

## Research Article

# Identification of Transcriptional Modules and Key Genes in Chickens Infected with *Salmonella enterica* Serovar Pullorum Using Integrated Coexpression Analyses

Bao-Hong Liu<sup>1,2</sup> and Jian-Ping Cai<sup>1,2</sup>

<sup>1</sup>State Key Laboratory of Veterinary Etiological Biology, Key Laboratory of Veterinary Parasitology of Gansu Province, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, Gansu, China

<sup>2</sup>Jiangsu Co-Innovation Center for Prevention and Control of Important Animal Infectious Diseases and Zoonoses, Yangzhou, China

Correspondence should be addressed to Jian-Ping Cai; [caijianping@caas.cn](mailto:caijianping@caas.cn)

Received 14 November 2016; Revised 1 March 2017; Accepted 27 March 2017; Published 26 April 2017

Academic Editor: Ansgar Poetsch

Copyright © 2017 Bao-Hong Liu and Jian-Ping Cai. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Salmonella enterica* Pullorum is one of the leading causes of mortality in poultry. Understanding the molecular response in chickens in response to the infection by *S. enterica* is important in revealing the mechanisms of pathogenesis and disease progress. There have been studies on identifying genes associated with *Salmonella* infection by differential expression analysis, but the relationships among regulated genes have not been investigated. In this study, we employed weighted gene coexpression network analysis (WGCNA) and differential coexpression analysis (DCEA) to identify coexpression modules by exploring microarray data derived from chicken splenic tissues in response to the *S. enterica* infection. A total of 19 modules from 13,538 genes were associated with the Jak-STAT signaling pathway, the extracellular matrix, cytoskeleton organization, the regulation of the actin cytoskeleton, G-protein coupled receptor activity, Toll-like receptor signaling pathways, and immune system processes; among them, 14 differentially coexpressed modules (DCMs) and 2,856 differentially coexpressed genes (DCGs) were identified. The global expression of module genes between infected and uninfected chickens showed slight differences but considerable changes for global coexpression. Furthermore, DCGs were consistently linked to the hubs of the modules. These results will help prioritize candidate genes for future studies of *Salmonella* infection.

## 1. Introduction

Chickens are an important component in the global agricultural economy by serving as one of the primary sources of proteins for humans. However, the poultry industry has been consistently threatened by various diseases, including those caused by viral, bacterial, and parasitic infections. *Salmonella enterica* serovar Pullorum (*S. Pullorum*) is one of the most important pathogens of poultry causing severe systemic disease [1, 2]. To prevent and control *S. Pullorum* in chickens, the host responses against this pathogen have been studied for decades. Although significant advances have been made, especially in the identification of molecules and genes involved in the host immune response [3, 4] and mucosal inflammation [5, 6], as well as their differential expression during infection [7–10], the precise pathways

regulating immunity to *Salmonella* infection using a systems biology approach have not been investigated. Although gene differential expression analysis (DEA) provides important information, such as identification of genes that are expressed at different times during infection, which inform our understanding of pathogenesis, identifying gene interactions using a systems biology approach greatly enhances our knowledge at the mechanistic and regulatory levels. A large amount of information regarding gene interactions is available in microarray datasets and by applying network approaches the gap between individual genes and systems can be bridged [11–13]. The modularity in biological systems allows for both the study of independent components and identification of gene relationships within modules. Modern approaches, such as weighted gene coexpression network analysis (WGCNA) [14], can identify modules with expression levels that are

highly correlated across samples and have been used to identify new candidate regulatory molecules and networks in *Salmonella*-infected pigs [15]. Differentially coexpressed modules (DCMs) can also be identified [16]. The holistic changes in modules would be reflected in transcriptional and coexpression changes for individual genes. In general, gene expression levels change during disease or infection, but some have reported that seemingly nonsignificant DEGs may also play a key role in a disease because their interactions with other genes change considerably [17]. These genes can be identified via differential coexpression analysis (DCEA), which can mine individual genes using a holistic approach [17–19]. Hence, combining the WGCNA and DCEA methods can identify interacting modules and differentially coexpressed genes (DCG) during infection, compared with controls. Here, we mined the molecular network relationships of the differential coexpression modules and genes using microarray data from spleens of *S. Pullorum*-infected and uninfected chickens using WGCNA and DCEA (Figure 1). The results complement traditional DEA and add to our understanding of the regulatory mechanisms that occur during *Salmonella* infection.

## 2. Materials and Methods

**2.1. Microarray Data Harvesting and Processing.** A comprehensive transcriptomics dataset derived from microarray analysis of spleens from chickens challenged with  $10^8$  CFU of *Salmonella enterica* serovar Pullorum or mock-challenged with the same volume of distilled water (controls) was obtained from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>) (accession number: GSE59663). The dataset was generated with the Agilent oligo microarray chips containing 43,663 probe sets. In this study, we first streamlined the dataset by excluding 14,920 probe sets that were either unmappable to any gene IDs or mapped to multiple gene IDs. In the case of multiple probe sets mapped to one identical gene, the probe set, which is most often associated with the highest expression level, was maintained to ensure that only one probe set was left to investigate one gene. If more than one probe set was left after the above steps, their intensities were averaged. Finally, a one-to-one match between 13,538 probe sets and 13,538 genes was achieved.

Three biological replicates (chips) for each time point were available in the challenged group for these datasets. However, at each time point in the control group, only one chip was used to hybridize with the equally mixed mRNA sample containing the three control samples. We averaged the replicates for each time point, except at day 21, with the two replicates included and forming the dataset for the challenged group with 10 samples; this dataset was equivalent to the dataset of the control group. The dataset was quantile normalized by the function of `normalizeQuantiles` in R package `limma` [20].

**2.2. Construction of Weighted Gene Coexpression Network and Identification of Modules.** Weighted gene coexpression network analysis (WGCNA) was used to detect coexpression

modules from the dataset of challenged samples [14, 21]. The R function of `blockwiseModules` was implemented with the following parameters: `power = 12`, `minModuleSize = 100`, and `networkType = "signed."` Microarray data were processed as described below.

The pairwise Pearson's correlation coefficients were calculated for all the genes in the challenged groups, followed by the construction of an adjacency matrix using the power function:

$$\alpha_{ij} = (0.5 + 0.5 \times \text{cor}(x_i, x_j))^\beta, \quad (1)$$

where  $x_i$  and  $x_j$  were the  $i$ th and  $j$ th gene expression traits, respectively, which formed a signed weighted correlation network; and  $\beta$  used default value (i.e.,  $\beta = 12$ ). The topological overlap measure (TOM) was calculated as follows:

$$\text{TOM}_{ij} = \frac{\sum_{u \neq i, j} \alpha_{iu} \alpha_{uj} + \alpha_{ij}}{\min(k_{\text{total}, i}, k_{\text{total}, j}) + 1 - \alpha_{ij}}, \quad (2)$$

where  $k_{\text{total}}$  is the sum of connection strengths for a gene with the other network genes.  $u$  is the other network genes.

Afterwards, 1-TOM was calculated as a biological important measure for network interconnectedness. Genes with highly similar coexpression relationships were grouped together by performing hierarchical clustering on the topological overlap. Subsequently, genes were hierarchically clustered using 1-TOM as the distance measure and modules were determined by choosing a height cutoff of 0.995 for the resulting dendrogram. Highly similar modules were identified by clustering and merged together using a dynamic tree-cutting algorithm [14]. Eigengene refers to the first principal component for a given module and could be calculated to draw a module trajectory curve [14].

**2.3. Identification of Differentially Coexpressed Modules.** Differentially coexpressed modules (DCMs) were identified using gene-set coexpression analysis (GSCA) that adopted the length-normalized Euclidean distance to measure the coexpression difference for the pairwise correlations between infected and control groups [16].

$$D_m = \sqrt{\frac{1}{P_m} \sum_{p=1}^{P_m} (r_p^c - r_p^i)^2}, \quad (3)$$

where  $P_m$  was the number of gene pairs from the pairwise correlation for all the module genes.  $r_p^c$  and  $r_p^i$  were the correlation coefficients for a gene pair in the control and infected groups, respectively.

The null distribution for distance was constructed by permuting samples across conditions for 10,000 times to yield gene-set specific  $p$  values. Modules with  $p$  value  $< 0.01$  were considered as significantly differentially coexpressed.

**2.4. Identification of Differentially Coexpressed Genes.** The differential coexpression analysis (DCEA) was implemented by using R package `DCGL`, which is a useful tool to identify

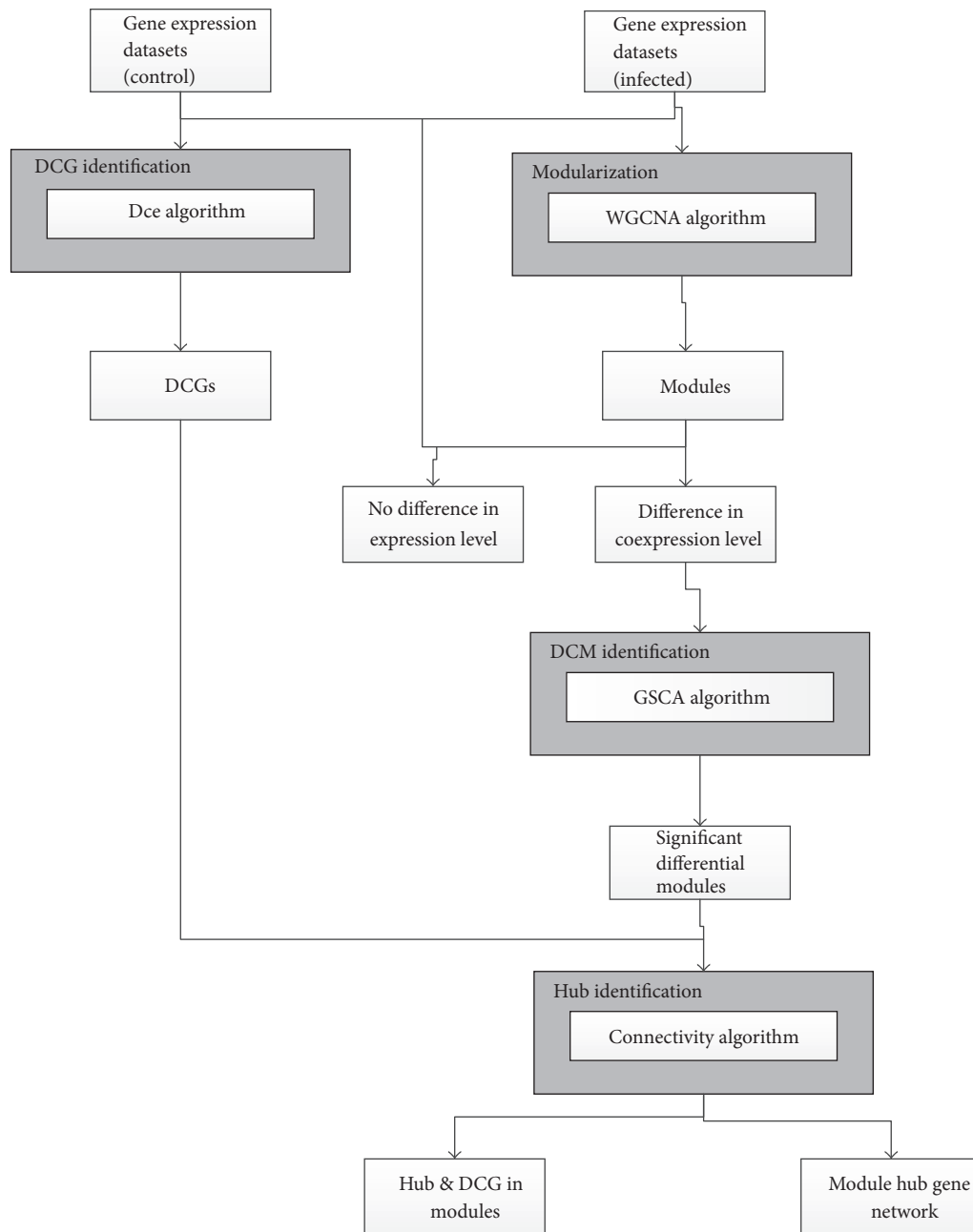


FIGURE 1: Workflow of the comprehensive gene coexpression network analysis.

differentially coexpressed genes (DCGs) and differentially coexpressed links (DCLs) [17–19]. The R function DCE was applied and then the  $p$  values were adjusted for a false discovery rate (FDR) using the Benjamini-Hochberg method to reduce a large amount of false positive results [22]. The genes with  $FDR < 0.001$  were selected as DCGs.

**2.5. Gene Ontology (GO) and Pathway Enrichment for Coexpression Modules.** GO enrichment and KEGG pathway analyses for network modules were performed using Database for Annotation, Visualization, and Integrated Discovery (DAVID, v6.7) program using all chickens genes as the background [23, 24]. The modified Fisher’s exact test with

an adjustment for multiple tests by Benjamini-Hochberg method was used to identify significantly enriched terms for module genes [22].

**2.6. Network Visualization.** The complex network bioinformatics software Cytoscape (v3.1.1) was used to visualize the pairwise relationships between genes [25].

### 3. Results

**3.1. Weighted Gene Coexpression Network Analysis.** Using blockwiseModules R function ( $\beta = 12$ ), a total of 19 modules ranging from 100 to 3,000 genes were recovered for the 13,538

TABLE 1: Module preservation and functions.

Module	Size	Z summary	Function
<i>Lightyellow</i>	126	5.42	Nucleus ( $8.40E - 4$ )
<i>Lightgreen</i>	145	0.67	Jak-STAT signaling pathway ( $9.73E - 3$ )
<i>Grey60</i>	145	7.49	Extracellular matrix ( $5.10E - 4$ ) Cytoskeleton ( $7.84E - 3$ )
<i>Lightcyan</i>	155	0.52	Anchored to membrane ( $4.19E - 3$ )
<i>Midnightblue</i>	159	5.78	Cytoplasm ( $1.75E - 4$ ); Organelle membrane ( $4.67E - 3$ ) Endomembrane system ( $9.27E - 3$ )
<i>Cyan</i>	181	8.71	Cell adhesion molecules ( $3.24E - 5$ ) Cell adhesion ( $2.75E - 3$ )
<i>Salmon</i>	248	2.14	Neuroactive ligand-receptor interaction ( $2.60E - 6$ )
<i>Tan</i>	287	4.86	Lysosome ( $4.80E - 3$ )
<i>Greenyellow</i>	293	6.58	Ligase activity ( $3.41E - 3$ )
<i>Purple</i>	298	10.93	Proteasome complex ( $6.56E - 6$ ) Regulation of cytokine biosynthetic process ( $3.83E - 3$ ) Toll-like receptor signaling pathway ( $9.52E - 3$ )
<i>Magenta</i>	357	1.82	G-Protein coupled receptor activity ( $2.88E - 4$ )
<i>Pink</i>	418	13.36	Cellular macromolecular complex assembly ( $1.00E - 03$ )
<i>Black</i>	599	1.97	Postsynaptic membrane ( $2.35E - 3$ ) Synapse ( $2.72E - 3$ )
<i>Red</i>	648	14.89	Signal transducer activity ( $2.07E - 4$ ) Multicellular organism development ( $2.70E - 4$ )
<i>Green</i>	1056	27.63	Cell cycle phase ( $9.63E - 13$ ) DNA replication ( $3.67E - 7$ ) Response to DNA damage stimulus ( $1.45E - 6$ ) DNA repair ( $4.11E - 6$ ) Cytoskeleton organization ( $3.77E - 4$ )
<i>Yellow</i>	1122	32.83	Glucose catabolic process ( $4.03E - 5$ ) Glycolysis/gluconeogenesis ( $1.49E - 4$ ) Glycolysis ( $1.79E - 4$ ) Glucose metabolic process ( $3.95E - 4$ )
<i>Brown</i>	1349	18.99	ABC transporters ( $1.00E - 03$ )
<i>Blue</i>	2581	43.01	Immune system process ( $1.31E - 4$ ) Induction of apoptosis ( $1.98E - 4$ ) Antigen processing and presentation ( $2.06E - 4$ ) Lysosome ( $3.08E - 4$ ) Defense response to bacterium ( $6.06E - 3$ )
<i>Turquoise</i>	2998	51.56	Nervous system development ( $2.20E - 15$ ) Focal adhesion ( $2.04E - 9$ ) Wnt signaling pathway ( $7.56E - 9$ ) Regulation of actin cytoskeleton ( $1.64E - 7$ ) TGF-beta signaling pathway ( $4.42E - 7$ )

Note. The column "Size" gives the gene numbers contained in every module. "Z summary" gives the z score of module preservation. "Function" gives the module functions enriched by DAVID.

distinct genes in the *S. Pullorum*-infected group (Table 1). Each module was assigned a unique color, including gray color for the 373 unassigned genes. Genes in the same module shared the same or similar expression patterns that were catalogued by the trajectory curves (Figure 2).

Subsequent analysis using DAVID identified biological features in modules that were potentially associated with the infection by *S. Pullorum* (Figure 6 and Table 1), such as the Jak-STAT signaling pathway (module lightgreen) [26], the extracellular matrix (ECM) (module grey60) [27],

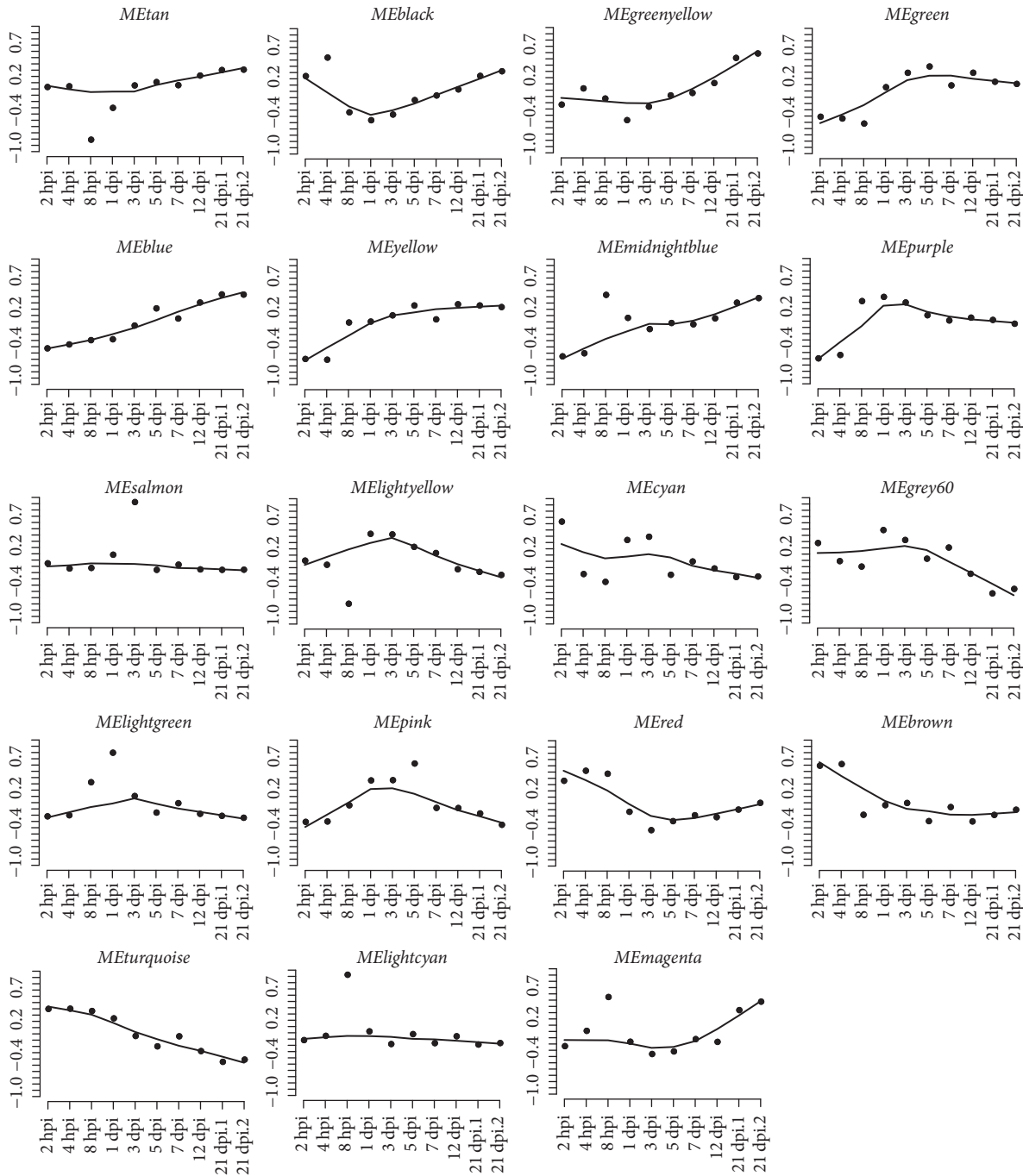


FIGURE 2: Module expression patterns.

cytoskeleton organization (module green), regulation of the actin cytoskeleton (module blue) [28], G-protein coupled receptor activity (module magenta), Toll-like receptor signaling pathways (module purple), and immune system processes (module blue). ECM genes and cell adhesion genes are significantly enriched in the module grey60 and cyan ( $FDR = 5.10e - 4$  and  $2.75e - 3$ ), respectively (Table 1). The grey60 and cyan modules also displayed significant similarity in expression patterns (eigengenes' correlation = 0.76;  $p = 0.01$ ). These observations were in congruent with those reported earlier by others on the crucial role of host cell ECM proteins

and bacterial outer membrane structures in the adhesion and invasion of *Salmonella* [27].

**3.2. Module Stability.** To test the reproducibility of the identified modules, we performed a sampling test, in which we randomly selected half of the samples to calculate the new intramodule connectivity. The sampling was repeated 100 times and then the module stability was expressed as the correlation of intramodule connectivity between the original and sampled ones [29]. Most modules displayed good stability; module salmon was the least stable (Figure 3).



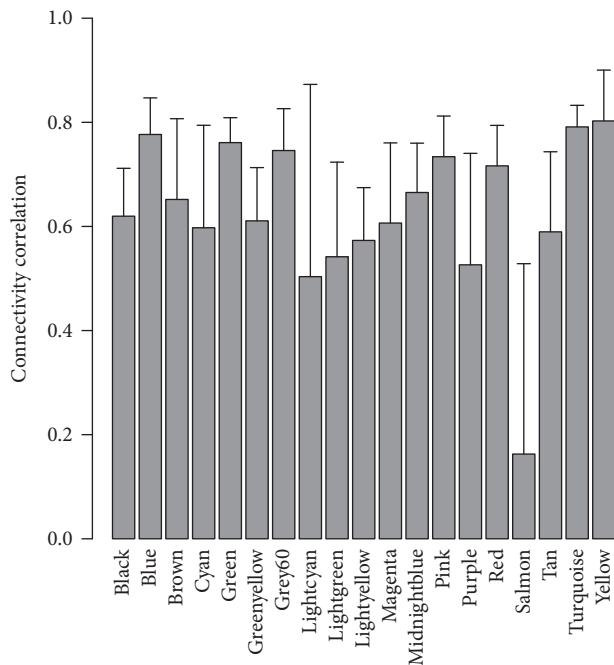


FIGURE 3: Correlation of intramodule connectivity for each module after 100 samplings.

**3.3. Module Preservation Analysis.** We investigated whether the *S. Pullorum*-infected module was preserved in the corresponding controls by testing whether the infection-associated coexpression network can be replicated in the control groups. The preservation scores for all the modules were listed in Table 1, in which  $Z$  summary scores  $<2$ , between 2 and 10, and  $>10$  indicate no evidence, weak-to-moderate evidence, and strong evidence for module preservation, respectively. Preservation analysis provided strong evidence to support the conservation of modules turquoise, brown, blue, yellow, green, red, pink, and purple, which all contained considerably large numbers of genes, but no evidence to support the preservation of modules lightcyan, lightgreen, magenta, and black associated with the membranes, the Jak-STAT signaling pathway, G-protein coupled receptor activity, and synapses, respectively (Table 1).

**3.4. Module Gene Expression and Coexpression Comparison.** We compared the module genes' expression and coexpression level between the infected and control groups. The violin plot in Figure 4(a) showed that the gene expression for modules in the infection versus control groups is not significantly different, and the distribution for the expression intensities is similar. Subsequently, we compared the gene coexpression level by calculating the gene connectivity for each module. The module turquoise exhibits the largest connectivities since it includes the largest number of genes (2,998 genes). Modules blue (2,581 genes), yellow (1,122 genes), brown (1,349 genes), and green (1,056 genes), which include a considerable number of genes, display the next highest connectivities. In addition, the coexpression levels are different between modules in the two conditions. The coexpressions are strengthened in the infected state (Figure 4(b)).

**3.5. Identification of Differentially Coexpressed Modules.** Gene-set coexpression analysis (GSCA) revealed that 14 of the 19 modules were significantly differentially coexpressed ( $p < 0.01$  by bootstrap sampling test) (Table 2). Among them, modules black ( $z = 1.97$ ), magenta ( $z = 1.82$ ), salmon ( $z = 2.14$ ), and lightcyan ( $z = 0.52$ ) were significantly differentially coexpressed. These observations were in agreement with the module preservation analysis, in which significantly differentially coexpressed modules (DCM) were only weakly preserved in the control group.

**3.6. Identification of Differentially Coexpressed Genes.** A total of 2,856 differentially coexpressed genes (DCG) were selected with a false discovery rate (FDR) of less than 0.001 using the DCe method in the DCGL package. And a total of 284,213 differentially coexpressed links (DCLs) were same signed, 82,619 were differently signed, and 272,491 were switched links.

Furthermore, we mapped the DCGs for each module and found that the DCMs enrich the DCGs. For example, a total of 152 DCGs appeared in the module magenta with  $Z$  summary of 1.82 ( $p = 0$ ), 231 DCGs in module black with  $Z$  summary of 1.97 ( $p = 0$ ), and 147 DCGs in module salmon with  $Z$  summary of 2.14 ( $p = 0$ ). In network biology, a hub gene is a good representative of a module. We identified the hub genes for all of the modules. Table 2 gives the gene names which are not only hub genes but also DCGs in each module.

## 4. Discussion

We constructed a gene network for the *S. Pullorum*-infected chickens using weighted gene coexpression network analysis (WGCNA) from the data of time-series microarray. This module detection strategy utilizes the biological variability inherent in the prospective cohort study to reveal the modular organization and function of transcriptional systems. The time series expression profiles allow the study of the transcriptional regulation of these gene coexpression networks during infection. A network-based analysis provides a systems-level understanding of the relationships between members of a network by focusing on genome-wide gene modules rather than individual genes [30]. Differential expression analysis (DEA) aims to identify genes that are expressed significantly higher or lower in one group compared with another. By contrast, WGCNA is not biased toward genes with significant changes in expression. Moreover, the dimensionality of microarray data in the present study was reduced from 13,538 genes to 19 modules, which significantly increased the ability to identify concordant changes in the expression of multiple genes.

The module expression analysis showed that module salmon was the least abundant but exhibited the largest variation in gene expression causing the instability in module construction. Gene expression was the most stable within module midnightblue (Figure 4(a)). The expression distribution for module genes in different conditions (infected versus control) was the same. The coexpression level was further compared. The coexpression comparison showed significant changes for different conditions. We investigated

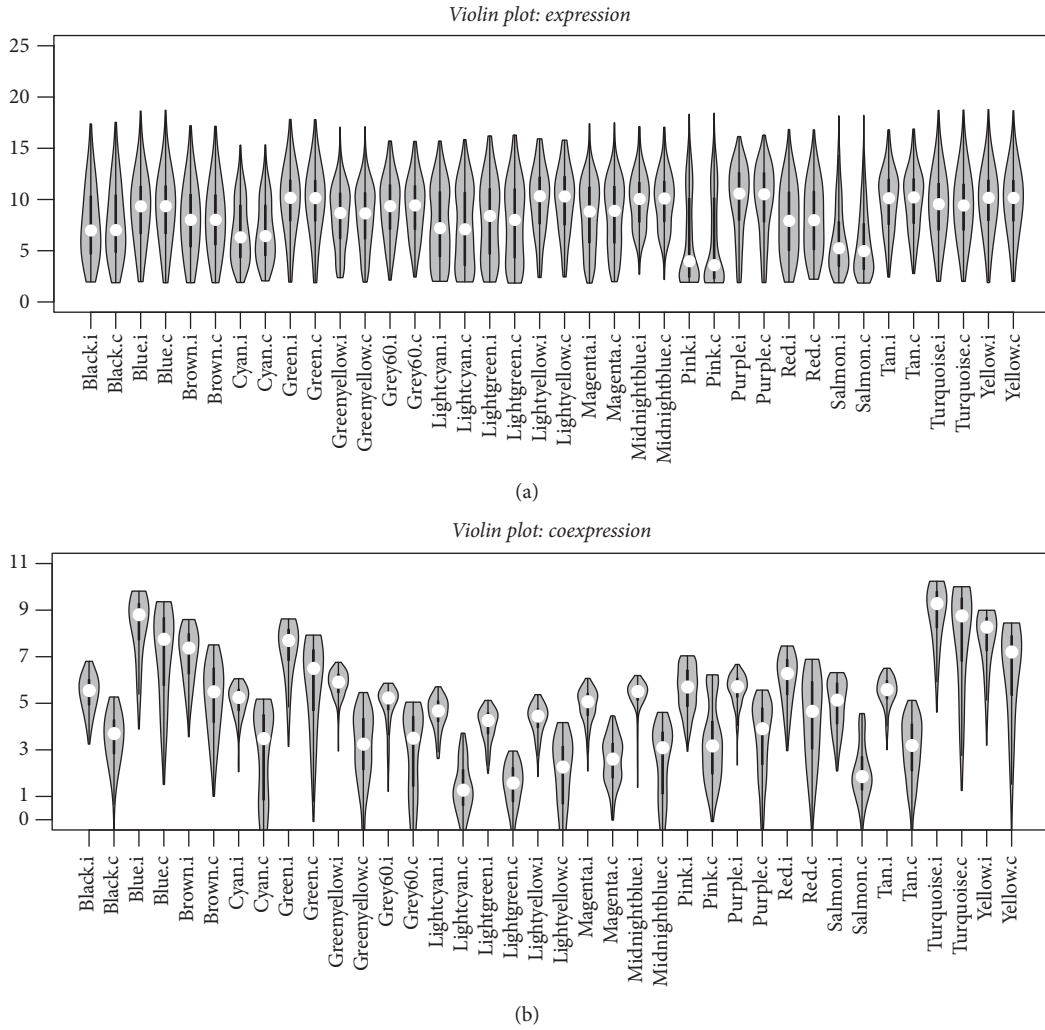


FIGURE 4: (a) Violin plot showing the gene expression differences between modules in the infected and control groups. (b) Violin plot showing the gene coexpression connectivity differences between modules in the infected and control groups. .i represents the module in the infected group, and .c represents the module in the control group.

the key modules and genes resulting in the differences. The GSCA analysis identified ten significantly differentially coexpressed modules (DCMs), which are in accordance with the module preservation results that are significant in coexpression differences with little evidence for preservation. Regulatory relationships among genes can be parsed as the pairwise correlations between gene expression levels, so the changes in coexpression patterns between two conditions may indicate dysfunctional regulatory systems in disease [31]. Module lightgreen associated with the Jak-STAT signaling pathway and lightcyan associated with membrane anchoring functions were the two weakest preserved modules (Table 1 and Figure 5). Thus, these modules may be associated with *S. Pullorum* infection in chickens.

Furthermore, we investigated the driven genes leading to the coexpression difference. The differential coexpression analysis (DCEA) method was applied, and 2,856 differentially coexpressed genes (DCGs) were identified. Compared to the differential expression analysis (DEA), it was found that the

overlapping of DCGs with the 234 DEGs (*t*-test *p* value less than 0.01) was significant (hypergeometric test  $p = 1.07e - 07$ ), indicating that differential expression and differential coexpression are somewhat related to each other, which is consistent with a previous report [17]. However, there are many *Salmonella* infection-related genes identified by the DCEA method. The top one DCG identified is *WASF1*, which is an important gene in the *Salmonella* infection pathway and was not identified as a DEG (expression fold change: 1.08; *t*-test *p* value of 0.34). The protein encoded by *WASF1*, a member of the Wiskott-Aldrich syndrome protein family, plays a critical role downstream of *Rac*, which is a Rho family of small GTPases, in regulating the actin cytoskeleton required for membrane ruffling. This gene associates with an actin nucleation core Arp2/3 complex while enhancing actin polymerization in vitro [32]. Another gene, *CDC42* (fold change = 1.02; *p* = 0.6), a member of the Rho subfamily of actin-organizing small GTP-binding proteins, interacts with *WASF1* and is essential for *S. Typhimurium* entry into host

TABLE 2: Differentially coexpressed modules enriched with differentially coexpressed genes (DCGs).

Module	Size	GSCA, <i>p</i>	Hub and DCGs
Black	599	0	GJA8, LOC421988, MYOCD
Magenta	357	0	CI0orf83, PPME1, NRTN, TMEM167B, ABCA5, CI0orf58
Salmon	248	0	CPA5, PRSS2, ARHGGEF19, INS, CELA2A, TCERG1L, LOC396296, LOC771434, KCNC2, XKR9
Lightcyan	155	5.55E - 16	LOC769741, TRPA1, CI0orf96, CASKIN2, GMCSE, cor6
Brown	1349	2.11E - 09	—
Cyan	181	1.77E - 08	RAB36, LOC415324, UPK3B, CWH43
Lightgreen	145	4.05E - 07	CIPI, ZC3H12D, TNS3, SOCS3, LOC415844
Pink	418	4.20E - 07	ABCG8
Lightyellow	126	2.10E - 05	DEPDC1, CDCA2
Tan	287	2.23E - 04	TP53INP1, CNRIP1, NEURL, SEPP1
Midnightblue	159	1.16E - 02	LOC416257
Greenyellow	293	2.04E - 01	INVS, ARSH
Grey60	145	2.10E - 01	—
Purple	298	2.51E - 01	WDR5, NIPA2
Red	648	8.66E - 01	—
Blue	2581	1	—
Green	1056	1	—
Turquoise	2998	1	—
Yellow	1122	1	—

Note. The column "Size" gives the gene numbers contained in every module. "GSCA,*p*" gives the *p* value calculated by GSCA for modules between the two different conditions and the *p* value less than 0.05 indicates that the module is significantly differentially coexpressed. "Hub and DCGs" shows the genes which are DCGs in the top ten hub genes.



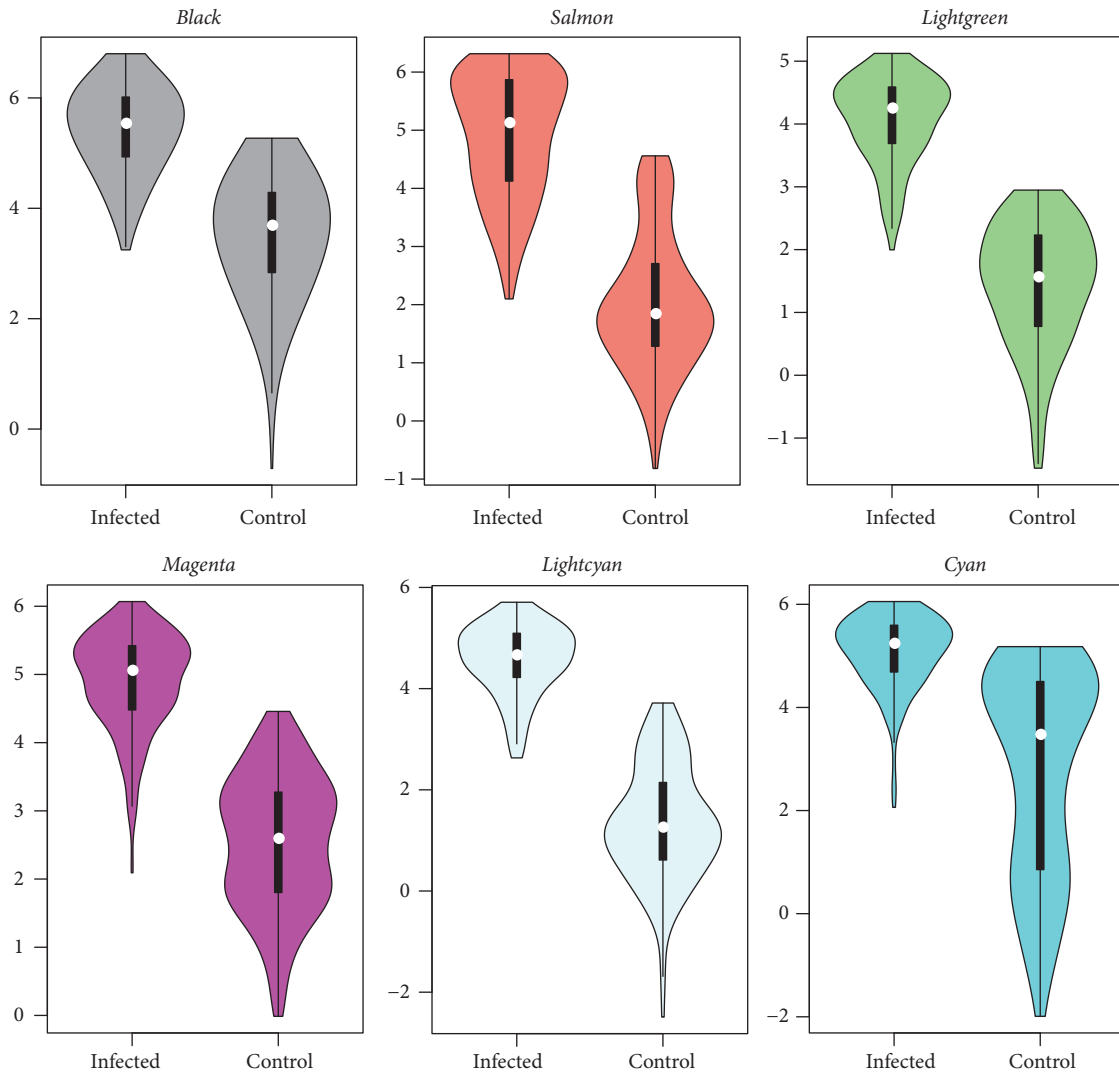


FIGURE 5: Violin plot of gene coexpression connectivity for significantly differentially coexpressed modules.

cells [33, 34]. *CDC42* was not selected as DCG with a false discovery rate (FDR) of  $1.55E - 3$ , but it interacts with genes *PAK7* (fold change = 1.95;  $p = 0.64$ ) [35, 36], *CDC42EP3* (fold change = 1.05;  $p = 0.81$ ) [37, 38], *PAK1* (fold change = 0.97;  $p = 0.67$ ) [39], *PARD6B* (fold change = 0.76;  $p = 0.05$ ) [40, 41], *PARD6A* (fold change = 1.75;  $p = 0.90$ ) [29, 40], and *IQGAP2* (fold change = 1.58;  $p = 0.12$ ) [42], which are all identified as DCGs. Carow and Rottenberg reported that gene *SOCS3*, which was also identified as a DCG (fold change = 1.54;  $p = 0.15$ ), is a major regulator of infection and inflammation and controls immune homeostasis in physiological and pathological conditions such as infection and autoimmunity [43]. *SOCS3* is a hub gene in the module lightgreen associated with the Jak-STAT signaling pathway, an important pathway for *Salmonella* infection [44]. It is well known that the Jak-STAT pathway can regulate cell growth, apoptosis, immunity, and inflammatory responses and because of its significance in the immune response, the Jak-STAT pathway is often exploited by pathogens [45]. In our study, we found that the Jak-STAT pathway genes were

significantly enriched in the module lightgreen which is not detectable in controls. So we think that *SOCS3* and the other Jak-STAT pathway genes may together regulate the activity of the organism in infection, which leads this module to be differentially coexpressed.

The above results showed some specific subnetworks for infection, in spite of a common network existing whether in the control or infected group. We constructed two coexpression networks from the top ten hub genes' expression profiles for each module from the two different conditions. As shown in Figures 6(a) and 6(b), common core networks, including the most preserved modules between infected and control groups, were present. However, some closely interacted subnetworks seen during infection disappeared in the control. These infection-specific subnetworks included genes that are members of the Jak-STAT signaling pathway (module lightgreen); others associated with membrane anchoring (module lightcyan), neuroactive ligand-receptor interaction (module salmon), and lysosomal processing (module tan), which suggested that these subnetworks dysregulated the

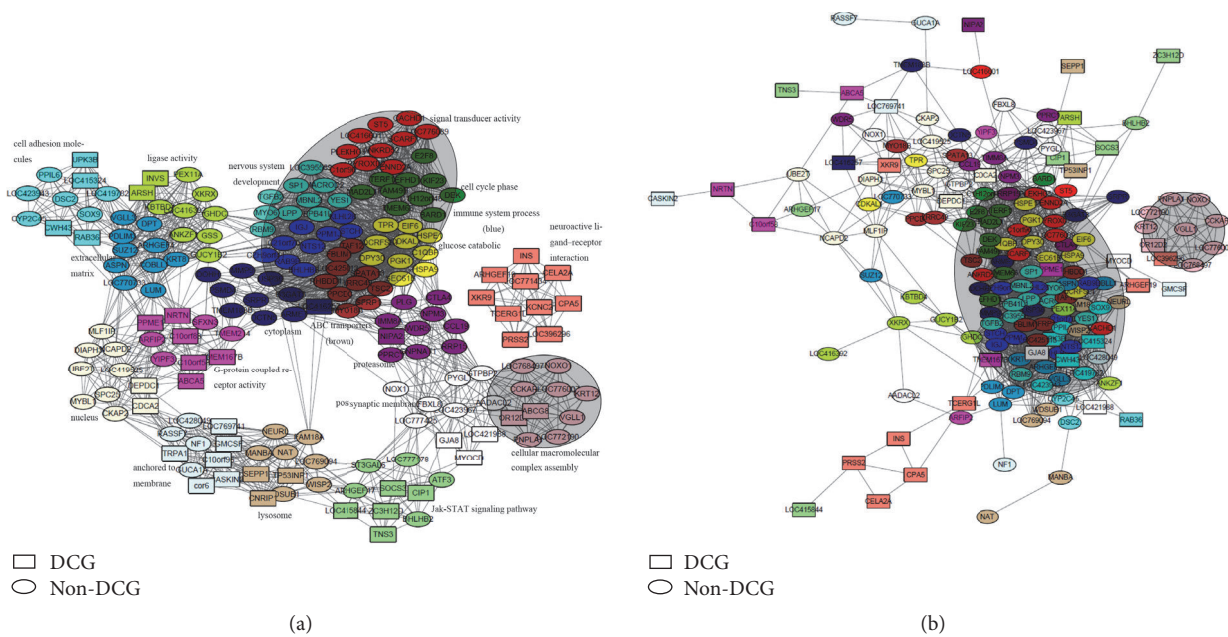


FIGURE 6: (a) Hub genes' network in the infected group. The node colors represent the module colors: the nodes with shape of rectangle are DCGs and the elliptical nodes are non-DCGs. (b) Hub genes' network in the control group. The node colors represent the module colors: the nodes with shape of rectangle are DCGs and the elliptical nodes are non-DCGs.

systems during infection. Although only one dataset was used here, due to the lack of published related microarray datasets, these present results advance our understanding of the cell biology and immunoregulatory pathways involved in *Salmonella* infection in the chicken host.

## Abbreviations

DAVID:	Database for Annotation, Visualization, and Integrated Discovery
DCEA:	Differential coexpression analysis
DCG:	Differentially coexpressed gene
DCL:	Differentially coexpressed link
DCM:	Differentially coexpressed module
DEA:	Differential expression analysis
ECM:	Extracellular matrix
FDR:	False discovery rate
GO:	Gene Ontology
GSCA:	Gene-set coexpression analysis
TOM:	Topological overlap measure
WGCNA:	Weighted gene coexpression network analysis.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Authors' Contributions

Bao-Hong Liu contributed to the design and conception of this study, conducted computational experiments, performed and interpreted data, and drafted the manuscript. Jian-Ping

Cai conceived this project and participated in its design, helped in interpreting the data, and drafted and revised the manuscript. All authors read and approved the final manuscript.

## Acknowledgments

The authors thank Professor Guan Zhu from Texas A&M University and Dr. Patricia Wilkins from CDC, USA, for revising the manuscript. And they also thank the researchers from Yangzhou University for submitting the dataset to GEO. This work was supported by Fundamental Research Program of Chinese Academy of Agricultural Sciences (0032015027) and the Innovative Special Project of Agricultural Sci-Tech and by the Special Fund for Agro-Scientific Research in Public Interest (201303044-7) to Professor Cai's team.

## References

- [1] R. T. Khan, M. Chevenon, K. E. Yuki, and D. Malo, "Genetic dissection of the *Ity3* locus identifies a role for *Ncf2* co-expression modules and suggests *Selp* as a candidate gene underlying the *Ity3.2* locus," *Frontiers in Immunology*, vol. 5, article 375, 2014.
- [2] G. C. Buckle, C. L. Walker, and R. E. Black, "Typhoid fever and paratyphoid fever: systematic review to estimate global morbidity and mortality for 2010," *Journal of Global Health*, vol. 2, no. 1, Article ID 010401, 2012.
- [3] R. L. Santos, "Pathobiology of *Salmonella*, intestinal microbiota, and the host innate immune response," *Frontiers in Immunology*, vol. 5, article 252, 2014.
- [4] C. Guadarrama, T. Villaseñor, and E. Calva, "The subtleties and contrasts of the *LeuO* regulator in *Salmonella* Typhi:

- implications in the immune response,” *Frontiers in Immunology*, vol. 5, article 581, 2014.
- [5] S. Patel and B. A. McCormick, “Mucosal inflammatory response to *Salmonella* typhimurium infection,” *Frontiers in Immunology*, vol. 5, article 311, 2014.
- [6] J. P. Mooney, B. P. Butler, K. L. Lokken et al., “The mucosal inflammatory response to non-typhoidal *Salmonella* in the intestine is blunted by IL-10 during concurrent malaria parasite infection,” *Mucosal Immunology*, vol. 7, no. 6, pp. 1302–1311, 2014.
- [7] M. M. Bellet, E. Deriu, J. Z. Liu et al., “Circadian clock regulates the host response to *Salmonella*,” *Proceedings of the National Academy of Sciences*, vol. 110, no. 24, pp. 9897–9902, 2013.
- [8] T.-H. Huang, J. J. Uthe, S. M. D. Bearson et al., “Distinct peripheral blood RNA responses to *Salmonella* in pigs differing in *Salmonella* shedding levels: intersection of IFNG, TLR and miRNA pathways,” *PLoS ONE*, vol. 6, no. 12, Article ID e28768, 2011.
- [9] A. Chaussé, O. Grépinet, E. Bouteau et al., “Susceptibility to *Salmonella* carrier-state: a possible Th2 response in susceptible chicks,” *Veterinary Immunology and Immunopathology*, vol. 159, no. 1-2, pp. 16–28, 2014.
- [10] A. Szmolka, Z. Wiener, M. E. Matulova, K. Varmuzova, and I. Rychlik, “Gene expression profiles of chicken embryo fibroblasts in response to *Salmonella* Enteritidis infection,” *PLoS ONE*, vol. 10, no. 6, Article ID e0127708, 2015.
- [11] V. van Noort, B. Snel, and M. A. Huynen, “The yeast coexpression network has a small-world, scale-free architecture and can be explained by a simple model,” *EMBO Reports*, vol. 5, no. 3, pp. 280–284, 2004.
- [12] J. M. Stuart, E. Segal, D. Koller, and S. K. Kim, “A gene-coexpression network for global discovery of conserved genetic modules,” *Science*, vol. 302, no. 5643, pp. 249–255, 2003.
- [13] H. K. Lee, A. K. Hsu, J. Sajdak, J. Qin, and P. Pavlidis, “Coexpression analysis of human genes across many microarray data sets,” *Genome Research*, vol. 14, no. 6, pp. 1085–1094, 2004.
- [14] B. Zhang and S. Horvath, “A general framework for weighted gene co-expression network analysis,” *Statistical Applications in Genetics and Molecular Biology*, vol. 4, article 17, 2005.
- [15] A. Kommadath, H. Bao, A. S. Arantes et al., “Gene co-expression network analysis identifies porcine genes associated with variation in *Salmonella* shedding,” *BMC Genomics*, vol. 15, no. 1, article 452, 2014.
- [16] Y. Choi and C. Kendzierski, “Statistical methods for gene set co-expression analysis,” *Bioinformatics*, vol. 25, no. 21, pp. 2780–2786, 2009.
- [17] H. Yu, B.-H. Liu, Z.-Q. Ye, C. Li, Y.-X. Li, and Y.-Y. Li, “Link-based quantitative methods to identify differentially coexpressed genes and gene Pairs,” *BMC Bioinformatics*, vol. 12, article 315, 2011.
- [18] B.-H. Liu, H. Yu, K. Tu, C. Li, Y.-X. Li, and Y.-Y. Li, “DCGL: an R package for identifying differentially coexpressed genes and links from gene expression microarray data,” *Bioinformatics*, vol. 26, no. 20, pp. 2637–2638, 2010.
- [19] J. Yang, H. Yu, B.-H. Liu et al., “DCGL v2.0: an R package for unveiling differential regulation from differential co-expression,” *PLoS ONE*, vol. 8, no. 11, Article ID e79729, 2013.
- [20] G. K. Smyth, “Linear models and empirical Bayes methods for assessing differential expression in microarray experiments,” *Statistical Applications in Genetics and Molecular Biology*, vol. 3, article 3, 2004.
- [21] P. Langfelder and S. Horvath, “WGCNA: an R package for weighted correlation network analysis,” *BMC Bioinformatics*, vol. 9, article 559, 2008.
- [22] Y. Benjamini, D. Drai, G. Elmer, N. Kafkafi, and I. Golani, “Controlling the false discovery rate in behavior genetics research,” *Behavioural Brain Research*, vol. 125, no. 1-2, pp. 279–284, 2001.
- [23] M. Ashburner, C. A. Ball, J. A. Blake et al., “Gene ontology: tool for the unification of biology,” *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.
- [24] D. W. Huang, B. T. Sherman, and R. A. Lempicki, “Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources,” *Nature Protocols*, vol. 4, no. 1, pp. 44–57, 2009.
- [25] P. Shannon, A. Markiel, O. Ozier et al., “Cytoscape: a software environment for integrated models of biomolecular interaction networks,” *Genome Research*, vol. 13, no. 11, pp. 2498–2504, 2003.
- [26] K.-I. Uchiya and T. Nikai, “*Salmonella* pathogenicity island 2-dependent expression of suppressor of cytokine signaling 3 in macrophages,” *Infection and Immunity*, vol. 73, no. 9, pp. 5587–5594, 2005.
- [27] A. Berndt, J. Müller, L. Borsi, H. Kosmehl, U. Methner, and A. Berndt, “Reorganisation of the caecal extracellular matrix upon *Salmonella* infection—relation between bacterial invasiveness and expression of virulence genes,” *Veterinary Microbiology*, vol. 133, no. 1-2, pp. 123–137, 2009.
- [28] D. G. Guiney and M. Lesnick, “Targeting of the actin cytoskeleton during infection by *Salmonella* strains,” *Clinical Immunology*, vol. 114, no. 3, pp. 248–255, 2005.
- [29] W. Liu and H. Ye, “Co-expression network analysis identifies transcriptional modules in the mouse liver,” *Molecular Genetics and Genomics*, vol. 289, no. 5, pp. 847–853, 2014.
- [30] J. A. Miller, M. C. Oldham, and D. H. Geschwind, “A systems level analysis of transcriptional changes in Alzheimer’s disease and normal aging,” *Journal of Neuroscience*, vol. 28, no. 6, pp. 1410–1420, 2008.
- [31] Y. Zhai, L. M. Franco, R. L. Atmar et al., “Host transcriptional response to influenza and other acute respiratory viral infections—a prospective cohort study,” *PLoS Pathogens*, vol. 11, no. 6, Article ID e1004869, 2015.
- [32] Z. Chen, D. Borek, S. B. Padrick et al., “Structure and control of the actin regulatory WAVE complex,” *Nature*, vol. 468, no. 7323, pp. 533–538, 2010.
- [33] L.-M. Chen, S. Hobbie, and J. E. Galán, “Requirement of CDC42 for *Salmonella*-induced cytoskeletal and nuclear responses,” *Science*, vol. 274, no. 5295, pp. 2115–2118, 1996.
- [34] K. Hallstrom and B. A. McCormick, “*Salmonella* interaction with and passage through the intestinal mucosa: through the lens of the organism,” *Frontiers in Microbiology*, vol. 2, p. 88, 2011.
- [35] A. Pandey, I. Dan, T. Z. Kristiansen et al., “Cloning and characterization of PAK5, a novel member of mammalian p21-activated kinase-II subfamily that is predominantly expressed in brain,” *Oncogene*, vol. 21, no. 24, pp. 3939–3948, 2002.
- [36] C. Dan, N. Nath, M. Liberto, and A. Minden, “PAK5, a new brain-specific kinase, promotes neurite outgrowth in N1E-115 cells,” *Molecular and Cellular Biology*, vol. 22, no. 2, pp. 567–577, 2002.
- [37] G. Joberty, R. R. Perlungher, and I. G. Macara, “The Borgs, a new family of Cdc42 and TC10 GTPase-interacting proteins,” *Molecular and Cellular Biology*, vol. 19, no. 10, pp. 6585–6597, 1999.

- [38] A. S. Alberts, N. Bouquin, L. H. Johnston, and R. Treisman, "Analysis of RhoA-binding proteins reveals an interaction domain conserved in heterotrimeric G protein  $\beta$  subunits and the yeast response regulator protein Skn7," *Journal of Biological Chemistry*, vol. 273, no. 15, pp. 8616–8622, 1998.
- [39] B. Zhang, J. Chernoff, and Y. Zheng, "Interaction of Rac1 with GTPase-activating proteins and putative effectors. A comparison with Cdc42 and RhoA," *The Journal of Biological Chemistry*, vol. 273, no. 15, pp. 8776–8782, 1998.
- [40] G. Joberty, C. Petersen, L. Gao, and I. G. Macara, "The cell-polarity protein Par6 links Par3 and atypical protein kinase C to Cdc42," *Nature Cell Biology*, vol. 2, no. 8, pp. 531–539, 2000.
- [41] Y. Noda, R. Takeya, S. Ohno, S. Naito, T. Ito, and H. Sumimoto, "Human homologues of the *Caenorhabditis elegans* cell polarity protein PAR6 as an adaptor that links the small GTPases Rac and Cdc42 to atypical protein kinase C," *Genes to Cells*, vol. 6, no. 2, pp. 107–119, 2001.
- [42] S. Brill, S. Li, C. W. Lyman et al., "The Ras GTPase-activating-protein-related human protein IQGAP2 harbors a potential actin binding domain and interacts with calmodulin and Rho family GTPases," *Molecular and Cellular Biology*, vol. 16, no. 9, pp. 4869–4878, 1996.
- [43] B. Carow and M. E. Rottenberg, "SOCS3, a major regulator of infection and inflammation," *Frontiers in Immunology*, vol. 5, article 58, 2014.
- [44] X. Liu, R. Lu, Y. Xia, S. Wu, and J. Sun, "Eukaryotic signaling pathways targeted by *Salmonella* effector protein AvrA in intestinal infection in vivo," *BMC Microbiology*, vol. 10, p. 326, 2010.
- [45] L. Yang and Y.-J. Zhang, "Antagonizing cytokine-mediated JAK-STAT signaling by porcine reproductive and respiratory syndrome virus," *Veterinary Microbiology*, 2016.