Iran Red Crescent Med J. 2017 March; 19(3):e41102. doi: 10.5812/ircmj.41102.

Published online 2016 November 21. Research Article

Determination of Biomarkers for Neonatal Sepsis Based on Differential Modules

Caixia Wang,[1,*] Shaoyong Luan,[1] Ming Li,[1] Ruiyun Zhang,[1] and Xiuxia Chen[1]

[1]Department of Pediatrics, Qingdao Municipal Hospital, Qingdao 266011, PR China

[*]*Corresponding author*: Caixia Wang, Department of Pediatrics, Qingdao Municipal Hospital, Qingdao 266011, PR China. Tel: +86-53282789387, Fax: +86-53282836421, E-mail: caixiawang623@yeah.net

Received 2016 July 27; **Revised** 2016 October 20; **Accepted** 2016 November 01.

Abstract

Background: The exact interacting factor that response to the infection for neonatal sepsis is still needed to urgently to be disclosed.
Objectives: This research was aimed to explore the potential biomarkers and illuminate the underlying molecular mechanisms associated with neonatal sepsis via identifying differential modules (DMs).
Methods: This is a case-control bioinformatics analysis using already published microarray data of neonatal sepsis. This study was conducted in Qingdao, China from September 2015 to May 2016. We recruited the gene expression profile of neonatal sepsis from the Array Express database (http://www.ebi.ac.uk/arrayexpress) under the accessing number of E-GEOD-25504, which included 27 neonatal samples with a confirmed blood culture-positive test for sepsis (bacterial infected cases) as well as 35 matched controls. Meanwhile, the human protein-protein interaction (PPI) data was collected from the database of Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, http://string-db.org). All of the data was preprocessed. Then, the differential co-expression network (DCN) was constructed by integrating co-expression analysis and differential expression analysis. Next, a systemic module searching strategy, which contained seed genes selection, module searching and refinement of modules, was performed by select DMs.
Results: Starting from the gene expression data and PPI data, the DCN that included 430 edges (covering 324 nodes) was constructed, in which each edge was assigned a weight value. From the DCN, we selected a total of 16 seed genes. Starting from these seed genes, a total of 3 modules were identified from the DCN based on the systemic module algorithm. Of them, only one module (Module 3) was considered as DM under $P < 0.05$. This DM was involved in the progress of ribosome biogenesis in eukaryotes.
Conclusions: In the present study, we identified a key gene RPS16 and a significant module involved in ribosome biogenesis in eukaryotes that were related to neonatal sepsis, which might be potential biomarkers for early detection and therapy for neonatal sepsis.

Keywords: Neonate, Sepsis, Gene Network, Biomarkers

1. Background

Neonatal sepsis, a systemic infection occurring in infants at ≤ 28 days of life (1) is a major cause of morbidity and mortality in newborns, especially in the developing countries (2). Even though there has been an improvement in neonatal care in the past few years, infection remains a leading cause of morbidity and mortality in neonates worldwide by currently causing about 1.6 million deaths per year (3). Moreover, over the past few years, to elucidate the mechanisms of neonatal sepsis, various kinds of efforts have concentrated on the host genetic variability (4-7). Some gene makers, such as CD64 (8), IL-6, CR, and PCT (9) have been indicated to be associated with neonatal sepsis. Moreover, the whole gene expression profiling of neonatal sepsis have been performed (10). Smith et al. (11) identified a 52-gene-classifier that could predict bacterial infection accurately by integrating co-expressed gene modules and immune and metabolic pathways. However, the exact interacting factor that response to the infection of neonatal sepsis is still unclear and is needed to urgently be disclosed.

Network analysis has opened new insights into the pathogenesis and progression of the diseases (12-14). It is believed that the dynamics of the molecular networks during the progression of the disease can contribute to tracking the biomarkers for this disease (15). Moreover, it has been reported that functional gene modules could help us understand the mechanism of diseases and provide opportunities to develop new therapies (16, 17). Currently, Ma et al. (18) proposed a differential module (DM) algorithm to identify differentially expressed gene modules with common members yet varied connectivity across multiple differential co-expression networks (DCNs).

Copyright © 2016, Iranian Red Crescent Medical Journal. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/) which permits copy and redistribute the material just in noncommercial usages, provided the original work is properly cited.

2. Objectives

In the present study, we aimed to explore the potential biomarkers for neonatal sepsis based on the module-search algorithm. To achieve this, the gene expression profile of neonatal sepsis and human protein-protein interaction (PPI) data was collected and preprocessed. Then, the DCN was constructed by integrating gene expression and PPI data. Next, a module searching algorithm, containing seed gene selection, module searching by seed gene expansion and module refinement was performed. Statistical significance analysis was conducted to select the DMs. This study integrated the gene expression data of neonatal sepsis with PPI data and co-expression information to track the DM using a novel module-search strategy. The results might contribute to understanding the progression and treatment of neonatal sepsis.

3. Methods

3.1. Acquisition of Expression Profile Data

This work is a case-control bioinformatics analysis using already published microarray data of neonatal sepsis. The gene expression dataset of neonatal sepsis under the accessing number of E-GEOD-25504 was recruited from the Array Express database (http://www.ebi.ac.uk/arrayexpress). In E-GEOD-25504, there were a total of 170 samples that existed on four platforms. To eliminate the batch effect, only samples existed on the platform of Illumina HumanHT-12 v3.0 Expression Bead Chip in training set were retained in our study. Meanwhile, we removed the samples with viral infections. Finally, a total of 62 samples were retained, including 27 neonatal samples with a confirmed blood culture-positive test for sepsis (bacterial infected cases) and 35 matched controls. The details of demographical variables and a confounding factor have been presented in the previous studies (10, 11). Moreover, in the previous studies, the researchers have performed a power calculation using the Illumina chip platform, on an independent set of 30 neonatal samples at 9 months of age and showed that the study design has 90% power to detect a twofold change in expression with an α of 1% (false discovery rate (FDR) corrected) (10, 11). The gene expression data was preprocessed and the probes were mapped to the corresponding official gene symbol. Finally, we obtained 15,449 genes and their expression data from the microarray dataset.

3.2. Construction of PPI Network

The human PPI data was downloaded from the database of search tool for the Retrieval of interacting genes/proteins (STRING, http://string-db.org), which provided a comprehensive, yet quality-controlled collection of protein-protein associations for a large number of organisms (19). A total of 787,896 interactions (covering 16,730 nodes) were downloaded from the STRING database. By intersecting the genes obtained from microarray profile to the PPI data, we constructed a background PPI network with 453,107 interactions.

3.3. Construction of the DCN

For each interaction of two genes under the disease condition, we calculated the absolute value of Pearson correlation coefficient (PCC), respectively, to assess the co-expression situation. In an attempt to eliminate indirect correlation because of a third gene, the utilization of the first order partial PCC was implemented (20). Then, to select the co-expressed genes, the interactions whose absolute value of PCC value was greater than the pre-defined threshold δ (δ = 0.9) were selected to construct the co-expression network.

Meanwhile, the one-side t-test was utilized to determine the P value of the differential gene expression between disease and normal conditions. Then, the EdgeR (21), a Bio conductor package for differential expression analysis of digital gene expression data, was utilized to detect differential gene expression for the microarray data. The weight wm, n on edge (m, n) in the co-expression network was calculated as following Formula 1:

$$W_{m,n} = \left\{ \begin{array}{c} \frac{(logp_m+logp_n)^{\frac{1}{2}}}{\left(2^{*^{max}}0,_{0,}^{1}\in V \quad logP_1\right)^{\frac{1}{2}}}, if \quad cor\left(m,n\right) \geq \delta, \\ if \quad cor\left(m,n\right) < \delta \end{array} \right\} \quad (1)$$

Where m, n represented the two genes on each edge. p_m and p_n were respective P values of differential expression for genes m and n. V was on behalf of the gene set of the co-expression network and cor (m, n) represented the absolute value of PCC between gene m and n.

In this case, a DCN was built and each edge was assigned a weight value. Under this weighting scheme, genes that were co-expressed and significantly differentially expressed were assigned high weights, which might indicated that these genes exhibited differential activities between disease and normal conditions. Therefore, these genes might play important roles during the occurrence and development of neonatal sepsis.

3.4. Identification of Modules in DCN

The module algorithm has been introduced to explore gene modules with common members but varied connectivity across multiple molecular interaction networks (22).

In this article, we performed the following three steps to identify the candidate modules from DCN.

3.4.1. Seed Prioritization

In this part, we firstly ranked genes in DCN using the topological feature of the genes contained in the network. Specifically, we computed the importance of each gene m in the DCN. The importance of a gene in the network depends on the number of its neighbors, strength of connection and importance of its neighbors. In this case, each gene was given a z-sore in the current study. All genes contained in the DCN were ranked in descending order based on the z-sores and the top 5% genes were considered as seed genes.

3.4.2. Candidate Modules

In this step, the modules were searched via the expansion and entropy minimization of seed genes. Starting from each seed gene, the module search step iteratively included genes whose addition led to the maximum decrease in the graph entropy-based objective function and the search stopped until there was no decrease in the objective function. For the known seed gene v ϵ V, it was thought as a module C. Then, we joined the neighbors v' ϵ V of gene v into this module to form a new module C'. We calculated the entropy decrease ΔH as the connectivity between C and C'. The function was defined: ΔH (C′, C) = H (C) - H (C′). If ΔH (C′, C) $>$ 0, we thought the neighbor gene v′ increased the connectivity of module C. Next, we joined all the others neighbor gene v' which can result in ΔH (C', C) $>$ 0 into the module C until there were no additional ones. In this condition, it meant that each gene could belong to one or more modules and each module contained at least one seed gene.

In addition, we performed the refinement of modules obtained above. The modules with node sizes $<$ 5 were removed. What was more, the modules were merged if the overlapping degree of nodes was between two modules $\geq$ 0.5.

3.5. Significant Analysis

The statistically significance of candidate modules was calculated on the basis of the null distribution of modules generated through randomized networks. Firstly, the network with the same number of edges in DCN was randomly captured from the background PPI network as randomized networks by degree-preserved edge shuffling. This step was performed 100 times randomly. Then, the module searching was conducted from the randomized networks. In addition, the empirical P value of each module was evaluated as the probability of the module with the smaller score by chance. The Formula 2 was shown as followed:

$$P \quad Value = Sum\frac{(HR > HD)}{HR} \tag{2}$$

Where HR stood for the number of modules from randomized networks and HD was the number of modules from DCN. In addition, the Benjamini and Hochberg (23) method was introduced to adjust the P value. Finally, the modules with adjusted P $\leq$ 0.05 were considered as the DMs.

3.6. Functional Analysis

To study functional inference of DMs, pathway enrichment analysis was conducted in Genelibs (http://www.genelibs.com/gb) based on the Kyoto encyclopedia of genes and genomes (KEGG) pathway database. The Fisher's exact test was utilized to determine the P values and the Benjamini-Hochberg method was performed to conduct multiple testing on the P values. The pathways whose adjusted P$<$ 0.05 were considered as the significant pathways.

4. Results

In our study, in order to explore the dynamic molecular changes in neonatal sepsis, we utilized the module-search algorithm to identify sub-networks (modules) in DCN. By topological analysis of DCN, the seed genes were identified and the modules were identified on the basis of the calculation of entropy change. Finally, we determined the DMs and their functional inference by statistical analysis and pathway enrichment analysis, respectively.

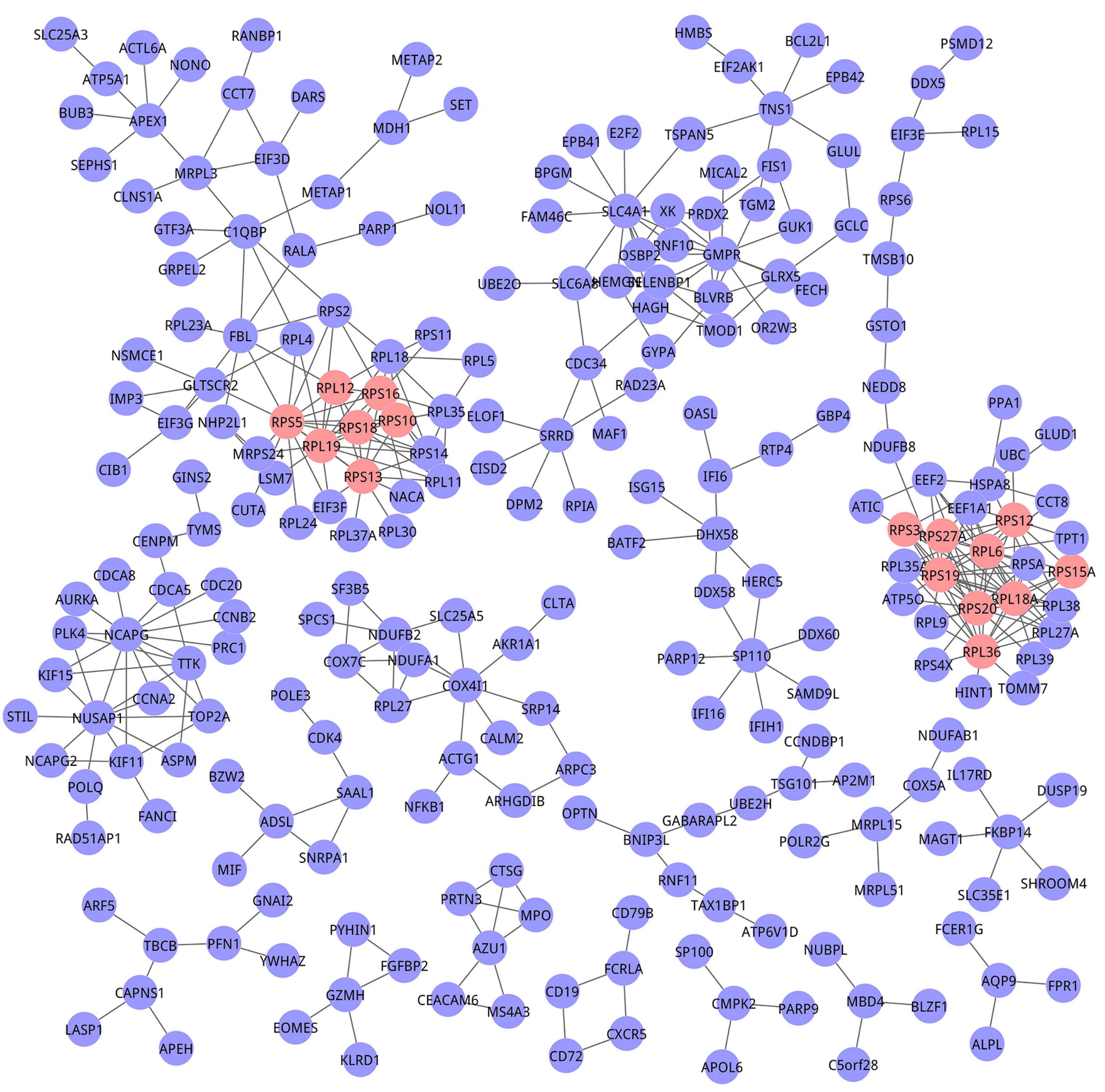
4.1. Construction of DCN

Prior to the analysis, gene expression data of neonatal sepsis and human PPIs were obtained from the open databases and then preprocessed, respectively. Then we constructed the background PPIs by integrating transcriptome data and PPI data, and calculated the absolute value of the PCC of all gene intersections. Under the pre-defined threshold of $\delta >$ 0.9, a total of 430 interactions (covering 324 genes) were obtained. Then, the EdgeR was applied to assign a weight value to each interaction. As shown in Figure 1, the DCN with 430 interactions (covering 324 genes) was constructed and each interaction was assigned a weight value.

4.2. Identification of Candidate Modules From DCN

In this part, the genes contained in the DCN were ranked in descending order according to their z-scores and the top 5% genes according to the z-score distribution were considered as seed genes. We identified a total of 16 seed genes from the DCN. The details of the seed genes were

Figure 1. The Differential Co-Expression Network (DCN), Which Included 430 Interactions (324 genes) in Neonatal Sepsis



Orange nodes were behalf of the seed genes.

shown in Table 1. Amongst these 16 seed genes, there were 6 genes with z-score > 60, which were RPS20 (z-score = 78.70), RPL6 (z-score = 75.60), RPL18A (z-score = 74.92), RPS19 (z-score = 72.65), RPS16 (z-score = 61.93) and RPS5 (z-score = 61.42). Starting from these 16 seed genes, a total of 16 modules were identified. After module refining, a total of 3 candidate modules (Module 1, Module 2 and Module 3) were identified.

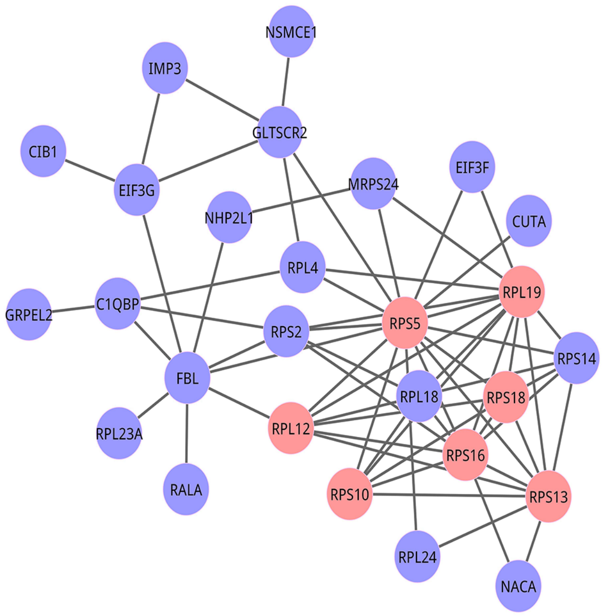
4.3. Selection of the DM

Based on the statistical analysis, we obtained the p-values of these 3 candidate modules. The result showed that only Module 3 (P = 0.04) was a DM in this work, the result was shown in Figure 2. We found that there were 64 edges (covering 27 nodes) in Module 3. Meanwhile, we noticed that 7 seed genes existed in Module 3, all which were ribosomes-related genes. *RPS16* (z-score = 61.93) showed the highest z-score in this module.

Table 1. The 16 Seed Genes With Their z-Scores in Descending Order

Seed Genes	z-Score	Seed Genes	z-Score
RPS20	78.70	RPS27A	58.05
RPL6	75.60	RPS15A	55.80
RPL18A	74.92	RPL12	55.40
RPS19	72.65	RPS13	53.16
RPS16	61.93	RPS10	51.74
RPS5	61.42	RPS18	51.30
RPL19	59.94	RPS3	50.12
RPL36	58.44	RPS12	49.66

Figure 2. The Differential Module, Which Included of 64 Interactions (27 Nodes)



Orange nodes represented the seed genes.

4.4. Functional analysis

Genes usually take part in the biological process in a functional cooperation way in a certain disease. In the present study, to disclose the functional inference of the DM, pathway enrichment analysis was conducted. Under the threshold value of P <0.05, ribosome biogenesis in eukaryotes (P = 0.0088) was identified. We inferred that the DM mainly affected the pathway ribosome biogenesis in eukaryotes to function during the occurrence and development of neonatal sepsis.

5. Discussion

Septicemia in neonates refers to generalized bacterial infection documented by a positive blood culture in the first 4 weeks of life (24). The infection in neonates is a global problem with significant morbidity and mortality (25). Preliminary studies indicate that the diagnosis of neonatal sepsis is complicated via nonspecific clinical symptomatology, a high false negative rate and a delay in obtaining blood culture results. Hence, searching an ideal biomarker for neonatal sepsis was with great help for recognizing the definite infection mechanisms and giving a guide for the principles of therapy at an early stage.

Recently, it has been indicated that the analysis of functional modules instead of individual genes would be more effective for system-wide identification of cellular functions (26). Moreover, the concept of network biology has been widely applied to a variety of disease. Xing et al. (27) indicated that specific gene modules derived from gene co-expression networks and may provide better understanding of molecular mechanisms in disease. In co-expression networks, two genes are connected and assumed to functionally interact if their expression profiles are correlated across multiple conditions. While networks contained only co-expression information it might reduce the statistical power for identifying genes that are perturbed under disease conditions.

By network topological measurements, we identified 16 seed genes from the DCN. It could easily be found that all seed genes were Ribosomes-related genes. Ribosome, the organelle that catalyzes protein synthesis, consists of a small 40S subunit and a large 60S subunit. Amongst the 16 seed genes, 11 seed genes belong to 40S subunit and 5 belong to 60S ribosomal subunit. Starting from these seed genes, we carried out a modules-search algorithm to identify the gene modules from DCN. Only one DM (Module 3) was identified, which was involved in the pathway of ribosome biogenesis in eukaryotes.

In Module 3, RPS16 was the initial seed gene. RPS16 (ribosomal protein S16) encode the ribosomal protein that is a component of the 40S subunit, which belongs to the S9P family of ribosomal proteins. Ribosomes, which synthesize the proteome of cells, are complex ribonucleoproteins in eukaryotes. Ribosomal proteins, in conjunction with rRNA, make up the ribosomal subunits involved in the cellular process of translation (28, 29). Ghosh et al. (30) indicated that RPS5-RPS16 communication is essential for efficient translation initiation in yeast S. cerevisiae. Karan et al. (31) reported that RPS16 was correlated with the progression of human prostate cancer. Yang et al. (32) revealed that protein synthesis processes of RPS16 were related to the progression of disc degeneration. However, there was only a few relevant researches for this gene in neonatal sepsis. In the present study, RPS16 and ribosome biogenesis in eukaryotes pathway were indicated to be the important gene and functional pathway in neonatal sepsis, respectively. Further experimental analysis should be implemented to reveal the underlying relationship between them and neonatal sepsis.

There are several limitations in the present study. The samples were retrieved from the open access database but not obtained from our hospital. We did not perform the microarray analysis of samples from patients with neonatal sepsis. This work is a pure bioinformatics analysis; the results were not verified by wet experiments. In the present study, we identified a key gene RPS16 and a significant module involved in ribosome biogenesis in eukaryotes that were related to neonatal sepsis, which might be potential biomarkers for early detection and therapy for neonatal sepsis. Further experiments for verification should be performed in the future.

Acknowledgments

This research was supported by the department of pediatrics, Qingdao Municipal hospital. We thank all the members of the research group. Meanwhile, we are grateful to Beijing Spirallink medical research institute for the professional translation and paper polishing.

References

1. Samsygina GA, Shabalov NP, Talalaev AG, Milovanov AP, Glukhovets NG, Glukhovets BI. [Sepsis in the newborn]. *Arkh Patol.* 2004;**Suppl**:1–48. [PubMed: 15285077].
2. Verma P, Berwal PK, Nagaraj N, Swami S, Jivaji P, Narayan S. Neonatal sepsis: epidemiology, clinical spectrum, recent antimicrobial agents and their antibiotic susceptibility pattern. *Int J Contemporary Pediatr.* 2015;**2**(3):176–80. doi: 10.18203/2349-3291.ijcp20150523.
3. Vergnano S, Sharland M, Kazembe P, Mwansambo C, Heath PT. Neonatal sepsis: an international perspective. *Arch Dis Child Fetal Neonatal Ed.* 2005;**90**(3):220–4. doi: 10.1136/adc.2002.022863. [PubMed: 15846011].
4. Jabandziev P, Smerek M, Michalek J, Fedora M, Kosinova L, Hubacek JA, et al. Multiple gene-to-gene interactions in children with sepsis: a combination of five gene variants predicts outcome of life-threatening sepsis. *Crit Care.* 2014;**18**(1):1. doi: 10.1186/cc13174. [PubMed: 24383711].
5. Wong HR. Genetics and genomics in pediatric septic shock. *Crit Care Med.* 2012;**40**(5):1618–26. doi: 10.1097/CCM.0b013e318246b546. [PubMed: 22511139].
6. Abu-Maziad A, Schaa K, Bell EF, Dagle JM, Cooper M, Marazita ML, et al. Role of polymorphic variants as genetic modulators of infection in neonatal sepsis. *Pediatr Res.* 2010;**68**(4):323–9. doi: 10.1203/00006450-201011001-00632. [PubMed: 20463618].
7. Chauhan M, McGuire W. Interleukin-6 (-174C) polymorphism and the risk of sepsis in very low birth weight infants: meta-analysis. *Arch Dis Child Fetal Neonatal Ed.* 2008;**93**(6):427–9. doi: 10.1136/adc.2007.134205. [PubMed: 18375611].
8. Streimish I, Bizzarro M, Northrup V, Wang C, Renna S, Koval N, et al. Neutrophil CD64 with hematologic criteria for diagnosis of neonatal sepsis. *Am J Perinatol.* 2014;**31**(1):21–30. doi: 10.1055/s-0033-1334453. [PubMed: 23456906].
9. Abdollahi A, Shoar S, Nayyeri F, Shariat M. Diagnostic Value of Simultaneous Measurement of Procalcitonin, Interleukin-6 and hs-CRP in Prediction of Early-Onset Neonatal Sepsis. *Mediterr J Hematol Infect Dis.* 2012;**4**(1):2012028. doi: 10.4084/MJHID.2012.028. [PubMed: 22708043].
10. Dickinson P, Smith CL, Forster T, Craigon M, Ross AJ, Khondoker MR, et al. Whole blood gene expression profiling of neonates with confirmed bacterial sepsis. *Genom Data.* 2015;**3**:41–8. doi: 10.1016/j.gdata.2014.11.003. [PubMed: 26484146].
11. Smith CL, Dickinson P, Forster T, Craigon M, Ross A, Khondoker MR, et al. Identification of a human neonatal immune-metabolic network associated with bacterial infection. *Nat Commun.* 2014;**5**:4649. doi: 10.1038/ncomms5649. [PubMed: 25120092].
12. Saris CG, Horvath S, van Vught PW, van Es MA, Blauw HM, Fuller TF, et al. Weighted gene co-expression network analysis of the peripheral blood from Amyotrophic Lateral Sclerosis patients. *BMC Genomics.* 2009;**10**:405. doi: 10.1186/1471-2164-10-405. [PubMed: 19712483].
13. Alavi Majd H, Talebi A, Gilany K, Khayyer N. Two-Way Gene Interaction From Microarray Data Based on Correlation Methods. *Iran Red Crescent Med J.* 2016;**18**(6):24373. doi: 10.5812/ircmj.24373. [PubMed: 27621916].
14. Alavian SM. Networking for overcoming on viral hepatitis in middle east and central asia: "Asian Hepatitis Network". *Hepat Mont.* 2007;**7**(4):181–2.
15. Chen L, Liu R, Liu ZP, Li M, Aihara K. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. *Sci Rep.* 2012;**2**:342. doi: 10.1038/srep00342. [PubMed: 22461973].
16. Zhu L, Zhu F. Identification association of drug-disease by using functional gene module for breast cancer. *BMC Med Genomics.* 2015;**8 Suppl 2**:3. doi: 10.1186/1755-8794-8-S2-S3. [PubMed: 26045063].
17. Ghiassian SD, Menche J, Barabasi AL. A DIseAse MOdule Detection (DIAMOnD) algorithm derived from a systematic analysis of connectivity patterns of disease proteins in the human interactome. *PLoS Comput Biol.* 2015;**11**(4):1004120. doi: 10.1371/journal.pcbi.1004120. [PubMed: 25853560].
18. Ma X, Gao L, Karamanlidis G, Gao P, Lee CF, Garcia-Menendez L, et al. Revealing Pathway Dynamics in Heart Diseases by Analyzing Multiple Differential Networks. *PLoS Comput Biol.* 2015;**11**(6):1004332. doi: 10.1371/journal.pcbi.1004332. [PubMed: 26083688].
19. von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, et al. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res.* 2005;**33**(Database issue):433–7. doi: 10.1093/nar/gki005. [PubMed: 15608232].

20. Watson-Haigh NS, Kadarmideen HN, Reverter A. PCIT: an R package for weighted gene co-expression networks based on partial correlation and information theory approaches. *Bioinformatics.* 2010;**26**(3):411–3. doi: 10.1093/bioinformatics/btp674. [PubMed: 20007253].
21. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;**26**(1):139–40. doi: 10.1093/bioinformatics/btp616. [PubMed: 19910308].
22. Ma X, Gao L, Tan K. Modeling disease progression using dynamics of pathway connectivity. *Bioinformatics.* 2014;**30**(16):2343–50. doi: 10.1093/bioinformatics/btu298. [PubMed: 24771518].
23. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Statistical Society Series B.* 1995:289–300.
24. Singh M, Deorari AK, Khajuria RC, Paul VK. Perinatal & neonatal mortality in a hospital. *Indian J Med Res.* 1991;**94**:1–5. [PubMed: 2071176].
25. Amel Jamehdar S, Mammouri G, Sharifi Hoseini MR, Nomani H, Afzalaghaee M, Boskabadi H, et al. Herpes simplex virus infection in neonates and young infants with sepsis. *Iran Red Crescent Med J.* 2014;**16**(2):14310. doi: 10.5812/ircmj.14310. [PubMed: 24719742].
26. Oh K, Hwang T, Cha K, Yi GS. Disease association and interconnectivity analysis of human brain specific co-expressed functional modules. *Biol Res.* 2015;**48**(1):1. doi: 10.1186/s40659-015-0061-4.
27. Xing YH, Zhang JL, Lu L, Li DG, Wang YY, Huang S, et al. Identification of specific gene modules in mouse lung tissue exposed to cigarette smoke. *Asian Pac J Cancer Prev.* 2015;**16**(10):4251–6. doi: 10.7314/APJCP.2015.16.10.4251. [PubMed: 26028081].
28. Ban N, Beckmann R, Cate JH, Dinman JD, Dragon F, Ellis SR, et al. A new system for naming ribosomal proteins. *Curr Opin Struct Biol.* 2014;**24**:165–9. doi: 10.1016/j.sbi.2014.01.002. [PubMed: 24524803].
29. Rodnina MV, Wintermeyer W. The ribosome as a molecular machine: the mechanism of tRNA-mRNA movement in translocation. *Biochem Soc Trans.* 2011;**39**(2):658–62. doi: 10.1042/BST0390658. [PubMed: 21428957].
30. Ghosh A, Jindal S, Bentley AA, Hinnebusch AG, Komar AA. Rps5-Rps16 communication is essential for efficient translation initiation in yeast S. cerevisiae. *Nucleic Acids Res.* 2014;**42**(13):8537–55. doi: 10.1093/nar/gku550. [PubMed: 24948608].
31. Karan D, Kelly DL, Rizzino A, Lin MF, Batra SK. Expression profile of differentially-regulated genes during progression of androgen-independent growth in human prostate cancer cells. *Carcinogenesis.* 2002;**23**(6):967–75. doi: 10.1093/carcin/23.6.967. [PubMed: 12082018].
32. Yang Z, Chen X, Zhang Q, Cai B, Chen K, Chen Z, et al. Dysregulated COL3A1 and RPL8, RPS16, and RPS23 in Disc Degeneration Revealed by Bioinformatics Methods. *Spine (Phila Pa 1976).* 2015;**40**(13):745–51. doi: 10.1097/BRS.0000000000000939. [PubMed: 25893343].