Biointerface Research in Applied Chemistry
Platinum Open Access Journal (ISSN: 2069-5837)

Article
Volume 11, Issue 5, 2021, 12901 - 12915
https://doi.org/10.33263/BRIAC115.1290112915

Identification of Specific Gene Modules and Candidate Signatures in Necrotizing Enterocolitis Disease: Network-based Gene Co-expression Approach

Md. Rabiul Auwul [1,*], Chongqi Zhang [1,*], Md Rezanur Rahman [2], Md. Shahjaman [3]

[1] School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China; rabiulauwul@gmail.com (M.R.A.); cqzhang@gzhu.edu.cn (C.Z.);
[2] Department of Biochemistry and Biotechnology, School of Biomedical Science, Khwaja Yunus Ali University, Sirajgonj-6751, Bangladesh; rezanur12@yahoo.com (M.R.R.);
[3] Department of Statistics, Begum Rokeya University, Rangpur-5400, Bangladesh; shahjaman_brur@yahoo.com (M.S.);
* Correspondence: rabiulauwul@gmail.com (M.R.A), cqzhang@gzhu.edu.cn (C.Z.);

Scopus Author ID 35212374200
Received: 28.12.2020; Revised: 18.01.2021; Accepted: 24.01.2021; Published: 31.01.2021

Abstract: Necrotizing enterocolitis (NEC) is a serious disease of the gastrointestinal systems that primarily affects premature newborns' intestine in neonatal intensive care units. The present study aimed to detect NEC molecular signatures and pathways from comprehensive bioinformatics analysis of NEC's RNA-seq transcriptomics. We performed systems biology analysis of RNA-seq transcriptomics data (with accession GSE64801) of NEC from nine NEC and five healthy controls. Differential expression of gene expression was performed using a combination of three R packages "DESeq2", "edgeR", “edger robust”. Gene co-expression analysis was performed using a weighted WGCNA package to identify gene modules, Gene Ontology (GO), pathway analysis, protein-protein interaction, gene-transcription factor, and gene-microRNA interaction analysis was performed. The differential expression analysis identified 966 differentially expressed genes (DEGs) in NEC from the RNA-seq dataset related to corresponding controls. The WGCNA showed the presence of three key gene modules. The GO analysis showed genes are enriched in metabolic processes, regulation of immune response and immune systems, cell communication, and cellular process. The immune and complement pathways are related to co-expressed key modules that were detected. The protein-protein interactions analysis showed the presence of key hub genes related to the modules. Integration of these co-expressed gene modules with regulatory networks showed the presence of significant key transcription factors and microRNAs as hub molecules. The present study's findings suggested the immune systems and complement cascade are key mechanisms of NEC pathogenesis. The comprehensive network analysis showed several key hub molecules that might be potential biomarkers and drug targets in NEC.

Keywords: necrotizing enterocolitis; RNA-seq; weighted gene co-expression network; systems biology; protein-protein interaction.

1. Introduction

Necrotizing enterocolitis (NEC) is one of the prevalent life-threatening diseases of the gastrointestinal systems that primarily affects premature newborns' intestines in neonatal intensive care units [1,2]. NEC is characterized by inflammation and necrosis of the intestine and high morbidity [3,4]. Several complications such as the development of short bowel syndrome, cholestasis liver disease, and perturbed growth and neurodevelopmental outcomes

are evident in NEC survivors [5,6]. The major risk factor recommended for the NEC's development is the pre-term birth, aberrant bacterial colonization and eternal feeding [7,8]. However, prematurity has been considered the established risk factor in NEC's development so far [9,10], but the precise mechanism has not been uncovered yet. Neonatologists face the most serious difficulty determining accurate early clinical signs and symptoms of NEC [9,11]. While there are numerous NEC diseases with different symptoms, the most common type of the disease is the inflammatory gut syndrome of prematurely born babies, referred to as "classic NEC" [1,3,9]. That being said, NEC is somewhat unspecific in its early clinical indications and may often be mischaracterized as other gastrointestinal issues [1,9,12]. Due to its abrupt onset, NEC has often been detected at the advanced stage [13,14]. An early diagnosing method for recognizing pre-term babies at risk of developing NEC or the outbreak of symptoms to help diagnose the problem would be a technique to avoid or treat NEC.

Several efforts were made to discover key genes with NEC [15–22] or distinguish them from related pathologies [23] though the suitable biomarker still needs to be established [24]. The gene expression profiling was done in NEC compared to controls via RNA-sequencing (RNA-Seq) and microarray gene expression profiling. Several deregulated genes and pathways contributing to NEC was identified [25–29]. Despite significant findings from these studies, there is an urgent need to uncover transcriptional landscape, associated biomolecules, and molecular pathways in NEC. Adopting gene co-expression profiling coupled with systems biology methodologies would identify significant co-expressed gene modules and pathways to shed light on NEC's mechanism. Thus, the weighted gene co-expression of RNA-seq transcriptomes in NEC would provide novel insights into NEC's molecular mechanism [30,31]. WGCNA is extensively used for candidate signatures identification in various diseases [32–40], having a great prospect in NEC disease.

In the present study, we used publicly available RNA-seq datasets (with accession no GSE64801) to identify differentially expressed genes (DEGs). The co-expression network was then constructed via the WGCNA algorithm and identified the NEC's highly correlated gene modules. The Gene Ontology (GO), pathway enrichment analysis, protein-protein interaction (PPI) network was conducted for these gene modules. The transcriptomics factors (TFs) and microRNAs (miRNAs) networks of co-expressed genes in each key module were conducted in this study. Our results provide novel insights into NEC's pathogenesis and the potential molecular targets for novel interventional approaches.

2. Materials and Methods

The workflow of this study is illustrated in Figure 1.

2.1. Data collection.

The RNA-Seq gene expression data with accession GSE64801, which was deposited by Trembly *et al.* from NEC compared to control based on the platform GPL11154 [Illumina HiSeq 2000] were collected from Gene Expression Omnibus (GEO) database [25]. A total of 14 samples, with 5 healthy controls and 9 NEC diagnosed pre-term infant samples included in this dataset.

2.2. Transcriptomic data processing and differential expression analysis.

The raw count data were normalized, and differential expression was performed to detect the differentially expressed genes (DEGs) NEC samples compared to healthy controls using edgeR [41], DESeq2 [42], and edgeR robust [43] algorithms through R packages. We take the common DEGs of these three algorithms for further analysis. For all the algorithms, the p-value < 0.05 and absolute [log2 (FC)]>=1 were used as the cutoff criteria for DEGs identification.

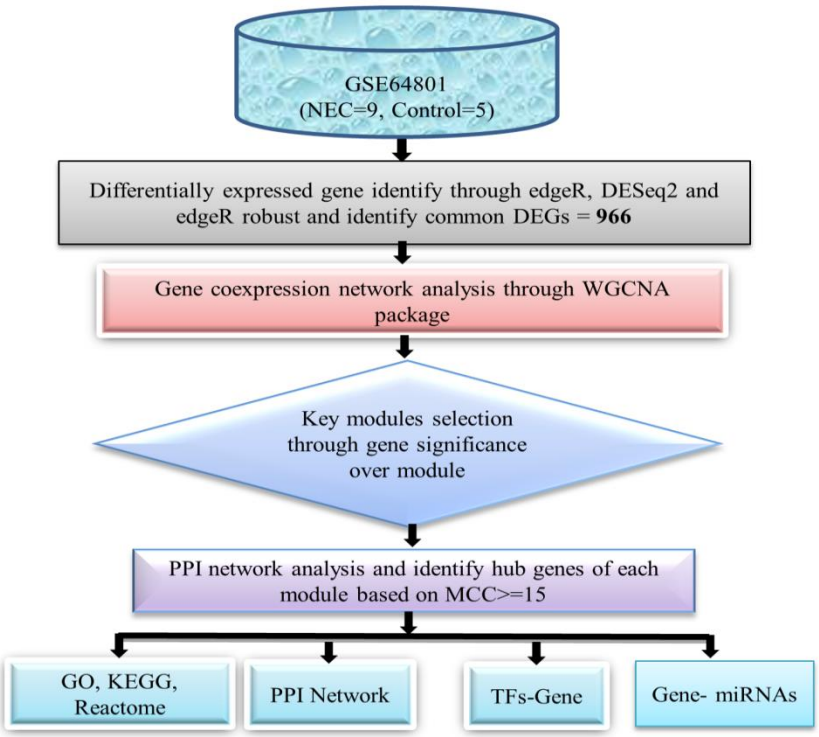

Figure 1. Flowchart of the methodology of this paper.

2.3. Weighted gene co-expression networks construction and key module selection.

For the gene co-expression network construction, firstly, we remove the outlier samples (if there exist) by constructing the sample cluster dendrogram by hclust R function. We then used the WGCNA package [44] of R in this reduced dataset for constructing a gene co-expression network. We used the pickSoftThreshold function for finding numerous soft thresholding powers β over R2 and pick the value of β for which the R2 value is being higher. Then, we constructed the adjacency matrix and Topological Overlap Matrix (TOM) by using the transformed gene expression matrix. The dissimilarity of TOM (dissTOM) was also conducted to construct a network heatmap plot and further analysis.

For the module selection, the dendrogram of genes constructed with dissTOM matrix using hclust R function with different colors. The Dynamic Tree Cut technique is used to getting modules for branch cutting. The parameters deepSplit=2 and minClusterSize=30 are being used for avoiding the generation of small or large modules. We used MEDissThres = 0.45 for merging similar modules.

For the selection of key modules, we measured the module significance of all genes in the module. The absolute value for ascertaining a correlation-based gene significance measure was used for gene significance over each module. We considered the first three modules with the highest gene significance as the key modules.

2.4. Gene Ontology and pathway analysis.

There is high connectivity of genes inside the co-expression modules. The genes perform similar roles within the same module. The functional enrichment analysis for the genes was studied in each selected key modules by GO and pathway analysis, for the GO analysis was identified and visualized by using clusterProfiler R package [45]. The pathway enrichment analysis was done via Enrichr [46] and identified significant pathways involved in gene co-expressed modules. We considered KEGG and Reactome pathway as data annotation sources. A statistical threshold criterion with an adjusted p-value <0.05 was used to select for the GO and pathway analysis.

2.5. PPI network analysis to detect hub genes in the key modules.

The DEG from each key module was used to construct the Protein-Protein Interaction (PPI) networks with the widespread web-based tool STRING [47], and the networks were visualized with Cytoscape open-source platform [48]. Then, we identified the top-ranked genes known as hub genes. The Maximal Clique Centrality (MCC) measurement [49] was used for the detection of hub genes through the CytoHubba plugin [49] in Cytoscape. The higher the value of MCC of the nodes, the higher the number of edges connected in those hub proteins.

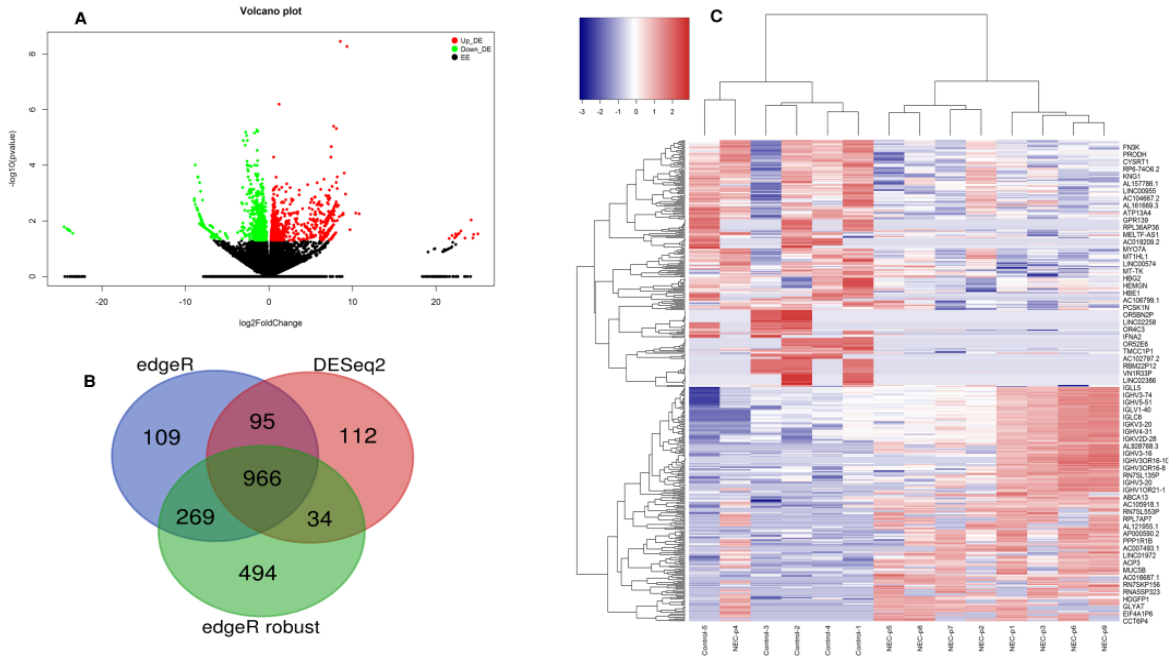

Figure 2. Differential expression profiles of the GSE64801 dataset. (**A**) The volcano plot of the GSE64801 dataset; (**B**) Venn diagram of DEGs identified using edgeR, DESeq2 and robust edgeR approach; (**C**) Heatmap of the identified differentially expressed genes (DEGs), which were mutually identified by three methods termed "common DEGs".

2.6. Identification of transcription factors and miRNAs that may regulate co-expressed gene modules.

We sought to detect the transcription factors (TFs) and microRNAs (miRNAs) that may control the gene modules' genes. We used a freely accessible database, JASPAR [50], for executing TFs-DEGs interaction through NetworkAnalyst [51]. For the identification of the significant miRNAs, we used the freely accessible database for miRNAs-target interactions databases, namely, Tarbase [52] and mirTarbase [53], to build miRNAs-DEGs interaction network via NetworkAnalyst [51]. Then, we used Cytoscape for visualization and identification

of significant hub TFs and miRNAs identification. We used the CytoHubba plugin in Cytoscape based on the MCC cutoff >=15 to select the hubs.

3. Results and Discussion

3.1.Results.

3.1.1. Differential expression analysis.

The transcriptomics dataset GSE64801 contains gene expression of 58037 genes from nine NEC and five healthy control samples. We performed differential expression of gene expression of RNA-Seq data with three R packages, namely DESeq2 (robust version), edgeR, and edgeR (robust version), to identify differentially expressed genes (DEGs), each separately (Figure 2). Then, we considered common DEGs identified by these packages; these common DEGs are considered robust. Figure2A shows the expression with direction genes, including the up-regulated, down-regulated, and equally expressed genes. The statistical analysis revealed 1439, 1207 and 1763 DEGs through edgeR, DESeq2 and edgeR robust algorithms respectively (p-value < 0.05 and absolute [log2 (FC)]>=1 cutoff criteria). Then, we filtered 966 DEGs, including 398 up-regulated and 568 down-regulated genes, which were commonly detected by these algorithms (Figure 2B). The unsupervised clustering analysis demonstrated these 966 DEGs clustered into two major clusters (Figure 2C). Moreover, these DEGs clustered NEC samples and controls perfectly (Figure 2C). These robust 966 DEGs were considered for the construction of a gene co-expression network.

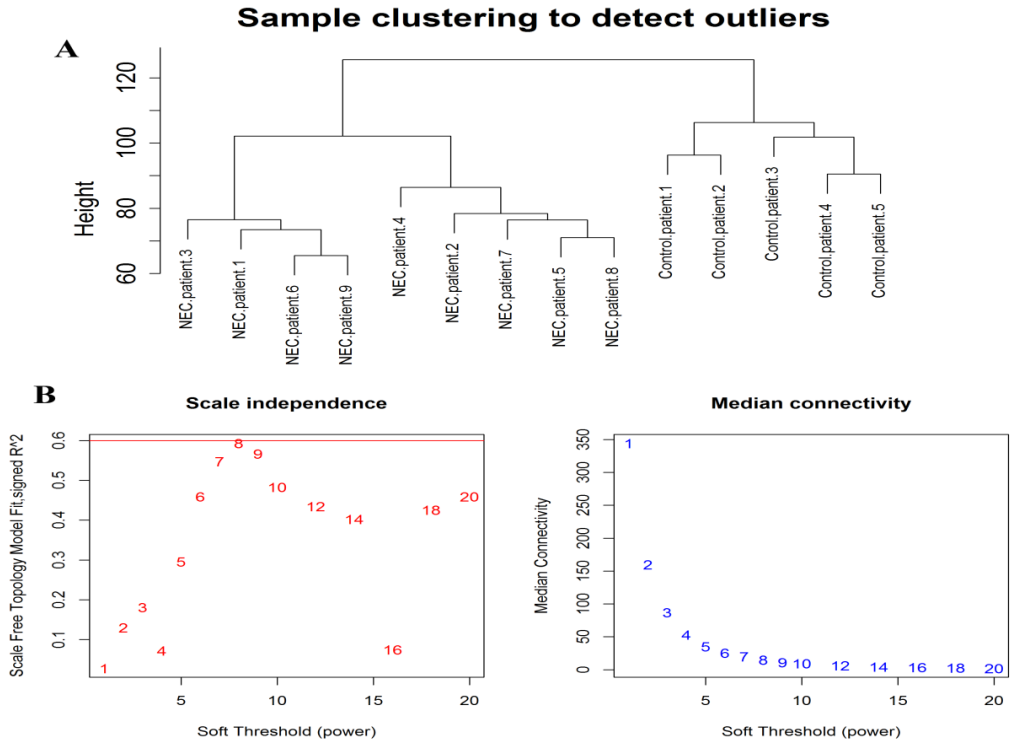

Figure 3. Samples clustering and determination of soft-thresholding power: (**A**) the cluster dendrogram of the samples; (**B**) analysis of the scale-free fit index (left) and the median connectivity (right) for various soft-thresholding powers.

3.1.2. Weighted gene co-expression networks construction and key module selection.

In this part, the gene co-expression network analysis had been executed with 966 DEGs for the GSE46801 dataset. First of all, the cluster dendrogram of samples was visualized by hclust R function to detect outlier samples, and no outlier sample was detected (Figure 3A). To identify the modules through WGCNA, we found the optimized soft thresholding powers β=8 as the scale-free topology criteria (Figure 3B). Considering this soft threshold power value, we constructed the co-expression networks. The analyses revealed the presence of nine co-expressed gene modules through the dynamic tree cut technique using deepSplit=2 and minClusterSize=30 parameters. Then, we sought to detect the key modules using the module merging approach. We used MEDissThres = 0.45 for merging similar modules and identified five key modules: black, blue, red, turquoise, and yellow (Figure 4A). We observed 403, 303, 127, 47, and 41 genes for blue, turquoise, yellow, red, and black modules, respectively, and the 45 non-co-expressed genes tied up in the grey module. The network heatmap of all genes with these 5 modules has been shown in Figure 4B. The interaction among these 5 modules has been shown in Figure 4C and Figure 4D. For the selection of key modules, we measured the module significance of all genes over the module. Figure 4E showed the barplot of gene significance of each module, and we chose blue, turquoise, and yellow modules as the key modules for their higher gene significance than the grey, black and red modules.

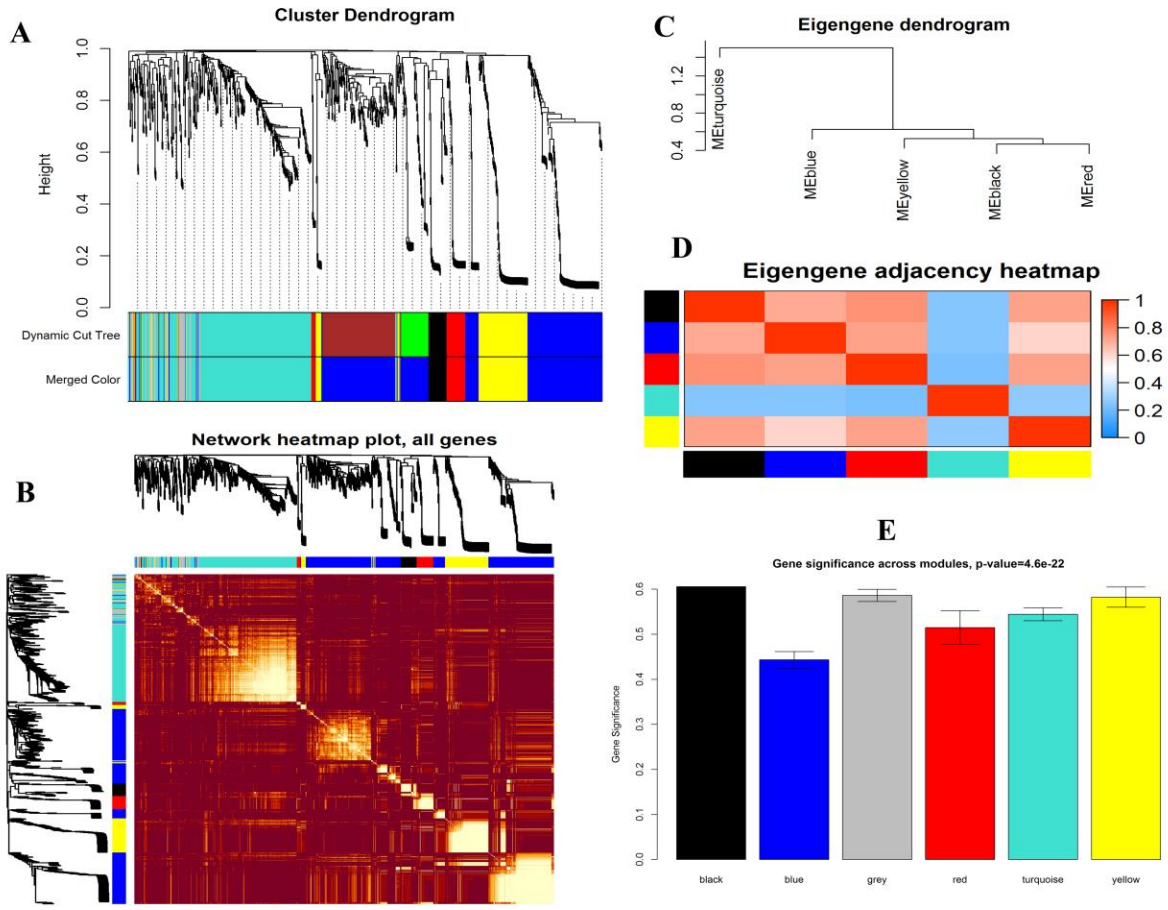

Figure 4. Gene co-expression modules construction with WGCNA: (**A**) the cluster dendrogram of the most connected genes furnished into gene co-expression modules based on dissimilarity matrix; (**B**) interaction relationship analysis co-expressed genes. Different colors of the horizontal axis and vertical axis represent different modules; (**C**) clustering dendrogram of module eigengene summarizes the modules yielded in the clustering analysis. (D) Module eigengene adjacencies network heatmap that summarizes the modules yielded in the clustering analysis. (E) The barplot of the gene significance across the modules.

3.1.3. Functional annotation analysis of gene modules.

The GO and pathway analyses have been conducted to obtain further biological insight into each key module's genes. For the blue module, the genes are significantly enriched in the ‘catabolic process’, ‘small molecules metabolic process’, ‘cellular localization’, ‘cytoplasm’, ‘catalytic activity’, ‘drug binding’, ‘ion binding’ (Figure 5A).

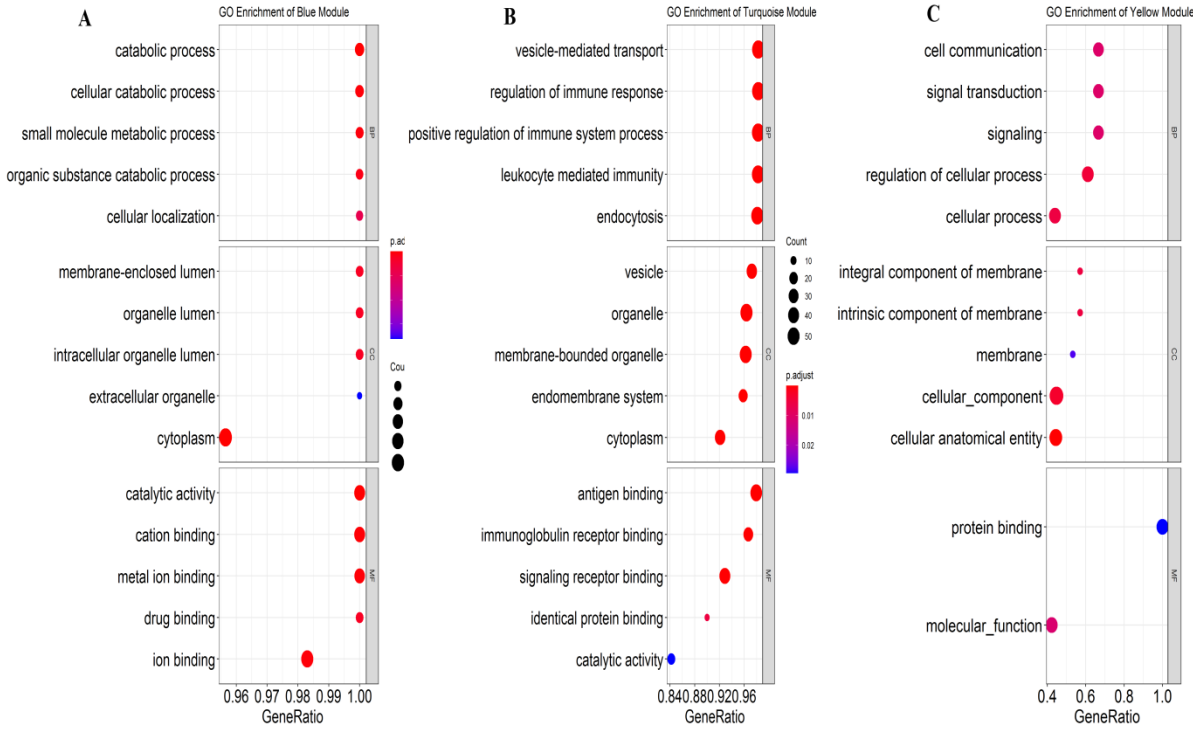

Figure 5. GO analyses of the selected three key modules: (**A**) GO analysis of all genes in the blue module; (**B**) GO analysis of all genes in the turquoise module; (**C**) GO analysis of all genes in the yellow module.

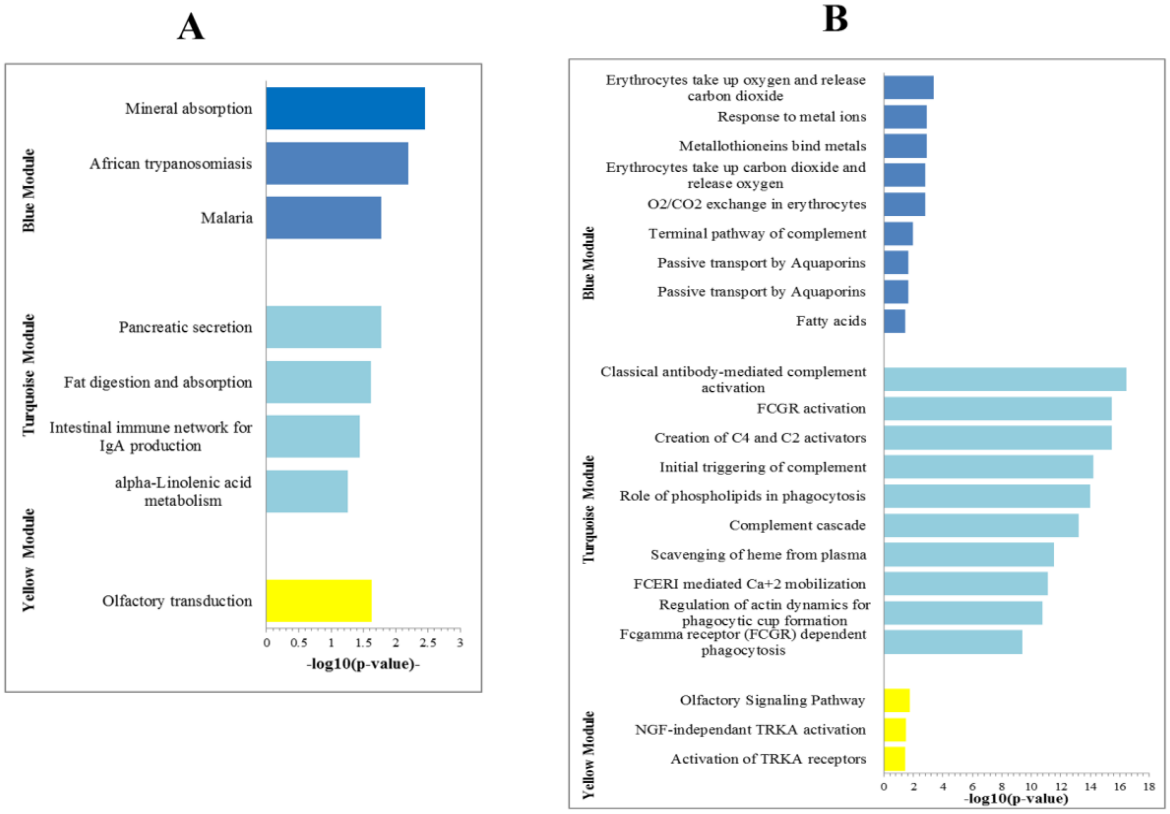

Figure 6. Pathway enrichment analyses of the selected three key modules (blue, turquoise and yellow): (**A**) KEGG pathway analysis of three modules; (**B**) Reactome pathway analysis of three modules.

The 'regulation of immune response', 'leukocyte mediated immunity', 'regulation of cellular process', 'endomembrane system', 'antigen-binding' and 'immunoglobulin receptor binding' are the significantly enriched terms of a turquoise module (Figure 5B). The significant enriched terms for the yellow module are the 'cell communication', 'signal transduction', 'cellular process', 'cellular anatomical entry', 'protein binding' and 'molecular function' (Figure 5C).

The genes of the blue module showed significant KEGG and Reactome pathways enrichment in "mineral absorption", "malaria", "african trypanosomiasis", "erythrocytes take up oxygen and release carbon dioxide", "response to metal ions' and "metallothioneins bind metals". For the turquoise module, the "pancreatic secretion", "fat digestion and absorption", "intestinal immune network for IgA production", "classical antibody-mediated complement activation", "FCGR activation and creation of C4 and C2 activators", "initial triggering of component" were the mainly enriched pathways. The yellow module demonstrated significant pathways, namely "the olfactory transduction", "olfactory signaling pathway", "NGF-independent TRKA activation", and "Activation of TRKA receptors" (Figure 6A and 6B and Table 1).

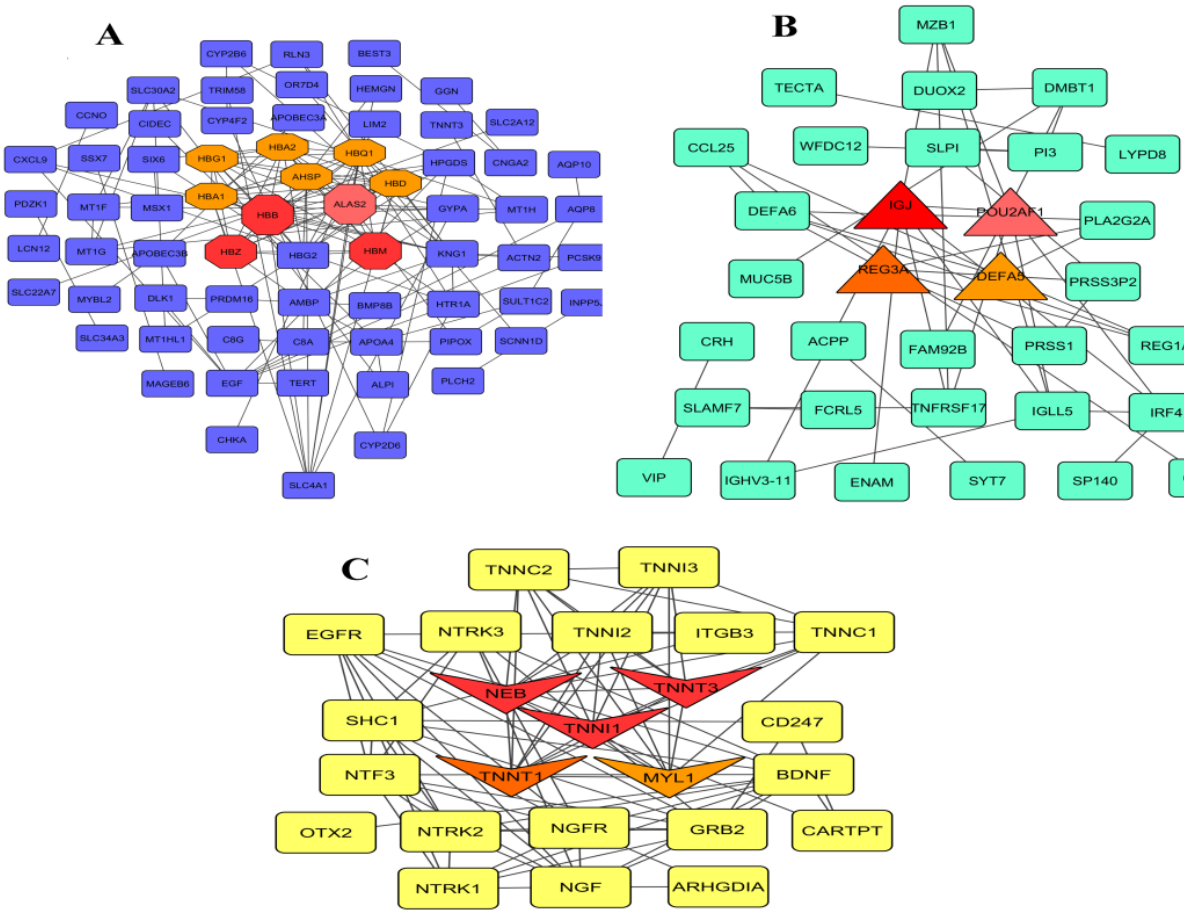

Figure 7. PPI analysis of the hub genes of (**A**) blue module; (**B**) turquoise module; (**C**) Yellow module.

3.1.4. PPI network analysis with hub genes detection in three key modules.

The DEGs for the three key modules were used to construct the PPI networks with the STRING database. The networks were visualized with Cytoscape software. The hub genes for

each module was detected with the MCC measurement criteria (MCC>=15). For the blue module, we construct a sub-network and identified 10 hub genes, namely, HBB, HBM, HBZ, ALAS2, HBA1, HBG1, HBA2, ASHP, HBQ1, and HBD (Figure 7A). We constructed a sub-network and identified IGJ, REG3A, POU2AF1, and DEFA5 hub genes for the turquoise module (Figure 7B). For the yellow module showed NEB, TNNT3, TNNI1, TNNT1, and MYL1 hub genes (Figure 7C).

3.1.5. Transcription factor and miRNAs identification for the three key modules.

To identify the transcriptomics factors and miRNAs of each key module of NEC disease, we analyzed the TFs-Genes and miRNAs-Gene network interaction in this section. The top 5 hub TFs and miRNAs were taken through MCC's raking values on CytoHubba in Cytoscape from the networks. The most significant TFs for the blue module are FOXC1, GATA2, YY1, TFAP2A, and FOXL1 (Figure 8A). The most significant TFs for the turquoise module are namely, REXO1L1P, FOXC1, CCL28, YY1, GATA2 (Figure 8B), and for the yellow module FOXC1, CECR1, OTX2, GATA2, and TMPRSS15 (Figure 8C). The most significant miRNAs detected for the blue module are namely, mir-27a-3p, mir-146a-5p, mir-335-5p, mir-124-3p and mir-129-2-3p (Figure 9A). The most significant miRNAs identified for the turquoise module are namely, mir-27a-3p, mir-335-5p, mir-124-3p, mir-334a-5p, mir-146a-5p (Figure 9B) and for the yellow module mir-27a-3p, mir-124-3p, mir-146a-5p, mir-34a-5p and mir-372-3p (Figure 9C).

3.2. Discussion

Despite much research to understand NEC's transcriptional changes to identify significant genes and pathways, a system biology approach to decode the NEC's molecular signature is not available yet. In this study, we employed gene co-expression analysis to decode the critical genes and pathways of NEC. The NEC RNA-seq transcriptomes were analyzed with integrative bioinformatics methods and revealed dynamic changes in transcription termed 966 DEGs as significant gene signature. The gene co-expression analysis demonstrated five significant key gene modules. These findings suggest dynamic transcriptional signatures of NECs. The GO and pathway annotation highlight the catabolic processes, catalytic process, dysregulated immune systems. The complement systems pathways, initial triggering of complement pathways, complement cascade, passive transport of aquaporin, etc., pathways were identified as novel therapeutic targets. These findings suggest the involvement of inflammation in NEC pathogenesis consistent with previous findings [23].

The hub genes identified by topological analysis of the PPI analysis are considered potential biomarkers and therapeutic targets. Therefore, we performed a PPI network analysis of the identified genes of three key modules. The topological assessment of the PPI network showed the presence of hub genes (HBB, HBM, HBZ, ALAS2, HBA1, HBG1, HBA2, ASHP, HBQ1, and HBD) in the blue module. Mutation in the hub gene HBB was identified by previous studies implicated in hemoglobinopathies such as thalassemia [54]. The hub HBZ plays a crucial role in the oncogenesis [55]. The hub genes HBM and ALAS2 were demonstrated down-regulated in NEC pigs [56]. The hub HBG1 is typically expressed in neonatal liver, spleen tissues and participates in various hemoglobinopathies (GeneCard database).

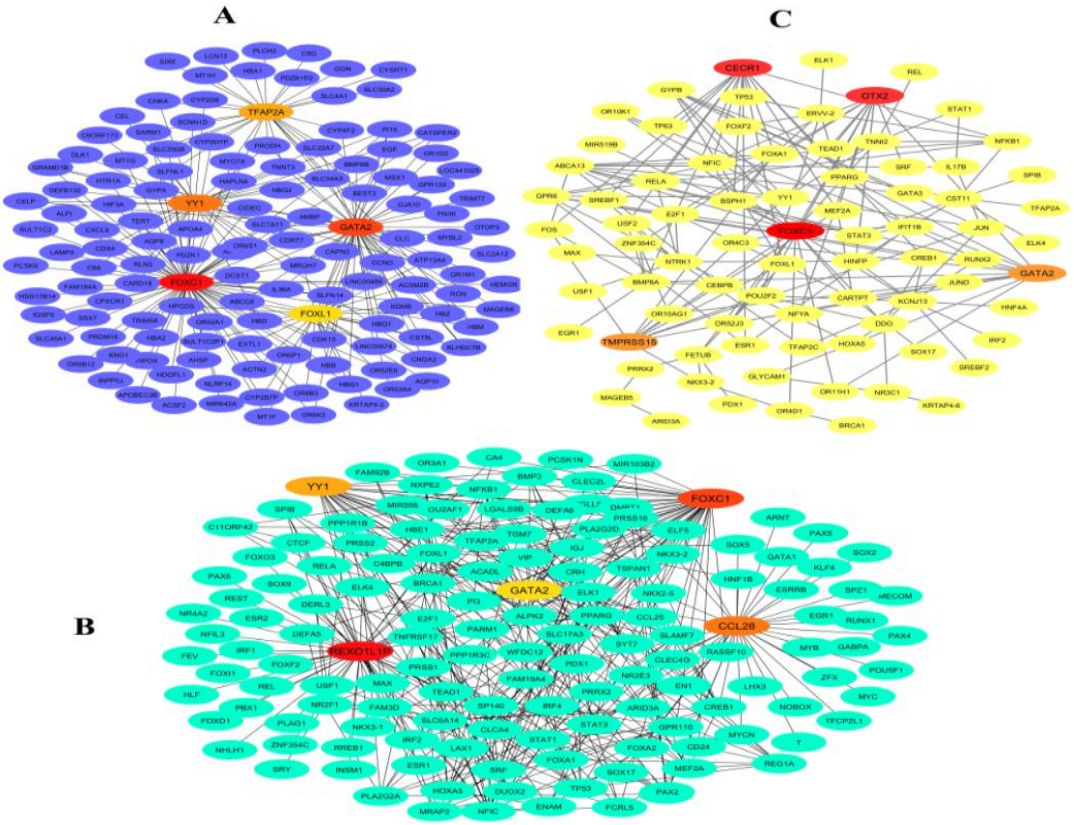

Figure 8. TFs-Genes interactions analysis with hub TFs of (**A**) blue module; (**B**) turquoise module; (**C**) yellow module.

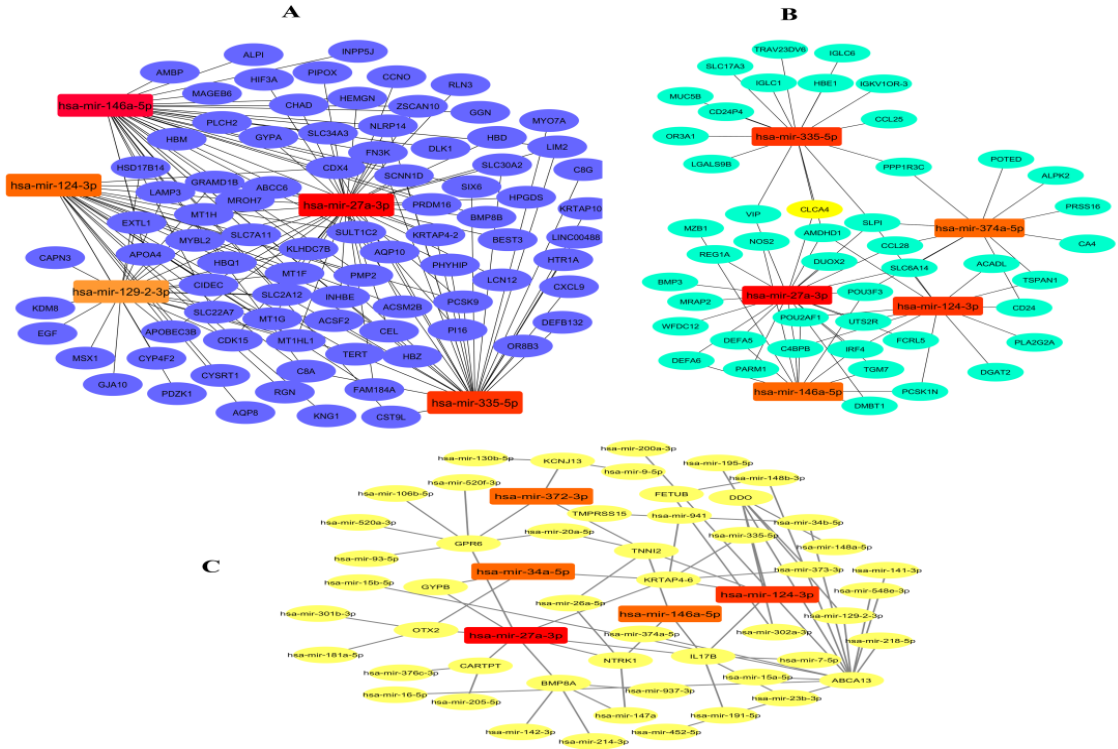

Figure 9. Gene-microRNA interactions analysis key gene co-expressed modules: (**A**) blue module; (**B**) turquoise module; (**C**) yellow module.

Table 1. Molecular pathways associated with co-expressed gene modules.

Database	Pathways	P-values	Related Genes
Blue Module KEGG	Mineral absorption	0.003557	SLC46A1;MT1F;MT1G;MT1H;MT1HL1
	African trypanosomiasis	0.006345	HBB;HBA2;HBA1;KNG1
	Malaria	0.016866	GYPA;HBB;HBA2;HBA1

Database	Pathways	P-values	Related Genes
Blue Module Reactome	Erythrocytes take up oxygen and release carbon dioxide	4.22E-04	HBB;SLC4A1;HBA1
	Metallothioneins bind metals	0.001188	MT1F;MT1G;MT1H
	Response to metal ions	0.001188	MT1F;MT1G;MT1H
	O2/CO2 exchange in erythrocytes	0.00156	HBB;SLC4A1;HBA1
	Erythrocytes take up carbon dioxide and release oxygen	0.00156	HBB;SLC4A1;HBA1
	Terminal pathway of complement	0.010465	C8G;C8A
	Passive transport by Aquaporins	0.02339	AQP10;AQP8
	Fatty acids	0.035766	CYP2D6;CYP4F2
Turquoise Module KEGG	Pancreatic secretion	0.016717	PRSS1;PLA2G2D;PLA2G2A;PRSS2;CLCA4
	Fat digestion and absorption	0.024021	PLA2G2D;DGAT2;PLA2G2A
	Intestinal immune network for IgA production	0.036096	CCL25;TNFRSF17;CCL28
	alpha-Linolenic acid metabolism	0.054599	PLA2G2D;PLA2G2A
Turquoise Module Reactome	Classical antibody-mediated complement activation	3.72E-17	IGHV3-23;IGLC6;IGKV1-5;IGKV1D-16;IGLV4-60; IGHG3;IGLV4-69;IGHG1;IGLV1-40;IGHG2; IGLV7-46;IGLV2-23;IGKV4-1;IGLV2-18;IGLC1
	Creation of C4 and C2 activators	3.56E-16	IGHV3-23;IGKV1-5;IGLC6;IGKV1D-16;IGLV4-60; IGHG3;IGLV4-69;IGHG1;IGHG2;IGLV1-40; IGLV7-46;IGLV2-23;IGKV4-1;IGLV2-18;IGLC1
	FCGR activation	3.56E-16	IGHV3-23;IGKV1-5;IGLC6;IGKV1D-16;IGLV4-60; IGHG3;IGLV4-69;IGHG1;IGHG2;IGLV1-40; IGLV7-46;IGLV2-23;IGKV4-1;IGLV2-18;IGLC1
	Initial triggering of complement	5.95E-15	IGHV3-23;IGKV1-5;IGLC6;IGKV1D-16;IGLV4-60;I GHG3;IGLV4-69;IGHG1;IGHG2;IGLV1-40; IGLV7-46;IGLV2-23;IGKV4-1;IGLV2-18;IGLC1
	Role of phospholipids in phagocytosis	1.04E-14	IGHV3-23;IGKV1-5;IGLC6;IGKV1D-16;IGLV4-60 ;IGHG3;IGLV4-69;IGHG1;IGHG2;IGLV1-40; IGLV7-46;IGLV2-23;IGKV4-1;IGLV2-18;IGLC1
	Complement cascade	5.84E-14	IGHV3-23;IGKV1-5;IGLC6;C4BPB;IGKV1D-16 ;IGLV4-60;IGHG3;IGLV4-69;IGHG1;IGHG2;IGLV1-40 ;IGLV7-46;IGLV2-23;IGKV4-1;IGLV2-18;IGLC1
	Scavenging of heme from plasma	2.89E-12	IGLV4-69;IGLV1-40;IGLV7-46;IGHV3-23;IGLC6 ;IGKV1-5;IGLV2-23;IGKV4-1;IGLV2-18 ;IGLC1;IGKV1D-16;IGLV4-60
	FCERI mediated Ca+2 mobilization	7.32E-12	IGHV3-23;IGKV1-5;IGLC6;IGKV1D-16;IGLV4-60; IGLV4-69;IGLV1-40;IGLV7-46;IGLV2-23;IGKV4-1 ;IGHE;IGLV2-18;IGLC1
	Regulation of actin dynamics for phagocytic cup formation	1.82E-11	IGHV3-23;IGKV1-5;IGLC6;IGKV1D-16;IGLV4-60 ;IGHG3;IGLV4-69;IGHG1;IGHG2;IGLV1-40; IGLV7-46;IGLV2-23;IGKV4-1;IGLV2-18;IGLC1
	Fcgamma receptor (FCGR) dependent phagocytosis	4.08E-10	IGHV3-23;IGKV1-5;IGLC6;IGKV1D-16;IGLV4-60; IGHG3;IGLV4-69;IGHG1;IGHG2;IGLV1-40; IGLV7-46;IGLV2-23;IGKV4-1;IGLV2-18;IGLC1
Yellow Module KEGG	Olfactory transduction	0.023333	OR10K1
Yellow Module Reactome	Olfactory Signaling Pathway	0.017423	OR10K1;OR4D1;OR52J3;OR1S1;OR11H1;OR10AG1; OR4C3
	NGF-independant TRKA activation	0.031352	NTRK1
	Activation of TRKA receptors	0.037505	NTRK1

The hub HBA2 was identified as a critical gene involved in NEC's inflammatory process [25]. The PPI analysis demonstrated IGJ, REG3A, POU2AF1, and DEFA5 hub genes for the turquoise module. The hub REG3A was identified as a critical gene involved in NEC's inflammatory process [25]. The hub POU2AF1 was previously identified as deregulated in ileal Peyer's small intestine patches [57]. The hub DEFA5 was identified as a critical gene involved

in NEC's inflammatory process [25]. The yellow module showed that NEB, TNNT3, TNNI1, TNNT1, and MYL1 are the hub genes. The hub NEB is involved in acute pancreatitis of leukemia patients [58]. Genetic variation of TNNT3 is involved in breast cancer [59]. Genetic variation in TNNT1 is associated with bodyweight [60]. A genetic variant of MYL1 is implicated in metabolic traits [61]. The transcriptomics factors (TFs) identified by topological analysis of the TFs-Gene interaction analysis are considered potential TFs in NEC. We performed the TFs-Gene interaction analysis for the selected three key modules. The most significant TFs detected for the blue module are FOXC1, GATA2, YY1, TFAP2A, and FOXL1. The most significant TFs detected for the turquoise module are REXO1L1P, FOXC1, CCL28, YY1, and GATA2 and the yellow module FOXC1, CECR1, OTX2, GATA2, and TMPRSS15. The identified microRNAs may control the co-expressed genes. The identified hub genes, TFs, miRNAs, and pathways may be considered for NEC's drug targets and potential biomarkers. Despite the crucial findings obtained from this co-expression study, we acknowledge the limitations of wet-lab experimental validation of these identified molecules should be performed before establishing them for clinical use.

4. Conclusions

The NEC is a severe disease of gastrointestinal systems with high morbidity. The present study employed gene co-expression bioinformatics methods to analyze RNA-Seq transcriptomes of NEC matched with controls to identify significant co-expressed gene modules and pathways. The WGCNA gene co-expression analysis demonstrated three key co-expressed gene modules. Further, the PPI analysis showed the presence of several hub genes (HBB, HBM, HBZ, ALAS2, HBA1, HBG1, HBA2, ASHP, HBQ1, and HBD) in the blue module; IGJ, REG3A, POU2AF1, DEFA5 hub genes for the turquoise module; the yellow module showed NEB, TNNT3, TNNI1, TNNT1 and MYL1 hub genes. The GO and pathway analysis highlights the inflammation process in the development and progression of NEC. We also detected TFs and miRNAs that may regulate the identified co-expressed gene modules. The identified hub genes, regulators, and pathways shed new light on the pathogenesis of NEC. Despite these candidate biomarkers' computational significance, we now propose a wet-lab experiment to establish them as biomarkers for clinical use.

Funding

This research received no external funding.

Acknowledgments

This work is supported by the National Nature Sciences Foundation of China (12071096).

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Neu, J.; Walker, W. Necrotizing enterocolitis. *N Engl J Med.* **2011**, *364*, 255–64, doi:10.1056/NEJMra1005408.
2. Bellodas Sanchez, Jenny Kadrofske, M. Necrotizing enterocolitis. *https://onlinelibrary.wiley.com/doi/epdf/10.1111/nmo.13569* **2019**, *31*, e13569,

doi:doi.org/10.1111/nmo.13569.
3. Berman, L.; Moss, R. Necrotizing enterocolitis: an update. *Semin Fetal Neonatal Med.* **2011**, *16*, 145–50, doi:doi.org/10.1016/j.siny.2011.02.002.
4. Verma, R.P.; Kota, A. Necrotizing Enterocolitis. In *Pediatric Surgery, Flowcharts and Clinical Algorithms*; IntechOpen, **2019**.
5. Huda, S.; Chaudhery, S.; Ibrahim, H.; Pramanik, A. Neonatal necrotizing enterocolitis: Clinical challenges, pathophysiology and management. *Pathophysiology* **2014**, *21*, 3–12, doi:10.1016/j.pathophys.2013.11.009.
6. Bazacliu, C.; Neu, J. Necrotizing Enterocolitis: Long Term Complications. *Curr. Pediatr. Rev.* **2019**, *15*, 115–124, doi:doi.org/10.2174/1573396315666190312093119.
7. Afrazi, A.; Sodhi, C.; Richardson, W.; Neal, M.; Good, M.; Siggers, R. New insights into the pathogenesis and treatment of necrotizing enterocolitis: Toll-like receptors and beyond. *Pediatr. Res.* **2011**, *69*, 183–8, doi:10.1203/PDR.0b013e3182093280.
8. Grave, G.; Nelson, S.; Walker, W.; Moss, R.; Dvorak, B.; Hamilton, F.; Al., E. New therapies and preventive approaches for necrotizing enterocolitis: report of a research planning workshop. *Pediatr ic Res.* **2007**, *62*, 510–4, doi:10.1203/PDR.0b013e318142580a.
9. Henry, M.; Moss, R. Necrotizing enterocolitis. *Annu. Rev. Med.* **2009**, *60*, 111–24, doi:10.1146/annurev.med.60.050207.092824.
10. Humberg, A.; Fortmann, I.; Siller, B.; et al. Pre-term birth and sustained inflammation: consequences for the neonate. *Semin. Immunopathol.* **2020**, *42*, 451–468, doi:doi.org/10.1007/s00281-020-00803-2.
11. Lin, L.; Xia, X.; Liu, W.; Wang, Y.; Hua, Z. Clinical characteristics of neonatal fulminant necrotizing enterocolitis in a tertiary Children's hospital in the last 10 years. *PLoS One* **2019**, *14*, e0224880, doi:doi.org/10.1371/journal.pone.0224880.
12. Ng, P.; Ang, I.; Chiu, R.; Li, K.; Lam, H.; Wong, R.; Al., E. Host-response biomarkers for diagnosis of late-onset septicemia and necrotizing enterocolitis in pre-term infants. *J. Clin. Invest.* **2010**, *120*, 2989–3000.
13. Bell, M.; Ternberg, J.; Feigin, R.; Keating, J.; Marshall, R.; Barton, L.; Al., E. Neonatal necrotizing enterocolitis. Therapeutic decisions based upon clinical staging. *Ann. Surg.* **1978**, *187*, 1–7.
14. Pietz, J.; Achanti, B.; Lilien, L.; Stepka, E.; Mehta, S. Prevention of necrotizing enterocolitis in pre-term infants: a 20-year experience. *Pediatrics* **2007**, *119*, e164–e170, doi:10.1542/peds.2006-0521.
15. Young, C.; Sharma, R.; Handfield, M.; Mai, V.; Neu, J. Biomarkers for infants at risk for necrotizing enterocolitis: clues to prevention? *Pediatr ic Res.* **2009**, *65*, 91R–7R.
16. Sharma, R.; Hudak, M. A clinical perspective of necrotizing enterocolitis: past, present, and future. *Clin. Perinatol.* **2003**, *40*, 27–51.
17. Wang, K.; Tao, G.; Sun, Z.; Sylvester Karl, G. Recent Potential Noninvasive Biomarkers in Necrotizing Enterocolitis. *Gastroenterol. Res. Pract.* **2019**, *2019*, 9, doi:doi.org/10.1155/2019/8413698.
18. Chen, W.; Yan, X.; Tian, T.; Yan, R.; Wang, X.; Yu, Z.; Li, Y.; Zhang, L.; Han, S. Integrated analysis of a lncRNA-mRNA network reveals a potential mechanism underlying necrotizing enterocolitis. *Mol. Med. Rep.* **2020**, *22*, 423–435.
19. Jones, I.; Hall, N. Contemporary Outcomes for Infants with Necrotizing Enterocolitis-A Systematic Review. *J. Pediatr.* **2020**, *86*, 86-92.e3, doi:doi.org/10.1016/j.jpeds.2019.11.011.
20. Alsaied, A.; Islam, N.; Thalib, L. Global incidence of Necrotizing Enterocolitis: a systematic review and Meta-analysis. *BMC Pediatr.* **2020**, *20*, 344, doi:doi.org/10.1186/s12887-020-02231-5.
21. Sonja, D.; Lea, T.; Christel, W.; Julia, H.; Hanna, M.; Manuel, B. Clinical Characteristics of Necrotizing Enterocolitis in Pre-term Patients With and Without Persistent Ductus Arteriosus and in Patients With Congenital Heart Disease. *Front. Pediatr.* **2020**, *8*, 257, doi:doi.org/10.3389/fped.2020.00257.
22. Eleni, A.; Charalampos, A.; Helen, G.; Kosmas, S. Emerging Biomarkers for Prediction and Early Diagnosis of Necrotizing Enterocolitis in the Era of Metabolomics and Proteomics. *Front. Pediatr.* **2020**, *8*, 838, doi:doi.org/10.3389/fped.2020.602255.
23. Chan, K.; Leung, K.; Tam, Y.; Lam, H.; Cheung, H.; Ma, T.; Al., E. Genome-wide expression profiles of necrotizing enterocolitis versus spontaneous intestinal perforation in human intestinal tissues: dysregulation of functional pathways. *Ann. Surg.* **2014**, *260*, 1128–37, doi:https://doi.org/10.1097/SLA.0000000000000374.
24. Ng, P.; Ma, T.; Lam, H. The use of laboratory biomarkers for surveillance, diagnosis and prediction of clinical outcomes in neonatal sepsis and necrotising enterocolitis. *Arch. Dis. Child. - Fetal Neonatal Ed.* **2015**, *100*, F448–52.
25. Tremblay, É.; Thibault, M.; Ferretti, E.; Babakissa, C.; Bertelle, V.; Bettolli, M.; Burghardt, K.M.; Colombani, J.; Grynspan, D.; Levy, E.; et al. Gene expression profiling in necrotizing enterocolitis reveals pathways common to those reported in Crohn ' s disease. *BMC Med. Genomics* **2016**, *9*,

doi:10.1186/s12920-016-0166-9.

26. Ng, P.; Chan, K.; Leung, K.; Al., E. Comparative MiRNA Expressional Profiles and Molecular Networks in Human Small Bowel Tissues of Necrotizing Enterocolitis and Spontaneous Intestinal Perforation. *PLoS One* **2015**, *10*.
27. Guanglin, C.; Yang, L.; Yang, S.; Lingling, Z.; Hua, Z.; Qiyang, S.; Chunxia, D.; Hongxin, L.; Zechao, W.; Yankai, X.; et al. Identification of candidate genes for necrotizing enterocolitis based on microarray data. *Gene* **2018**, *661*, 152–159, doi:10.1016/j.gene.2018.03.088.
28. Jung, K.; Koh, I.; Kim, J.; Al., E. RNA-Seq for Gene Expression Profiling of Human Necrotizing Enterocolitis: a Pilot Study. *J Korean Med Sci.* **2017**, *32*, 817–824.
29. Jung, K.; Kim, J.; Cheong, H.; et al. Gene expression profile of necrotizing enterocolitis model in neonatal mice. *Int. J. Surg.* **2015**, *23*, 28–34.
30. Liu, W.; Li, L.; Ye, H. Weighted gene co-expression network analysis in biomedicine research. *Chin. J. Biotechnol.* **2017**, *33*, 1791–1801.
31. Zhao, W.; Langfelder, P.; Fuller, T.; Dong, J.; Li, A.; Hovarth, S. Weighted gene co-expression network analysis: state of the art. *J. Biopharm. Stat.* **2010**, *20*, 281–300, doi:10.1080/10543400903572753.
32. Feng, T.; Li, K.; Zheng, P.; Wang, Y.; Lv, Y.; Shen, L.; Chen, Y.; Xue, Z.; Li, B.; Jin, L.; et al. Weighted Gene Coexpression Network Analysis Identified MicroRNA Coexpression Modules and Related Pathways in Type 2 Diabetes Mellitus. *Oxid. Med. Cell. Longev.* **2019**, doi:10.1155/2019/9567641.
33. Wang, L.; Hu, J.; Zhou, J.; Guo, F.; Yao, T.; Zhang, L. Weighed Gene Coexpression Network Analysis Screens the Potential Long Noncoding RNAs and Genes Associated with Progression of Coronary Artery Disease. *Comput. Math. Methods Med.* **2020**, doi:doi.org/10.1155/2020/8183420.
34. He, B.; Xu, J.; Tian, Y.; Liao, B.; Lang, J.; Lin, H.; Mo, X.; Lu, Q.; Tian, G.; Bing, P. Gene Coexpression Network and Module Analysis across 52 Human Tissues. *Biomed Res. Int.* **2020**, 14, doi:doi.org/10.1155/2020/6782046.
35. Qu, S.; Shi, Q.; Xu, J.; Yi, W.; Fan, H. Weighted Gene Coexpression Network Analysis Reveals the Dynamic Transcriptome Regulation and Prognostic Biomarkers of Hepatocellular Carcinoma. *Evol. Bioinforma.* **2020**, *16*, doi:doi.org/10.1177/1176934320920562.
36. Pan, Y.; Zhang, Q.; Deng, X.; An, N.; Du, X.; Liu, J. Gene co-expression network analysis revealed biomarkers correlated with blast cells and survival in acute myeloid leukemia. *Mol. Clin. Oncol.* **2020**, *12*, 475–484, doi:doi.org/10.3892/mco.2020.2006.
37. Jin, X.; Li, J.; Li, W.; Wang, X.; Du, C.; Geng, Z.; Geng, Y.; Kang, L.; Zhang, X.; Wang, M.; et al. Weighted gene co-expression network analysis reveals speci fi c modules and biomarkers in Parkinson ' s disease. *Neurosci. Lett.* **2020**, *728*, 134950, doi:10.1016/j.neulet.2020.134950.
38. Hu, R.; Yu, Q.; Zhou, S.; Yin, Y.; Hu, R.; Lu, H.; Hu, B. Co-expression Network Analysis Reveals Novel Genes Underlying Alzheimer's Disease Pathogenesis. *Front. Aging Neurosci.* **2020**, *12*, 605961, doi:doi.org/10.3389/fnagi.2020.605961.
39. Lucchetta, M.; Pellegrini, M. Finding disease modules for cancer and COVID-19 in gene co-expression networks with the Core&Peel method. *Sci. Rep.* **2020**, *10*, 17628, doi:doi.org/10.1038/s41598-020-74705-6.
40. Min, J.; Jingtian, L.; Quan, Z.; Xiufeng, X.; Ruidong, L.; Peikang, D.; Chun, M.; Yi, L.; Lijuan, W.; Wanpeng, Q.; et al. Identification of Four Potential Biomarkers Associated With Coronary Artery Disease in Non-diabetic Patients by Gene Co-expression Network Analysis. *Front. Genet.* **2020**, *11*, 542, doi:doi.org/10.3389/fgene.2020.00542.
41. Robinson, M.; McCarthy, D.; Smyth, G. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140, doi:10.1093/bioinformatics/btp616.
42. Love, M.I.; Huber, W.; Anders, S. oderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Geneome Biol.* **2015**, *550*.
43. Zhou, X.; Lindsay, H.; Robinson, M. Robustly detecting differential expression in RNA sequencing data using observation weights. *Nucleic Acids Res.* **2014**, *42*, e91.
44. Langfelder, P.; Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **2008**, *9*, 559.
45. Yu, G.; Wang, L.; Han, Y.; He, Q. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omi. A J. Integr. Biol.* **2012**, *16*, 284–287, doi:10.1089/omi.2011.0118.
46. Kuleshov, M. V.; Jones, M.R.; Rouillard, A.D.; Fernandez, N.F.; Duan, Q.; Wang, Z.; Koplev, S.; Jenkins, S.L.; Jagodnik, K.M.; Lachmann, A.; et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **2016**, *44*, W90–W97, doi:10.1093/nar/gkw377.
47. Szklarczyk, D.; Morris, J.H.; Cook, H.; Kuhn, M.; Wyder, S.; Simonovic, M.; Santos, A.; Doncheva, N.T.; Roth, A.; Bork, P.; et al. The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **2017**, *45*, doi:10.1093/nar/gkw937.

48. Smoot, M.E.; Ono, K.; Ruscheinski, J.; Wang, P.L.; Ideker, T. Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics* **2011**, *27*, 431–432, doi:10.1093/bioinformatics/btq675.
49. Chin, C.; Chen, S.; Wu, H.; Ho, C.; Ko, M.; Lin, C. cytoHubba : identifying hub objects and sub- networks from complex interactome. *BMC Syst. Biol.* **2014**, *8*, :S11, doi:10.1186/1752-0509-8-S4-S11.
50. Khan, A.; Fornes, O.; Stigliani, A.; Gheorghe, A.; Castro-Mondragon, J.A.; Van Der Lee, R.; Bessy, A.; Chèneby, J.; Kulkarni, S.R.; Tan, G.; et al. JASPAR 2018: Update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* **2018**, *46*, D260–D266, doi:10.1093/nar/gkx1126.
51. Xia, J.; Gill, E.E.; Hancock, R.E.W. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. *Nat. Protoc.* **2015**, *10*, 823–844.
52. Sethupathy, P.; Corda, B.; Hatzigeorgiou, A.G. TarBase: A comprehensive database of experimentally supported animal microRNA targets. *RNA* **2006**, *12*, 192–197.
53. Hsu, S.-D.; Lin, F.-M.; Wu, W.-Y.; Liang, C.; Huang, W.-C.; Chan, W.-L.; Tsai, W.-T.; Chen, G.-Z.; Lee, C.-J.; Chiu, C.-M.; et al. miRTarBase: a database curates experimentally validated microRNA–target interactions. *Nucleic Acids Res.* **2011**, *39*, D163–D169, doi:10.1093/nar/gkq1107.
54. Carlice-dos-Reis, T.; Viana, J.; Moreira, F.; Cardoso, G.; Guerreiro, J.; Santos, S. Investigation of mutations in the HBB gene using the 1,000 genomes database. *PLoS One* **2017**, *12*, e0174637, doi:0.1371/journal.pone.0174637.
55. Zhao, T.; Matsuoka, M. HBZ and its roles in HTLV-1 oncogenesis. *Front. Microbiol.* **2012**, *3*, doi:10.3389/fmicb.2012.00247.
56. Sun, J.; Pan, X.; Christiansen, L.I.; Yuan, X.L.; Skovgaard, K.; Chatterton, D.; Kaalund, S.S.; Gao, F.; Sangild, P.T.; Pankratova, S. Necrotizing enterocolitis is associated with acute brain responses in pre-term pigs. Journal of neuroinflammation. *J. Neuroinflammation* **2018**, *15*, doi:10.1186/s12974-018-1201-x.
57. Mach, N.; Berri, M.; Esquerré, D.; Chevaleyre, C.; Lemonnier, G.; Billon, Y. Extensive Expression Differences along Porcine Small Intestine Evidenced by Transcriptome Sequencing. *PLoS One* **2014**, *9*, e88515, doi:10.1371/journal.pone.0088515.
58. Liu, C.; Yang, W.; Devidas, M.; Al., E. Clinical and Genetic Risk Factors for Acute Pancreatitis in Patients With Acute Lymphoblastic Leukemia. *J Clin Oncol.* **2016**, *34*, 2133–2140, doi:10.1200/JCO.2015.64.5812.
59. Turnbull, C.; Ahmed, S.; Morrison, J.; Al., E. Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat Genet.* **2010**, *42*, 504–507, doi:10.1038/ng.586.
60. Akiyama, M.; Ishigaki, K.; Sakaue, S.; Al., E. Characterizing rare and low-frequency height-associated variants in the Japanese population. *Nat Commun* **2019**, *10*, doi:10.1038/s41467-019-12276-5.
61. Rueedi, R.; Ledda, M.; Nicholls, A.; Salek, R.; Marques-Vidal, P.; Morya, E.; Al., E. Genome-Wide Association Study of Metabolic Traits Reveals Novel Gene-Metabolite-Disease Links. *PLoS Genet* **2014**, *10*, e1004132, doi:10.1371/journal.pgen.1004132.