Research Article

Potential Genes and Pathways of Neonatal Sepsis Based on Functional Gene Set Enrichment Analyses

YuXiu Meng,[1] **Xue Hong Cai,**[2] **and LiPei Wang** ⓘ[1]

[1]*Department of Neonatology, First People's Hospital of Jining, Jining, Shandong 272000, China*
[2]*Department of Pediatrics, Traditional Chinese Medicine Hospital of Yanzhou, Jining, Shandong 272100, China*

Correspondence should be addressed to LiPei Wang; wanglipeijining@yeah.net

Background. Neonatal sepsis (NS) is considered as the most common cause of neonatal deaths that newborns suffer from. Although numerous studies focus on gene biomarkers of NS, the predictive value of the gene biomarkers is low. NS pathogenesis is still needed to be investigated. *Methods.* After data preprocessing, we used KEGG enrichment method to identify the differentially expressed pathways between NS and normal controls. Then, functional principal component analysis (FPCA) was adopted to calculate gene values in NS. In order to further study the key signaling pathway of the NS, elastic-net regression model, Mann–Whitney U test, and coexpression network were used to estimate the weights of signaling pathway and hub genes. *Results.* A total of 115 different pathways between NS and controls were first identified. FPCA made full use of time-series gene expression information and estimated F values of genes in the different pathways. The top 1000 genes were considered as the different genes and were further analyzed by elastic-net regression and MWU test. There were 7 key signaling pathways between the NS and controls, according to different sources. Among those genes involved in key pathways, 7 hub genes, PIK3CA, TGFBR2, CDKN1B, KRAS, E2F3, TRAF6, and CHUK, were determined based on the coexpression network. Most of them were cancer-related genes. PIK3CA was considered as the common marker, which is highly expressed in the lymphocyte group. Little was known about the correlation of PIK3CA with NS, which gives us a new enlightenment for NS study. *Conclusion.* This research might provide the perspective information to explore the potential novel genes and pathways as NS therapy targets.

1. Introduction

Neonatal sepsis is the most prevalent cause of death of the neonates with few certainly reported biomarkers for many years. At least 35% neonatal deaths were caused by infections each year. The neonates have usually suffered from early-onset NS, which occurs within the first 72 hours after birth [1, 2]. According to the recent report, the diagnosis of sepsis is humbled by the nonspecific and highly variable human inflammatory and anti-inflammatory processes [3]. The main risk factor that causes the neonatal death is infections, which include respiratory infections, drug-resistant infections, and neonatal tetanus [4]. In order to manage the infections, the research to develop primary and secondary prevention strategies based on different kinds of infections has been a hot field for NS study in recent decades [5, 6]. Future medical research should be based on reducing the

application and duration of antibiotics for NS. In view of the side effects caused by the treatment of NS, it is important to make the division for using the right standard of practice for the vulnerable group. The current classification criteria for susceptible populations are crucial to future research and to improve the development of neonatal management strategies. Medzhitov et al. found that the Toll-like receptor-2 and Toll-like receptor-4 are involved in the recognition process of the bacteria in neonates [7]. Septic neonates have a significant upregulation and obvious decline of several genes, which involved in innate immunity [8, 9]. The neonatal innate immune response to sepsis is driven by innate immunity genes (IL1R2, ILRN, and SOCS3) [10]. Current studies also investigated the relationship between the cytokine pattern and onset of NS and proved that the increased expression of proinflammatory cytokines, such as TNF-alpha, IL-6, and IL-10,

was associated with the acute and post-acute phase of NS, respectively [11].

Despite the numerous studies of NS pathogenesis based on genes, the valuable predictors have remained unveiled and contributed to being a major challenge to the research of NS. Gene transcriptomic profiles can be used to identify diagnostic and prognostic gene signatures in complex diseases and to reveal the pathogenesis of NS [9, 10]. Several systems biology approaches were built to dissect the physiological mechanism of sepsis. In particular, methods for discovering the context-specific activations of pathways [12, 13] were merged. However, for the time-course studies, it will be difficult to do clinical trials if the role of the gene of choice is not specific to the biological process of interest. In other words, a temporally differentially expressed gene should show a significant nonconstant expression pattern across time points. To address this, weighting methods were needed to be carried out to assess the functional similarities between a given gene and the sets in different time points [14].

In this report, a method based on the functional principal component analysis (FPCA) was proposed to discover arbitrary nonconstant trends in time-course data analysis [15]. After estimating the impact of the gene, an elastic-net regression model was used to analyze the weights of genes. Besides, a generalized Mann–Whitney U (MWU) test was also applied for gene set-level inferences. Finally, hub genes were determined by the topological feature of coexpression networks [16]. Using the proposed analysis method, susceptible pathways and crucial genes will be revealed. And they will facilitate the future investigation of NS.

2. Methods

2.1. Data Recruitment and Preprocess. Gene expression profiles of human peripheral blood cells at various time points from samples of meningococcal sepsis were deposited at Gene Expression Omnibus database with the data accession no. GSE11755, including NS patients and normal controls. These datasets were processed on Affymetrix Human Genome U133 Plus 2.0 Array platform. Totally, forty-one samples, which were drawn at four time points ($t = 0$, $t = 8$, $t = 24$, and $t = 72$ h after admission to the paediatric intensive care unit), were studied, and key pathways and hub genes were also identified. Next, based on the RNA microarray, gene expressions isolated from whole blood, lymphocytes, and monocytes were also analyzed, respectively. According to the different sources of microarray data, we adopted different groups: The first, we named All Sources, contained all the 41 samples (10 controls and 31 patients). The second, we named Blood Source, contained the microarray data derived from blood (3 controls and 8 patients). The third, we named Lymphocyte Source, contained the microarray data derived from lymphocytes (4 controls and 12 patients). The fourth, we named Monocyte Source, contained the microarray data derived from monocytes (3 controls and 11 patients). In the following analyses, we conducted four parallel analyses based on different groups. The study was approved by the local medical ethics committee.

For data preprocessing, a freely available R platform (http://cran.r-project.org/) was applied. GraphPad Prism 7.0 software was used to create images. And data preprocess of dataset was commenced with reading the data by the standard method carried out by Affy. Expressions of genes were normalized using the robust multiarray average (RMA) method, in order to eliminate the influence of nonspecific hybridization [17]. And then, genes were further filtered by quartile-based algorithm [18]. A total of 15144 genes were reported for each subject.

2.2. Pathway Enrichment Analysis. Pathway analysis was used to find the significant pathways of the NS and control groups according to Kyoto Encyclopedia of Genes and Genomes (KEGG) [19]. Fisher's exact test was adopted to select the significant pathways, and the threshold of significance was defined by FDR and p value. Significant pathways were extracted according to the thresholds of $p < 0.05$ and intersection gene count >1.

2.3. A Gene-Level Summary Statistic by the Functional Principal Component Analysis. In the present research, the FPCA model was used to identify temporally differentially expressed genes [20]. The gene expression profile obtained was assumed to be the scattered members from the true profile of gene expression. And the true profile will be further interfered by noisy signals. After subtracting the average expression value of genes, FPCA was used to center all the gene values. The gene expression profile of preprocessed data was weighted according to their corresponding mean expression and FPCA score across all the gene expression values.

The observed expression using the FPCA model is as follows:

$$\widehat{X}_i\,(t) = \widehat{\mu}_i + \sum_{l=1}^{L} \widehat{\xi}_{il}\widehat{\Phi}_l(t), \tag{1}$$

where $\widehat{\mu}_i$ is the average expression of the temporal sample, $\widehat{\Phi}_l(t)$ is the lth eigenfunction, and $\widehat{\xi}_{il}$ is the FPC value that quantifies how much $\widehat{X}_i(t)$ can be explained by $\widehat{\Phi}_l(t)$.

When it applied to the time-course gene expression, we used functional F-statistic to summarize the gene pattern information for each gene in the time points:

$$F_i = \frac{\mathrm{RSS}_i^0 - \mathrm{RSS}_i^1}{\mathrm{RSS}_i^1}, \tag{2}$$

where RSS_i^0 is the residual sum of squares of null hypotheses and RSS_i^1 is the residual sum of squares of alternative hypotheses. F_i can be viewed as a "signal-to-noise" ratio and revealed the importance of genes.

2.4. Estimating the Weights of Signaling Pathway Using the Elastic-Net Regression Model. In this study, we also took an approach with computationally efficient and highly flexible methods on the basis of an equivalent influence between the penalty function regression and a standard multivariate

TABLE 1: Top 6 differentially expressed pathways according to the KEGG analysis.

Pathway_name	p value	FDR	Gene count
Hsa04740: Olfactory transduction	$1.25E - 138$	$3.59E - 136$	59
Hsa05206: MicroRNAs in cancer	$1.07E - 23$	$1.53E - 21$	133
Hsa04080: Neuroactive ligand-receptor interaction	$2.84E - 10$	$2.03E - 08$	179
Hsa04110: Cell cycle	$2.42E - 10$	$2.03E - 08$	117
Hsa04380: Osteoclast differentiation	$1.45E - 09$	$8.27E - 08$	122
Hsa00830: Retinol metabolism	$2.47E - 09$	$1.0E - 07$	23

regression, in order to minimize optimization problem, which is known as the functional elastic-net regression problem [14]. This problem occurs because of the model selection methods in a functional linear regression model that is needless for the concurrent function regression.

The main function of the model is as follows:

$$\widehat{\beta}_i = \min_{\beta_i} \text{OBJ}\big(\beta_i | x_i(t), \widehat{\Phi}_i(t)\big)$$

$$\text{OBJ}\big(\beta_i | x_i(t), \widehat{\Phi}_i(t)\big) = \left\| x_i(t) - \widehat{\Phi}_i(t)^T \beta_i \right\|^2 + \lambda_1 \|\beta_i\|_1$$

$$+ \lambda_2 \|\beta_i\|^2,$$

(3)

where λ is the penalty coefficient and β_i is the vector of the set of linear coefficients. When $\widehat{\beta}_i$ is calculated and estimated, then the weights of the pathways can be obtained by

$$\widehat{\mathcal{W}}_{i,k} := \frac{\sum_{l=1}^{L}\left(\widehat{\beta}_{l,i}^k\right)^2}{\sum_{k \in K_i}\sum_{l=1}^{L}\left(\widehat{\beta}_{l,i}^k\right)^2}.$$

(4)

A similar approach can be used to estimate the weights of genes.

2.5. Weighted Mann–Whitney U (MWU) Test with Correlation Using Gene Set Enrichment Analysis (GSEA). MWU test is used to compare two independent samples. Given that two samples were exactly from the same groups, the mean was different. The aim of the MWU test was to analyze whether there was a significant difference between the means of the two groups. Recent reports showed that MWU test plays an important role in gene set enrichment analysis (GSEA) [21, 22]. The pathway enrichment analysis was carried out based on the genome-wide background and was applied to identify the biological functions of the significant clusters. KEGG pathway enrichment was also performed. Categories with more than 5 genes were presented, and p value < 0.01 were considered significant in pathway enrichment analysis [23].

2.6. Identification of Hub Genes Based on the Coexpression Networks. Adjacency matrixes were firstly constructed based on the intergenomic relationships evaluated by Spearman correlation coefficient [24]. Topological features were further studied to find key nodes in the network. Genes whose degree was greater than the average degree values and whose Spearman correlation coefficient was greater than 0.6 were considered as hub genes.

3. Results

3.1. Pathway Enrichment Analysis. Gene expression profile of human NS with the series of GSE11755 was downloaded from Gene Expression Omnibus. After preprocessing the expression profile data of the dataset, we collected data from a total of 41 samples, including six children with meningococcal sepsis. Blood was drawn at four time points and matched with controls. Pathway enrichment analysis of NS and controls was conducted on the basis of the KEGG pathway database. A total of 286 pathways covering 6893 genes were obtained. After Fisher's exact test, 115 differential pathways covering 3532 genes met the thresholds of $p < 0.05$ and intersection gene count > 1. Table 1 shows the top 6 differential signaling pathways in ascending order based on p value.

3.2. Integrated Analysis of Gene Signatures Using the FPCA Model. In the present research, the FPCA model was used to identify temporally differentially expressed genes and each gene would get an F value. Based on the 115 differential pathways (covering 3532 genes), we identified top 1000 gene signatures of NS using FPCA model, which were defined as dysregulated genes. FPCA narrowed the gene search range from 3532 to 1000. Greater F value means that the expression level differed greatly with others. Figures 1(a)–1(d) show the curve of gene signatures with F value. Among the dysregulated genes, the top 12 genes from All Sources, Blood Source, Lymphocyte Source, and Monocyte Source were CDC37, NCOA2, P2RY12, RXRB, EDEM2, ACTN4, STX12, PPM1A, PRKACB, DUSP10, VEGFA, and SLC44A2. Since NS is mainly caused by infections, the dysregulated genes in NS should be immune response related. However, there were few genes in the list that were immune related. Activation of the cytokines in a specific infection might not be derived from all the regulated genes that can activate those genes. Therefore, it is important to find the pathways which are activated by infections in NS. FPCA could effectively utilize the time-series information and overcome the traditional control design deficiencies [14]. F values would be used for the MWU test.

3.3. Estimating the Weights of Genes Using the Elastic-Net Regression Model. Genes that exist in multiple pathways were considered as overlapping genes. These genes are thought to play multiple roles in hypothesis testing, where the weight coefficients were overestimated. In the present study, elastic-net regression model was used to decompose
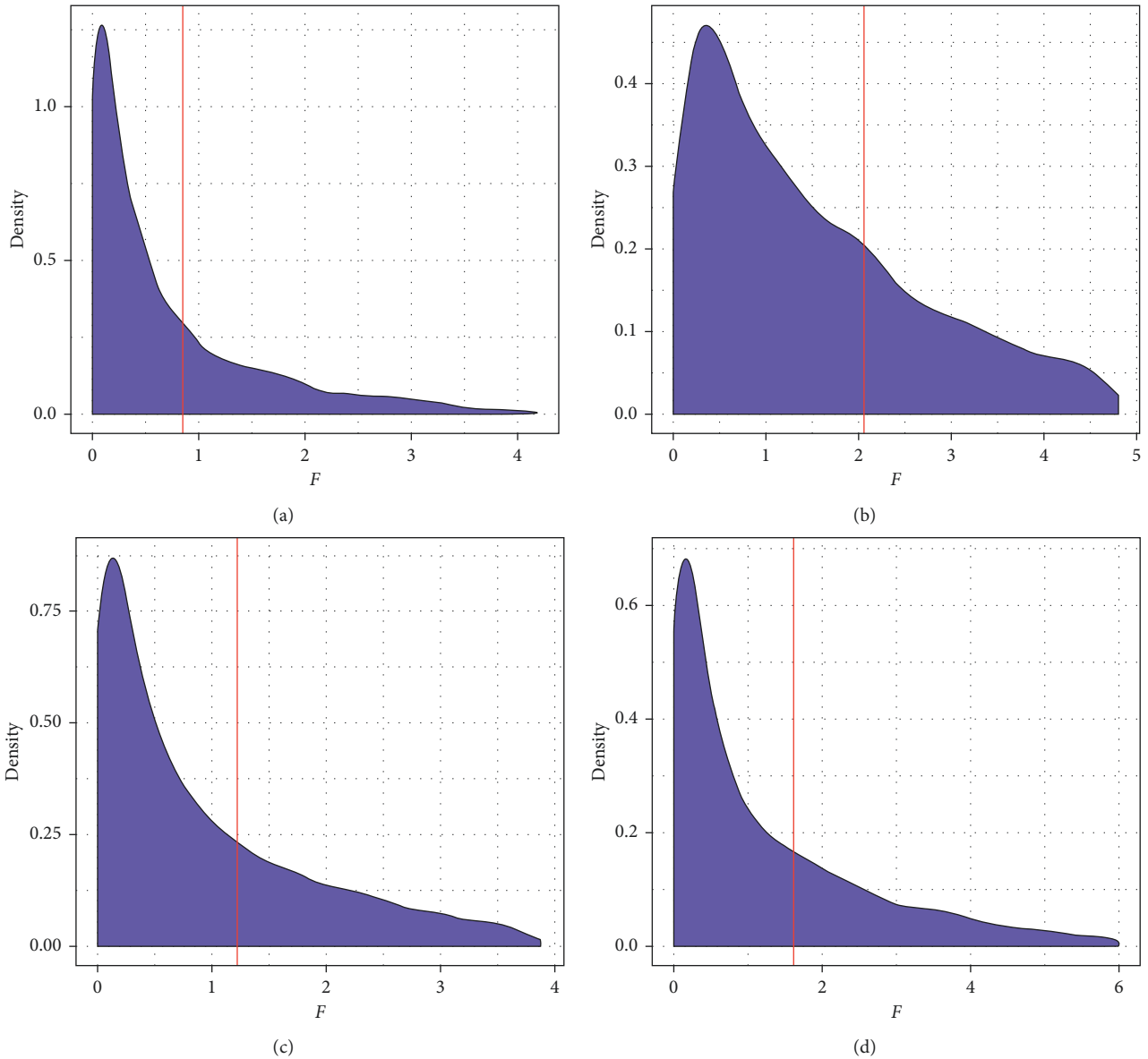
(a)

(b)

(c)

(d)

FIGURE 1: The distribution of F value of pathway genes. Time-series gene signatures data were analyzed by FPCA and each gene obtained an F value (x-coordinate, F value). Y-axis represents gene density. The genes were ranked in the order of F value, and the top 1000 of them were selected. The red line represents the threshold of top 1000 genes. (a) All Sources, (b) Blood Source, (c) Lymphocyte Source, and (d) Monocyte Source.

an overlapping gene between gene sets and eliminate the overlapping effects. After calculating the weight value of each gene and adding the weight values of the pathway genes, the total weight value of the pathways was obtained. Figure 2 shows the sum weight of each pathway. The weight value (w) of each gene would be used for the MWU test.

3.4. Functional Enrichment Analysis Using GSEA and MWU Model Test.
Based on the KEGG pathway enrichment, 115 differential pathways were obtained. In order to more accurately find key pathways and molecules, FPCA and elastic-net regression were performed to eliminate overlapping gene effects. Combined with the MWU test, key molecular

pathways in the gene transcription data of NS and controls were identified. Based on the t-test, pathways were ranked in the descending order. After the pathway data were tested by the MWU model, a total of 7 pathway terms met the condition p values < 0.05. There was no pathway met the conditions in the monocyte group. The resulting pathways are presented in Table 2.

According to the MWU test, there were 7 pathways were screened based on the p values < 0.05. We selected the top 3 significant pathways: hsa05220: Chronic myeloid leukemia; hsa04380: Osteoclast differentiation; and hsa05222: Small-cell lung cancer for further analysis. Besides, pathways including the proinflammatory cytokine genes were also studied, such as hsa05164: Influenza A (TNF, IL-6, IL-18,
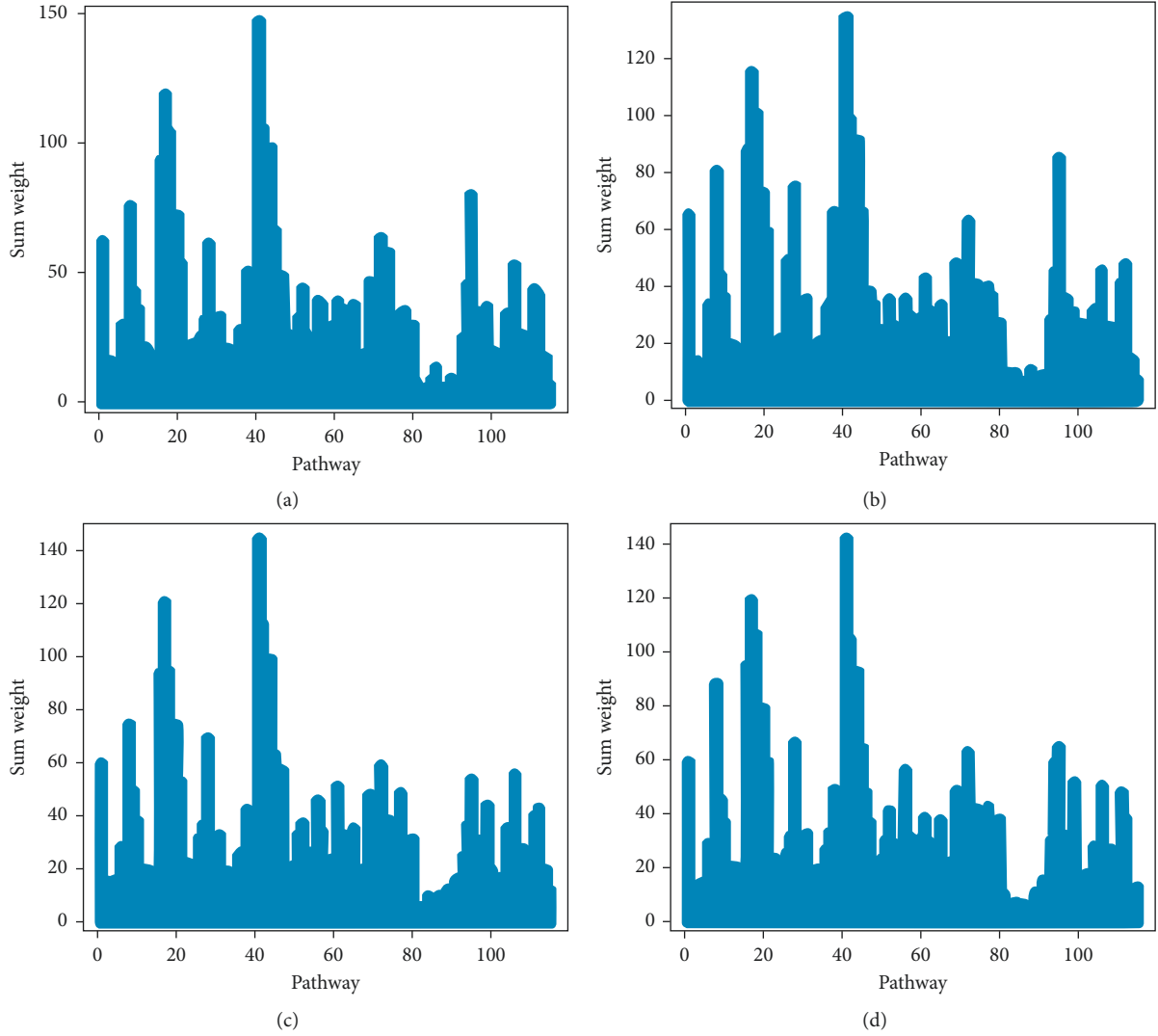
FIGURE 2: Sum weights of 115 differential pathways. *Y*-axis represents the sum weights of pathways. *X*-axis represents the number of pathways. (a) All Sources, (b) Blood Source, (c) Lymphocyte Source, and (d) Monocyte Source.

TABLE 2: *p* values of MWU test of sample groups.

Groups	KEGG pathways	*p* values
All Sources	hsa05220: Chronic myeloid leukemia	0.0457024
Blood Source	hsa05120: Epithelial cell signaling in *Helicobacter pylori* infection	0.0357933
	hsa04380: Osteoclast differentiation	0.0380088
	hsa04666: Fc gamma R-mediated phagocytosis	0.0415344
Lymphocyte Source	hsa05222: Small-cell lung cancer	0.0150380
	hsa04660: T cell receptor signaling pathway	0.0412070
	hsa05219: Bladder cancer	0.0463346
Monocyte Source	None	

hsa: *Homo sapiens* (human); *p* values < 0.05 significant difference.

and IFNA1; *p* value 0.3256); hsa04620: Toll-like receptor signaling pathway (TNF, IL-6, and IFNA1; *p* value 0.2185); hsa05168: Herpes simplex infection (TNF, IL-6, IFNA1, and IL-15; *p* value 0.4868). Unfortunately, the MWU test showed that there was no difference between the controls and patients in those proinflammatory cytokines included pathways. For the obtained genes in the top 3 pathways, Figure 3(a) reveals that the expression change of hsa05220: Chronic myeloid leukemia from All Sources was not obvious. The levels of hsa05120: Epithelial cell signaling in *Helicobacter pylori* infection from Blood Source after admission to the paediatric intensive care unit were significantly higher than control. Besides, the levels of hsa05222: Small-cell lung cancer from Lymphocyte Source were up at 72 h after admission to the paediatric intensive care unit.

3.5. Identification and Estimation of the Weights of the Hub Genes in the Pathway.

Networks provide effective models to study complex biological systems, such as gene and protein interaction networks. A weighted gene coexpression network was constructed using adjacency matrix based on superman coefficient. We further studied the topological
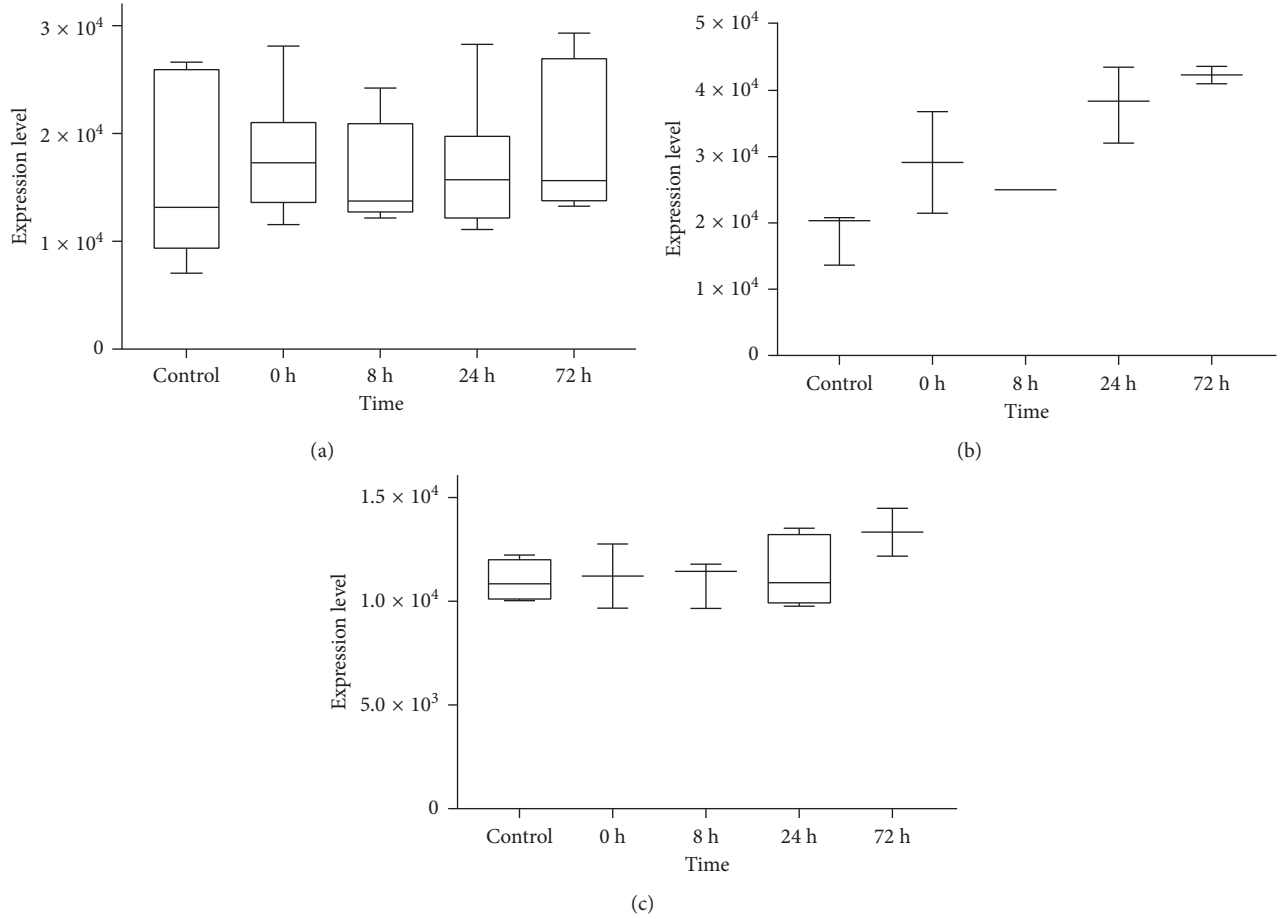
(a)

(b)

(c)

FIGURE 3: Expression levels of the top 3 significant signaling pathways. (a) hsa05220: Chronic myeloid leukemia from All Sources, (b) hsa05120: Epithelial cell signaling in *Helicobacter pylori* infection from Blood Source, and (c) hsa05222: Small-cell lung cancer from Lymphocyte Source. *Y*-axis represents expression levels of pathways. *X*-axis represents control and several time points after admission to the paediatric intensive care unit. The graphs were made with GraphPad Prism 7.0.

features to find key nodes in the networks. Genes whose degree was greater than the average degree values were considered as hub genes. Based on the three networks of 7 pathways from All Sources, Blood Source, and Lymphocyte Source, we mapped a Venn diagram. Figure 4 shows that the intersection of these three sets contained only one gene, PIK3CA. We defined PIK3CA as the common marker of NS. The intersection of All Sources and Blood Source had two genes, namely, PIK3CA and TGFBR2. A total of 4 genes (PIK3CA, CDKN1B, KRAS, and E2F3) existed in All Sources and Lymphocyte Source sets, simultaneously. There were 3 genes that shared in both Blood Sources and Lymphocyte Source: PIK3CA, TRAF6, and CHUK.

3.6. Expression Levels of Hub Genes and Common Inflammatory Factors. After analyzing the topological features of networks based on 7 pathways, 7 genes were considered as the hub genes which were described above. In order to investigate the relevance of the hub genes and NS, the expression levels of PIK3CA, TGFBR2, CDKN1B, KRAS, E2F3, TRAF6, and CHUK were further analyzed. As we all know that NS is mainly caused by infections, the levels of

proinflammatory genes were also observed, such as tumor necrosis factor alpha (TNF-α), interleukin-2 (IL-2), interleukin-6 (IL-6), interleukin-7 (IL-7), interleukin-10 (IL-10), and interferon alpha-1 (IFNA1). Besides, we examined expressions of housekeeping genes GAPDH and beta-catenin (not shown here) aiming at objectively reflecting the changes in hub genes. Figure 5 shows the expression levels of these genes in Box-whisker plot. We can easily found that PIK3CA levels from common and blood groups of patients after admission to the paediatric intensive care unit had no obvious changes compared with controls, while expression of PIK3CA from Lymphocyte Source significantly decreased. According to the reports, NS is mainly caused by infections; however, there was no significant difference between controls and patients in immune response-related gene expression levels (Figures 5(d)–5(i)). The expressions of CDKN1B, KRAS, E2F3, TRAF6, and CHUK were not displayed here.

4. Discussion

In recent years, many new mathematical model methods such as high-dimensional differential equations [25, 26], dynamic Bayesian network [27, 28], and Granger's model [29]
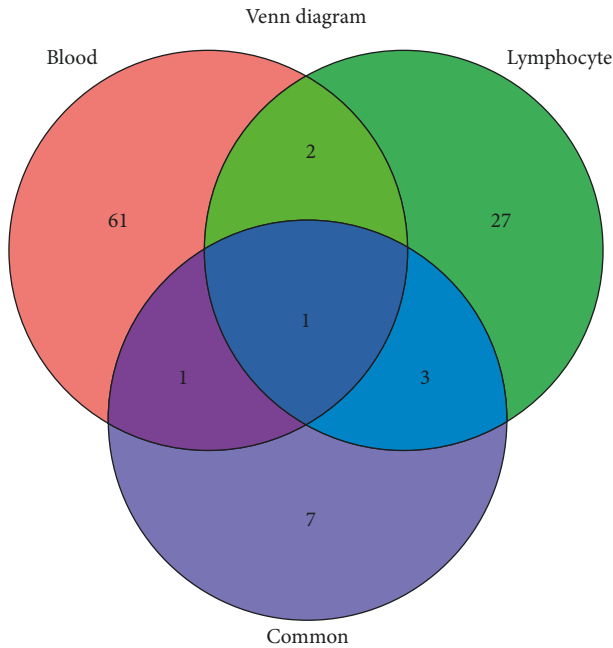
Figure 4: Venn diagram of hub genes based on the coexpression networks. Venn diagram showing the number of hub genes obtained from All Sources (Blue), Blood Source (Red), and Lymphocyte Source (Green).

were widely used in molecular biology and bioinformatics. According to reports, inchoate changes in gene expression underlying diseases or infections could be calculated by mathematical models. Low et al. [27, 28] used them to analyze the temporal causality between genes on account of changes expressed at many time points. Time-series gene expression experiments are getting more and more popular. This method plays an important role in studying translation and gene regulation. We provide a flexible way to detect common expression patterns in the individual subjects. Elastic-net regression model combined with the MWU test was used in this study. According to this method, both individual gene and gene set changes, which are induced by infection in a subject-specific way, will be detected.

In the classic MWU test, each variable is independent and there is no relationship between them. However, genes are interrelated, in particular within the related signaling pathway. Therefore, we must make some amendments to the classic MWU test, which can be used to accommodate with gene correlation. In our method, we assume that genes in the relative signaling pathway share a common pairwise correlation q and the irrelevant genes maintain independence. In the current study based on KEGG enrichment of gene signatures, the results showed that, among several KEGG pathways, the top 3 significant pathways were hsa05220: Chronic myeloid leukemia, hsa04380: Osteoclast differentiation, and hsa05222: Small-cell lung cancer, respectively. Besides, pathways including the proinflammatory cytokine genes were also studied, such as hsa05164: Influenza A, hsa04620: Toll-like receptor signaling pathway, and hsa05168: Herpes simplex infection. Unfortunately, the MWU test showed that there was no difference between the

control and common groups in those proinflammatory cytokines included pathways. In order to determine the hub genes, based on the topological characteristics of coexpression networks, PIK3CA was defined as the common marker of NS. Then, TGFBR2, CDKN1B, KRAS, E2F3, TRAF6, and CHUK were also selected as our target molecules.

PIK3CA, an oncogene, encodes the p110 catalytic subunit of class I phosphatidylinositol 3-kinases (PI3Ks), namely, PI3Kp110a. Approximately 4/5 of the mutations in PIK3CA occur in the two hot spots, exon 9 and exon 20. Its mutation not only can reduce the apoptosis of cells but also can promote the infiltration of tumors and increase the activity of its downstream kinase PI3Ks [30]. Under physiological conditions, PIK3CA is expressed in brain, lung, mammary gland, gastrointestinal tract, cervix, and other tissues and has many important physiological functions such as regulation of somatic cell proliferation, differentiation, and survival. PIK3CA is often inactive and usually not easily detected. However, PIK3CA was overexpressed after mutation, which could increase the catalytic activity of PI3Ks and promote cell canceration in tissues. PIK3CA mutation has become the molecular biomarker of many tumors [31–34]. PI3K-Akt-mTOR signaling is associated with the balance between cell proliferation and survival and plays a major role not only in tumor growth but also in the potential response of cancer treatment, such as wortmannin and LY294002 [35, 36]. Unfortunately, it seems that there is no direct correlation between PIK3CA and NS in the existing literature.

TGFBR2, transforming growth factor, beta receptor II, is a tumor suppressor gene. The encoded protein is a transmembrane protein that has a protein kinase domain, forms a heterodimeric complex with another receptor protein, and binds TGF-beta. Heterozygous mutations in TGFBR2 play an important role in Marfan syndrome, which is an extracellular matrix disorder with cardinal manifestations in the eye, skeleton, and cardiovascular systems [37]. Several recent reports showed that inducible ablation of TGF-β receptor type 2 signaling was able to limit hepatic stellate cells and fibrosis and attenuates tumor-associated inflammation [38]. TGF-β acts as a key regulator of immune cells, epithelium, in inflammatory bowel disease [39]. Many studies have shown that TGFBR2 signaling was associated with inflammatory-related diseases. But, whether TGFBR2 and NS are related is still a mystery. CDKN1B, a cyclin-dependent kinase inhibitor 1B, can bind to and prevent the activation of cyclin E-CDK2 or cyclin D-CDK4 complexes and thus controls the cell cycle progression at G1. KRAS is a gene that acts as an on/off switch in cell signaling and controls cell proliferation. Most of the target molecules we selected were related to the cell proliferation and tumorigenesis. While very few literature studies report the correlation between them and NS. This study gave us a new enlightenment for neonatal sepsis research. However, there were some limitations to our study. Firstly, PIK3CA and other molecular biomarkers were predictive biomarkers of n, and further experimental verification should be conducted to verify our results. Besides, whether the workflow was suitable for other analysis or another database is a question.
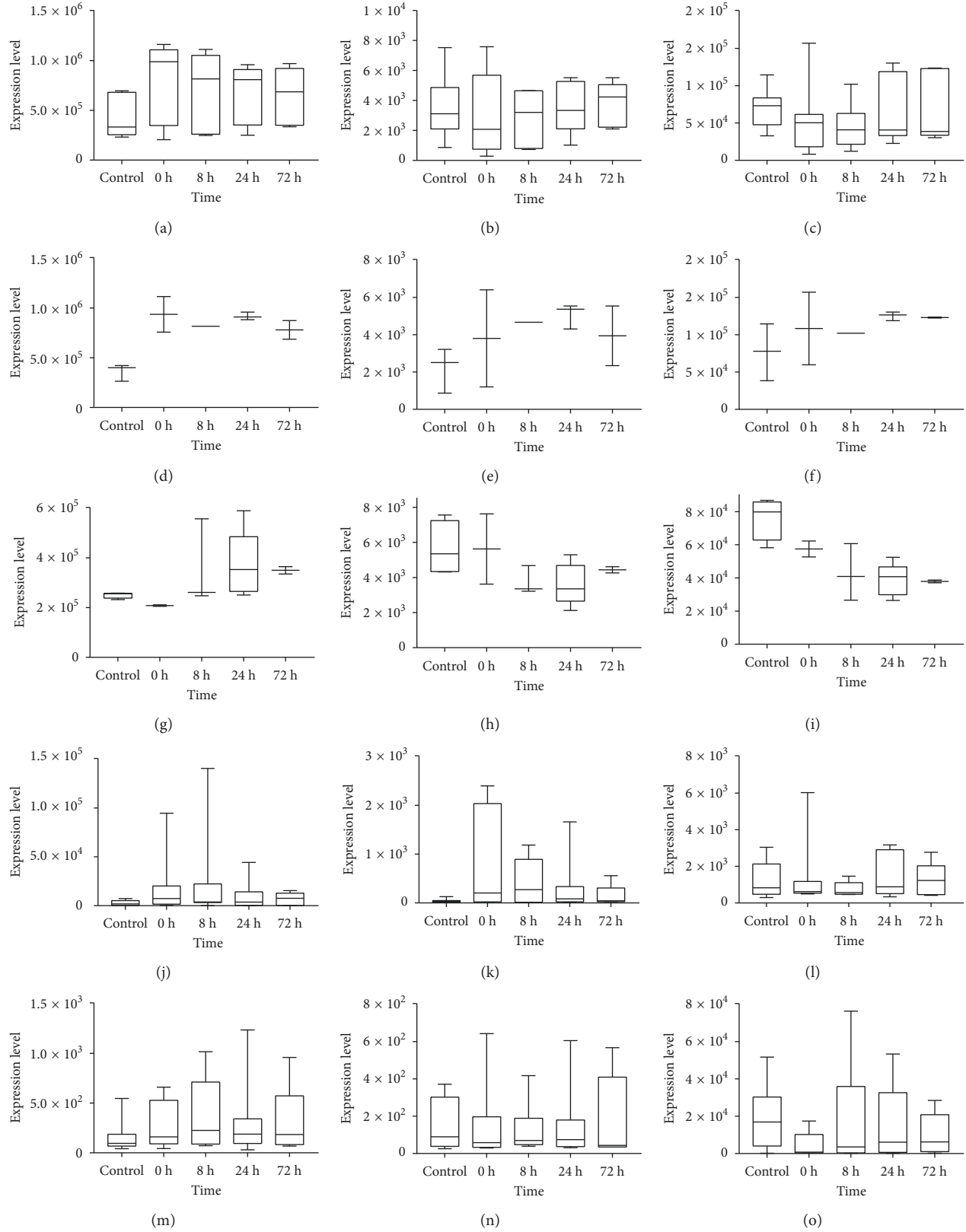
FIGURE 5: Box-whisker plot of expression levels of genes from GSE11755. (a) GAPDH (internal reference) from All Sources; (b) PIK3CA from All Sources; (c) TGFBR2 from All Sources. (d) GAPDH from Blood Source; (e) PIK3CA from Blood Source; (f) TGFBR2 from Blood Source. (g) GAPDH from Lymphocyte Source; (h) PIK3CA from Lymphocyte Source; (i) TGFBR2 from Lymphocyte Source. Levels of (j) IL-6, (k) IL-10, (l) TNF-α, (m) IL-18, (n) IL-7, and (o) IFNA1 from All Sources. Y-axis represents the expression levels of genes. X-axis represents control and several time points after admission to the paediatric intensive care unit. The box represents the express range and the central line was the median of the data. All graphs were made with GraphPad Prism 7.0.

5. Conclusion

In conclusion, a comprehensive process of data in datasets of NS was conducted in our research. Then, the function and signaling pathways of NS were presented systematically by the cutting edge models. Finally, based on the potential pathways and their topological characteristics of coexpression networks, several critical genes for NS were identified. PIK3CA was defined as the common marker of NS. However, the current study was based on the previous reports and more clinical evidence results were needed.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

YuXiu Meng and Xue Hong Cai are co-first authors.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] S. Skibsted, A. E. Jones, M. A. Puskarich et al., "Biomarkers of endothelial cell activation in early sepsis," *Shock*, vol. 39, no. 5, pp. 427–432, 2013.

[2] J. Song, D. Hu, C. He et al., "Novel biomarkers for early prediction of sepsis-induced disseminated intravascular coagulation in a mouse cecal ligation and puncture model," *Journal of Inflammation*, vol. 10, no. 1, p. 7, 2013.

[3] D. Jrovsky, I. C. Marchetti, M. A. da Silva Mori et al., "Early-onset neonatal pneumococcal sepsis: a fatal case report and brief literature review," *Pediatric Infectious Disease Journal*, vol. 37, no. 4, pp. e111–e112, 2017.

[4] A. Abbas and I. Ahmad, "First report of neonatal early-onset sepsis caused by multi-drug-resistant Raoultella ornithinolytica," *Infection*, vol. 46, no. 2, pp. 275–277, 2017.

[5] W. van Herk, S. el Helou, J. Janota et al., "Variation in current management of term and late-preterm neonates at risk for early-onset sepsis: an international survey and review of guidelines," *Pediatric Infectious Disease Journal*, vol. 35, no. 5, pp. 494–500, 2016.

[6] S. Vergnano, M. Sharland, P. Kazembe, C. Mwansambo, and P. T. Heath, "Neonatal sepsis: an international perspective," *Archives of Disease in Childhood–Fetal and Neonatal Edition*, vol. 90, no. 3, pp. F220–F224, 2005.

[7] R. Medzhitov, "Toll-like receptors and innate immunity," *Nature Reviews Immunology*, vol. 1, no. 2, pp. 135–145, 2001.

[8] J. L. Wynn and S. Guthrie, "Postnatal age is a critical determinant of the neonatal host response to sepsis," *Molecular Medicine*, vol. 21, pp. 496–504, 2015.

[9] M. Cernada, E. Serna, C. Bauerl, M. C. Collado, G. Perez-Martinez, and M. Vento, "Genome-wide expression profiles in very low birth weight infants with neonatal sepsis," *Pediatrics*, vol. 133, no. 5, pp. e1203–e1211, 2014.

[10] C. L. Smith, P. Dickinson, T. Forster et al., "Identification of a human neonatal immune-metabolic network associated with bacterial infection," *Nature Communications*, vol. 5, p. 4649, 2014.

[11] K. S. Khaertynov, S. V. Boichuk, S. F. Khaiboullina et al., "Comparative assessment of cytokine pattern in early and late onset of neonatal sepsis," *Journal of Immunology Research*, vol. 2017, Article ID 8601063, 8 pages, 2017.

[12] B. M. Hartmann, J. Thakar, R. A. Albrecht et al., "Human dendritic cell response signatures distinguish 1918, pandemic, and seasonal H1N1 influenza viruses," *Journal of Virology*, vol. 89, no. 20, pp. 10190–10205, 2015.

[13] D. Katanic, A. Khan, and J. Thakar, "PathCellNet: cell-type specific pathogen-response network explorer," *Journal of Immunological Methods*, vol. 439, pp. 15–22, 2016.

[14] Y. Zhang, D. J. Topham, J. Thakar, and X. Qiu, "FUNNEL-GSEA: FUNctioNal ELastic-net regression in time-course gene set enrichment analysis," *Bioinformatics*, vol. 33, no. 13, pp. 1944–1952, 2017.

[15] I. Sohn, K. Owzar, S. L. George, S. Kim, and S. H. Jung, "A permutation-based multiple testing method for time-course microarray experiments," *BMC Bioinformatics*, vol. 10, p. 336, 2009.

[16] S. Das, P. K. Meher, A. Rai, L. Mohan Bhar, and B. Nath Mandal, "Statistical approaches for gene selection, hub gene identification and module interaction in gene co-expression network analysis: an application to aluminum stress in soybean (Glycine max L.)," *PLoS One*, vol. 12, no. 1, Article ID e0169605, 2017.

[17] Y. Kim, B. Q. Doan, P. Duggal, and J. E. Bailey-Wilson, "Normalization of microarray expression data using within-pedigree pool and its effect on linkage analysis," *BMC Proceedings*, vol. 1, no. S1, p. S152, 2007.

[18] R. C. Gehrau, V. R. Mas, C. I. Dumur et al., "Donor hepatic steatosis induce exacerbated ischemia-reperfusion injury through activation of innate immune response molecular pathways," *Transplantation*, vol. 99, no. 12, pp. 2523–2533, 2015.

[19] J. Zhu and X. Yao, "Use of DNA methylation for cancer detection: promises and challenges," *International Journal of Biochemistry and Cell Biology*, vol. 41, no. 1, pp. 147–154, 2009.

[20] S. Hug, A. Raue, J. Hasenauer et al., "High-dimensional Bayesian parameter estimation: case study for a model of JAK2/STAT5 signaling," *Mathematical Biosciences*, vol. 246, no. 2, pp. 293–304, 2013.

[21] R. Lin, S. Dai, R. D. Irwin, A. N. Heinloth, G. A. Boorman, and L. Li, "Gene set enrichment analysis for non-monotone association and multiple experimental categories," *BMC Bioinformatics*, vol. 9, p. 481, 2008.

[22] A. P. Oron, Z. Jiang, and R. Gentleman, "Gene set enrichment analysis using linear models and diagnostics," *Bioinformatics*, vol. 24, no. 22, pp. 2586–2591, 2008.

[23] X. Chen, L. Wang, J. D. Smith, and B. Zhang, "Supervised principal component analysis for gene set enrichment of microarray data with continuous or survival outcomes," *Bioinformatics*, vol. 24, no. 21, pp. 2474–2481, 2008.

[24] J. Petereit, S. Smith, F. C. Harris, and K. A. Schlauch, "Petal: Co-expression network modelling in R," *BMC Systems Biology*, vol. 10, no. 2, p. 51, 2016.

[25] T. T. Cai and A. Zhang, "Inference for high-dimensional differential correlation matrices," *Journal of Multivariate Analysis*, vol. 143, pp. 107–126, 2016.

[26] N. C. Chung and J. D. Storey, "Statistical significance of variables driving systematic variation in high-dimensional data," *Bioinformatics*, vol. 31, no. 4, pp. 545–554, 2015.

[27] S. T. Low, M. S. Mohamad, S. Omatu, L. En Chai, S. Deris, and M. Yoshioka, "Inferring gene regulatory networks from perturbed gene expression data using a dynamic Bayesian

network with a Markov Chain Monte Carlo algorithm," in *Proceedings of the IEEE International Conference on Granular Computing*, Noboribetsu, Hokkaido, Japan, October 2014.

[28] W. C. Young, A. E. Raftery, and K. Y. Yeung, "Fast Bayesian inference for gene regulatory networks using ScanBMA," *BMC Systems Biology*, vol. 8, no. 1, p. 47, 2014.

[29] S. Basu, A. Shojaie, and G. Michailidis, "Network granger causality with inherent grouping structure," *Journal of Machine Learning Research*, vol. 16, pp. 417–453, 2012.

[30] Y. Samuels, Z. Wang, A. Bardelli et al., "High frequency of mutations of the PIK3CA gene in human cancers," *Science*, vol. 304, no. 5670, p. 554, 2004.

[31] N. A. Lockney, X. Pei, L. E. Blumberg et al., "PIK3CA activating mutations are associated with decreased local control in lung cancer brain metastases treated with radiation," *International Journal of Radiation Oncology Biology Physics*, vol. 96, no. 2, pp. S178–S179, 2016.

[32] C. D. Young, A. D. Pfefferle, P. Owens et al., "Conditional loss of ErbB3 delays mammary gland hyperplasia induced by mutant PIK3CA without affecting mammary tumor latency, gene expression, or signaling," *Cancer Research*, vol. 73, no. 13, pp. 4075–4085, 2013.

[33] Z. Xu, X. Huo, C. Tang et al., "Frequent mutations in MLH1, MET, KIT, PDGFRA, and PIK3CA genes in human gastrointestinal stromal tumors," *Pensar-Revista de Ciências Jurídicas*, vol. 11, no. 1, 2013.

[34] L. Xiang, W. Jiang, J. Li et al., "PIK3CA mutation analysis in Chinese patients with surgically resected cervical cancer," *Scientific Reports*, vol. 5, article 14035, 2015.

[35] I. A. Mayer and C. L. Arteaga, "The PI3K/AKT pathway as a target for cancer treatment," *Annual Review of Medicine*, vol. 67, no. 1, pp. 11–28, 2015.

[36] F. Atif, S. Yousuf, and D. G. Stein, "Anti-tumor effects of progesterone in human glioblastoma multiforme: role of PI3K/Akt/mTOR signaling," *Journal of Steroid Biochemistry and Molecular Biology*, vol. 146, pp. 62–73, 2015.

[37] R. D. Cario, E. Sticchi, S. Nistri, G. Pepe, and B. Giusti, "Role of TGFBR1 and TGFBR2 genetic variants in determining or modulating Marfan syndrome," *Nutrition, Metabolism and Cardiovascular Diseases*, vol. 27, no. 1, p. e17, 2017.

[38] D. R. Principe, B. DeCant, J. Staudacher et al., "Loss of TGFβ signaling promotes colon cancer progression and tumor-associated inflammation," *Oncotarget*, vol. 8, no. 3, p. 3826, 2017.

[39] S. Ihara, Y. Hirata, and K. Koike, "TGF-β in inflammatory bowel disease: a key regulator of immune cells, epithelium, and the intestinal microbiota," *Journal of Gastroenterology*, vol. 52, no. 7, pp. 1–11, 2017.