler Thromb Vasc Biol.* 2012;32(2):223–9.
39. Donato R. S100: a multigenic family of calcium-modulated proteins of the EF-hand type with intracellular and extracellular functional roles. *Int J Biochem Cell Biol.* 2001;33(7):637–68.
40. Lim SY, Raftery MJ, Geczy CL. Oxidative modifications of DAMPs suppress inflammation: the case for S100A8 and S100A9. *Antioxid Redox Signal.* 2011;15(8):2235–48.
41. Goyette J, Geczy CL. Inflammation-associated S100 proteins: new mechanisms that regulate function. *Amino Acids.* 2011;41(4):821–42.



42. Perera C, McNeil HP, Geczy CL. S100 Calgranulins in inflammatory arthritis. *Immunol Cell Biol.* 2010;88(1):41–9.
43. Vogl T, Tenbrock K, Ludwig S, et al. Mrp8 and Mrp14 are endogenous activators of Toll-like receptor 4, promoting lethal, endotoxin-induced shock. *Nat Med.* 2007;13(9):1042–9.
44. Cheng P, Corzo CA, Luetteke N, et al. Inhibition of dendritic cell differentiation and accumulation of myeloid-derived suppressor cells in cancer is regulated by S100 A9 protein. *J Exp Med.* 2008;205(10):2235–49.
45. Matsushita K, Takeuchi O, Standley DM, et al. Zc3h12a is an RNase essential for controlling immune responses by regulating mRNA decay. *Nature.* 2009;458(7242):1185–90.
46. Cline MS, Smoot M, Cerami E, et al. Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc.* 2007;2(10):2366–82.
47. Rhrissorrakrai K, Gunsalus KC. MINE: Module Identification in Networks. *BMC Bioinformatics.* 2011;12:192.
48. Alexopoulos LG, Saez-Rodriguez J, Cosgrove BD, Lauffenburger DA, Sorger PK. Networks inferred from biochemical data reveal profound differences in toll-like receptor and inflammatory signaling between normal and transformed hepatocytes. *Mol Cell Proteomics.* 2010;9(9):1849–65.
49. Habbal OA, Al-Jabri AA. Circadian rhythm and the immune response: a review. *Int Rev Immunol.* 2009;28(1):93–108.
50. Wilkins MR, Kummerfeld SK. Sticking together? Falling apart? Exploring the dynamics of the interactome. *Trends Biochem Sci.* 2008;33(5):195–200.
51. Davis MJ, Shin CJ, Jing N, Ragan MA. Rewiring the dynamic interactome. *Mol Biosyst.* 2012;8(8):2054–66, 2013.