International Journal of Biomedicine 15(4) (2025) 704-714
http://dx.doi.org/10.21103/Article15(4)_OA10

ORIGINAL ARTICLE

Genomics and Proteomics

INTERNATIONAL JOURNAL OF BIOMEDICINE

Using Integrated Bioinformatics Strategy to Identify Differentially Expressed Genes and Hub Genes of Human Hosts with Tuberculosis

Peng Yue[1], Yan Dong[2], Fukai Bao[1,2*], Aihua Liu[1*]

[1]*Faculty of Basic Medicine, Kunming Medical University, Kunming 650500, China*
[2]*Department of Microbiology and Immunology, Haiyuan College, Kunming Medical University, Kunming 650101, China*

Abstract

Background: To date, the molecular mechanisms underlying the occurrence, development, and prognosis of tuberculosis remain incompletely understood. The study aimed to identify the host hub involved in tuberculosis.

Methods and Results: Four gene expression profiles (GSE51029, GSE52819, GSE54992, and GSE65517) were downloaded from Gene Expression Omnibus (GEO). First, the selected data sets of the *Mycobacterium tuberculosis* (MTB) infection group and the healthy control group were analyzed through GEO2R, and the genes that met the following conditions: |log FC|> 1 and P-values <0.05, are considered differentially expressed genes (DEGs). Secondly, the DEGs shared by the 4 microarray datasets were further identified. Next, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed for functional enrichment analysis of these DEGs, the host hub genes were identified by the Cytohubba plugin, and module networks in DEG networks were screened by the plugin Molecular Complexity Detection (MCODE). Other bioinformatics methods were performed, including protein-protein interaction (PPI) network analysis and the construction of miRNA-hub gene networks and transcription factor (TF)-hub gene networks. Finally, the expression of the host hub genes was verified by real-time PCR.

Four GEO microarray datasets were integrated, and a total of 46 DEGs were identified. The results of the GO analysis showed that the biological functions of DEGs were primarily involved in regulating the immune response process, cytokine/chemokine activity, and receptor-ligand activity. DEGs were also significantly enriched in membrane rafts, the mitochondrial outer membrane, cytoplasmic vesicle cavities, and nuclear chromatin. KEGG enrichment analysis showed that the NOD-like receptor signaling pathway and the Toll-like receptor signaling pathway were 2 important pathways. In addition, 5 highly differentially expressed hub genes, *STAT1, TLR7, CXCL8, CCR2*, and *CCL20*, were screened out. Finally, based on the NetworkAnalyst database, we screened targeted miRNAs and TF of hub genes and found that hsa-miR-335-3p may play a key role in the regulation of these hub genes.

Conclusion: In summary, bioinformatics analyses were used to identify DEGs to find potential biomarkers that may be associated with tuberculosis. This study provides a set of candidate DEGs and 5 essential host hub genes that can be potentially useful for early detection, prognostic determination, risk assessment, and targeted tuberculosis therapy.(**International Journal of Biomedicine. 2025;15(4):704-714.**)

Keywords: tuberculosis • *Mycobacterium tuberculosis* • GEO dataset • miRNA • hub gene network • bioinformatics analysis

For citation: Yue P, Dong Y, Bao F, Liu A. Using Integrated Bioinformatics Strategy to Identify Differentially Expressed Genes and Hub Genes of Human Hosts with Tuberculosis. International Journal of Biomedicine. 2025;15(4):704-714. doi:10.21103/Article15(4)_OA10

Abbreviations

BP, biological processes; **CC**, cell components; **DEGs**, differentially expressed genes; **GEO**, Gene Expression Omnibus; **GO**, Gene Ontology; **IGRA**, interferon gamma release assay; **KEGG**, Kyoto Encyclopedia of Genes and Genomes; **MTB**, *Mycobacterium tuberculosis;* **MCODE**, Molecular Complexity Detection; **MF**, molecular functions; **PPI**, protein-protein interaction; **PBS**, phosphate-buffered saline; **STRING**, Search Tool for Retrieval of Interacting Genes; **TF**, transcription factor; **TB**, tuberculosis; **TST**, tuberculin skin test.

Introduction

Tuberculosis (TB) is a chronic infectious disease caused by *Mycobacterium tuberculosis* (MTB), which can involve many organs, but pulmonary TB is the most common infection. According to the World Health Organization, an estimated 10 million people worldwide were infected with TB in 2020, of whom 7.1 million were newly diagnosed and reported as TB cases.[1] At the same time, the incidence of TB varies from less than 5 to more than 500 cases per 100,000 population per year, and TB remains a deadly disease even in developing countries with well-established healthcare systems.

Infection with MTB causes clinical signs and symptoms when host defense is reduced or cell-mediated allergy is increased.[2] Respiratory symptoms include cough, sputum, hemoptysis, chest pain, varying degrees of chest tightness, or dyspnea. Sputum smear microscopy, bacterial culture, and MTB isolation are the most traditional, classical, and widely used diagnostic tools for TB; however, these tools require a significant amount of time, and their accuracy is not high, making it difficult to achieve early diagnosis and effective treatment of TB patients.[3] The tuberculin skin test (TST) is widely used to detect latent TB infection; however, IFN-γ release assay (IGRA), which relies on in vitro detection of a single cytokine induced by MTB-specific antigen, has been used as an alternative to the TST in the diagnosis of MTB infection.[4] However, the TST and IGRA are essentially unable to distinguish between active and latent TB infections.[5]

In recent years, rapidly developing microarray technology has been widely utilized to compare gene expression levels, predict disease progression, and facilitate accurate diagnosis and prognosis evaluations.[6,7] With the widespread use of gene expression microarray technology, a large amount of data has been published on public database platforms. The integration of these databases can be used to investigate further the molecular mechanisms involved in disease. Gene expression microarrays offer a novel approach to studying disease-related genes, providing promising prospects for molecular prediction, drug-based molecular targeting, and molecular therapy.[8] Therefore, it is essential to investigate the potential molecular mechanisms underlying TB's biological behavior to develop more effective early diagnostic techniques with high sensitivity, as well as more reliable and specific novel biomarkers to monitor recurrence and assess prognosis. There is a strong need to identify new potential diagnostic and therapeutic biomarkers for TB patients, which will not only provide new insights into the molecular and cellular processes involved in pathogenesis but also establish rapid, sensitive, and effective methods for diagnosing and treating TB.[9]

With the development of genomics technology, a large amount of data has been generated in the field of TB research.[10] In this study, we screened the differentially expressed genes(DEGs) between the MTB infection group and the healthy control group in 4 separate Gene Expression Omnibus (GEO) data sets. Then we performed Gene Ontology (GO) function enrichment analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis, and protein-protein interaction (PPI) network construction and module analysis. The results of this study may help to explore potential targets for the diagnosis and treatment of TB.

Materials and Methods

Microarray Dataset

The GEO database (http://www.ncbi.nlm.nih.gov/geo) is a free public genomics database that contains a variety of data, including microarray and next-generation sequencing data. We used the following keywords to search the GEO database: "tuberculosis"[MeSH Terms] OR tuberculosis [All Fields]) AND "Homo sapiens"[Organism]. The inclusion criteria of gene expression profile are as follows: The data set must include samples of peripheral blood mononuclear cells infected with MTB and normal controls; the sample size of each data set is not less than 6; there must be enough clinical information for analysis.

Based on the above search results, we obtained 4 microarray data sets from the GEO database (GSE51029, GSE52819, GSE54992, and GSE65517). The platform of GSE51029 is GPL4133 Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Feature Number version), the platform of GSE52819 is GPL6244 [HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version], the platform of GSE54992 is GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array, and the platform of GSE65517 is GPL10558 Illumina HumanHT-12 V4.0 expression beadchip. GSE51029 dataset includes 27 samples infected with MTB and 27 normal samples. GSE52819 includes 3 samples infected with MTB and 3 normal samples. GSE54992 includes 9 samples infected with MTB and 6 normal samples. GSE65517 includes 3 samples infected with MTB and 3 normal samples. The microarray data set information is shown in Table 1.

Table 1.

Detailed information about the GEO microarray data set for TB patients.

GEO profile	Source	Case	Control	Platform	Annotation platform
GSE51029	tuberculosis	27	27	GPL4133	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F
GSE52819	tuberculosis	3	3	GPL6244	Affymetrix Human Gene 1.0 ST Array
GSE54992	tuberculosis	9	6	GPL570	Affymetrix Human Genome U133 Plus 2.0 Array
GSE65517	tuberculosis	3	3	GPL10558	Illumina Human HT-12 V4.0 expression beadchip

Identifying Differentially Expressed Genes

The interactive network tool GEO2R is used to analyze the gene expression data of the microarray and find DEGs.[11] In this study, the selected data sets of the MTB infection group and the healthy control group were first analyzed by GEO2R.

Subsequently, the analysis results are downloaded in Microsoft Excel format, and the genes that meet the following conditions: |log FC|>1 and P-values<0.05, were considered DEGs.[12] Finally, we used the FunRich tool (version 3.1.3) to show the intersection of DEGs. In addition, we used the R language tool (version 4.0.3) corresponding to the software R Studio to draw the heat map and correlation circle map of DEGs. The gene expression matrix data used to draw the heat map and correlation circle map was derived from GSE51029.

Function and Pathway Enrichment Analysis of DEGs

We used Metascape online software for GO analysis and KEGG pathway enrichment analysis[13] and further explored the main biological functions of the identified DEGs through functional enrichment analysis based on GO and KEGG databases.[14] The purpose of GO analysis is to identify the characteristic biological features of genes, gene products, and sequences, including biological processes (BP), cell components (CC), and molecular functions (MF).[15] KEGG pathway enrichment analysis provides a complete set of biologically interpreted genome sequence and protein interaction network information.[16] The CC, BP, and MF categories, as well as the KEGG pathway, are classified and presented in the form of bubble charts. These bubble charts are created based on the p-value using the ggplot2 R software package and the statistical software R, where $P<0.05$ is considered statistically significant.

Protein-Protein Interaction (PPI) Network Construction and Module Analysis

The Search Tool for Retrieval of Interacting Genes (STRING) database[17] (http://www.string-db.org/) is an online tool used to identify and predict interactions between genes or proteins. These interactions encompass both physical and functional associations, with data primarily derived from computational predictions, high-throughput experiments, automated text mining, and co-expression networks.[18] We mapped DEGs to the PPI network and set the interaction score>0.4 as a threshold to build a PPI analysis network for DEGs. Analysis of DEGs in the context of protein interactions can help clarify the biochemical complexes or signal transduction components that control biological output,[19] and PPI analysis is very important to explain the underlying molecular mechanisms of key cell activities in pathogenicity.[9]

Next, we used Cytoscape software (version 3.7.2) to visualize the PPI network of DEGs. Each node in the network is a gene, protein, or molecule. The connections between nodes represent the interaction of these biological molecules, which can be used to identify genes that are differentially expressed and pathway relationships between the proteins encoded by DEGs.[8] Subsequently, the modules of the PPI network were screened through the plugin MCODE in Metascape online software and Cytoscape software.[20] The default parameters are as follows: degree cutoff=2, node score cutoff=0.2, k-score=2, and max. Depth=100. Finally, Cytohubba serves as another plugin for Cytoscape.[21] Eleven methods (among which MCC shows satisfactory performance) were used to study the important nodes in the network, and the first 5 genes were selected as the target central genes.

Construction of miRNA-Hub Gene Networks and TF-Hub Gene Networks

NetworkAnalyst is a comprehensive network visualization analysis platform for gene expression analysis. The NetworkAnalyst network tool is used to find TFs of central core genes and construct TF-hub gene regulatory networks.[22] We analyzed MiRNA or TF controlled gene expressions under defined disease conditions through interaction with target genes during the post-transcriptional stage.[23,24] We also applied NetworkAnalyst (https://www.net workanalyst.ca/) to integrate miRNA databases.[25] In the study, the targeted miRNAs of hub genes were defined based on the positive results from at least 3 miRNA-target prediction databases. The targeted TF of hub genes were defined according to the positive results of the ENCODE database (http://cistrome.org/BETA/).[26,27] Finally, we visualized the target miRNA-hub gene and TF-hub gene networks using Cytoscape software.

Confirm Expression by Real-Time PCR

THP-1 is the most important immune cell line to investigate TB defense, antigen presentation, and phagocytosis. This cell line was selected for our experiment.[28] During the stimulation phase of cell culture, the standard strain MTB H37Rv was used to infect macrophages induced and differentiated from human THP-1 monocytes in vitro.[29] Trizol reagent was used to extract total RNA from cells at 24h, 48h, and 72h, respectively.

According to the manufacturer's protocol, we used the PrimeScript RT kit (Takara, Dalian, China) to reverse transcribe total RNA into cDNA and use SYBR Premix Ex Taq II (2×) to perform real-time PCR on the CFX96 PCR detection system (BioRad, California, USA). The configuration system and sample addition operations were performed on ice, and 2 replicate wells were made for detecting the target genes and the internal reference gene of each sample. The reaction volume of 25 μL contains 12.5 μL SYBR® Premix Ex Taq II (Tli RNaseH Plus) (2×), 1 μL forward primer, 1 μL reverse primer, 1 μL cDNA template, and 8.5 μL RNase-free dH_2O.The amplification conditions are as follows: 95°C for 15 s, then 40 cycles, in which the denaturation process at 95°C lasts for 5 s, and the annealing process at 58.5°C lasts for 30 s. SYBR® TB Green Premix PCR mix and primers specific to our target gene were used to amplify the target region effectively. The $2^{-\Delta\Delta}$CT method was used to calculate gene relative expression and perform statistical analysis.[30,31] The primer sequences are shown in Table 2, and the β-actin gene was used as an internal reference gene.

Statistical Analysis

The data are presented as the mean ± standard deviation (SD). GraphPad PRISM version 8.0 (GraphPad Software, Inc., San Diego, CA) was used to perform statistical calculations and prepare graphs. The inter-group comparisons were performed using an unpaired 2-tailed Student's t-test. The probability value of $P<0.05$ was considered statistically significant.

Table 2.

Primer sequence used in Real-Time PCR.

Gene	Forward	Reverse	Purification way
STAT1	ATGCTGGCACCAGAACGAATGAG	TCACCACAACGGGCAGAGAGG	PAGE
TLR7	ACCAACTGACCACTGTCCCTGAG	TCGCAACTGGAAGGCATCTTGTAG	PAGE
CCL20	AACAGCACTCCCAAAGAACTGG	GCAGAGGTGGAGTAGCAGCA	PAGE
CCR2	CCAACGAGAGCGGTGAAGAAGTC	CGAGTAGAGCGGAGGCAGGAG	PAGE
CXCL8	ACTTTCAGAGACAGCAGAGCACAC	CACACAGTGAGATGGTTCCTTCCG	PAGE
β-actin	TGGCATCCACGAAACTACCT	CAATGCCAGGGTACATGGTG	PAGE

Results

Microarray Data Information and Identification of Candidate DEGs

The 4 microarray expression datasets, GSE51029, GSE52819, GSE54992, and GSE65517, were obtained from the GEO database. The above data sets were uploaded to GEO2R and standardized (Figure 1) to screen DEGs between the TB patient group and the normal control group and create a volcano map of the distribution of these DEGs in 4 data sets. The differential expression of multiple genes in the 2 sets of sample data in each array is shown in Figure 2.

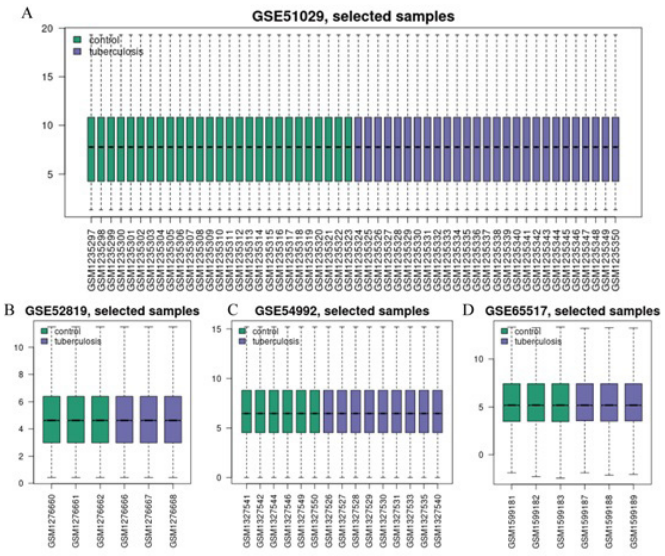

Figure 1. *Standardization of gene expression.*

(A) The standardization of GSE51029 data, (B) the standardization of GSE52819 data, (C) the standardization of GSE54992 data, and (D) the standardization of GSE65517 data. The green bars represent the normalized data for the healthy control group, and the blue bars represent those for the TB infection group.

In the GSE51029 data set, 5182 DEGs were obtained, including 1024 up-regulated genes and 4156 down-regulated genes. In the GSE52819 data set, 1959 DEGs were obtained, including 1044 up-regulated genes and 915 down-regulated genes. In the GSE54992 data set, 7461 DEGs were obtained, including 3918 up-regulated genes and 3543 down-regulated genes. In the GSE65517 data set, 2604 DEGs were obtained, including 1337 up-regulated genes and 1267 down-regulated genes.

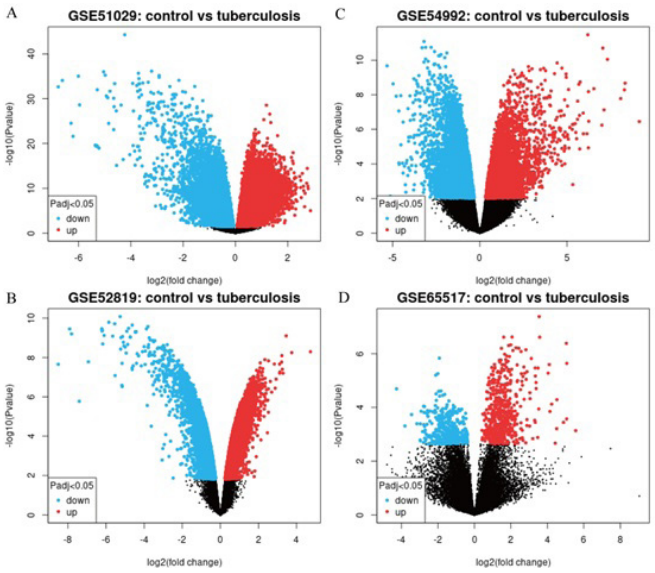

Figure 2. *Differential expression trends of data between two sets of samples.*

(A) GSE51029 data, (B) GSE52819 data, (C) GSE54992 data, and (D) GSE65517 data. The red points represent up-regulated genes screened based on \|fold change\| >1.0 and a corrected P-value of <0.05. The green points represent downregulation of gene expression, screened based on \|fold change\| >1.0 and a corrected P-value of <0.05. The black points represent genes with no significant difference.

According to the criteria of |log FC|> 1 and *P*-values <0.05, we used FunRich software to display the intersection of DEGs in the 4 microarray expression data, a total of 46 overlapping genes were found (Figure 3A), which were regarded as candidate DEGs and used for further analysis. We used the pheatmap software package in R Studio to visualize the heat map of 46 DEGs (Figore 3B). It is worth mentioning that by using the corrplot and circlize software packages in R Studio to visualize 46 DEGs, the generated correlation circle graph can show the correlation of multiple genes in one picture (Figure 3C). The outer circle represents genes, and the line between the 2 genes represents the correlation coefficient. Positive correlation is shown in red, and negative correlation is shown in green. The darker the color, the more significant the correlation.

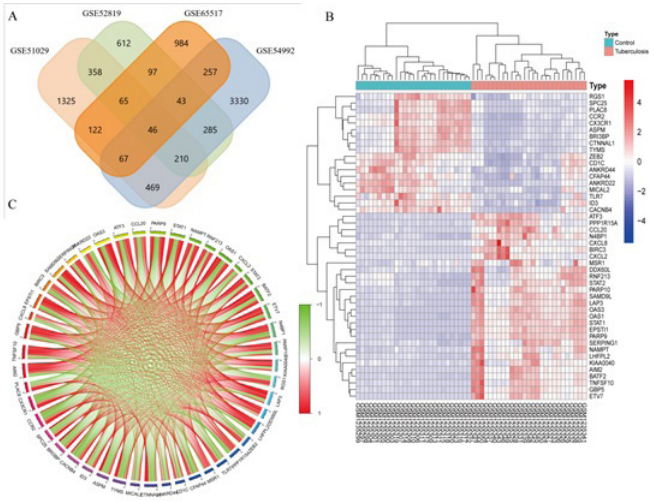

Figure 3. *Selection of candidate DEGs, hierarchical clustering heatmap, and corCircos.*
(A) Venn diagram for overlapping differentially expressed genes (DEGs) based on the four datasets, namely, GSE51029, GSE52819, GSE54992, and GSE65517. (B) The hierarchical clustering heatmap data derived from GSE51029: red indicates relatively up-regulated gene expression, and blue indicates relatively down-regulated gene expression. (C) Circos shows the relationships among multiple genes in a single picture. The positive correlation is shown in red, and the negative correlation is shown in green.

Enrichment Analysis of GO and KEGG Pathways

Enrichment analysis is the core of most existing gene annotation portals.[32] In the enrichment analysis process, the input gene list is compared with thousands of gene sets, which are defined by their participation in specific BP, protein localization, enzyme functions, pathway members, or other characteristics. We used RSQLite, org.Hs.eg.db, clusterProiler and other software packages in R software to analyze and visualize the GO function and KEGG pathway enrichment of 46 DEGs. The results are shown in Figure 4.

The top 20 most important items of enrichment analysis are illustrated in the form of bubble diagrams. The size and color of the bubbles indicate the number of DEGs and the importance of enrichment in the enrichment analysis of GO and KEGG pathways, respectively. The first 5 important terms of GO enrichment analysis indicate that in the BP category, DEGs participate in the defense response to the virus, the regulation of the immune response process, the cellular response to interferon-γ, the regulation of the innate immune response, and the mediation of the chemokine signaling pathway (Fig. 4A). For the MF category, DEGs are related to cytokine receptors, cytokine activity, G protein-coupled receptor binding, receptor ligand activity, and signal receptor activator activity (Fig. 4B). For the CC category, DEGs are significantly enriched in membrane rafts, mitochondrial outer membrane, cytoplasmic vesicle cavity, endocytic vesicles, and nuclear chromatin (Fig. 4C).

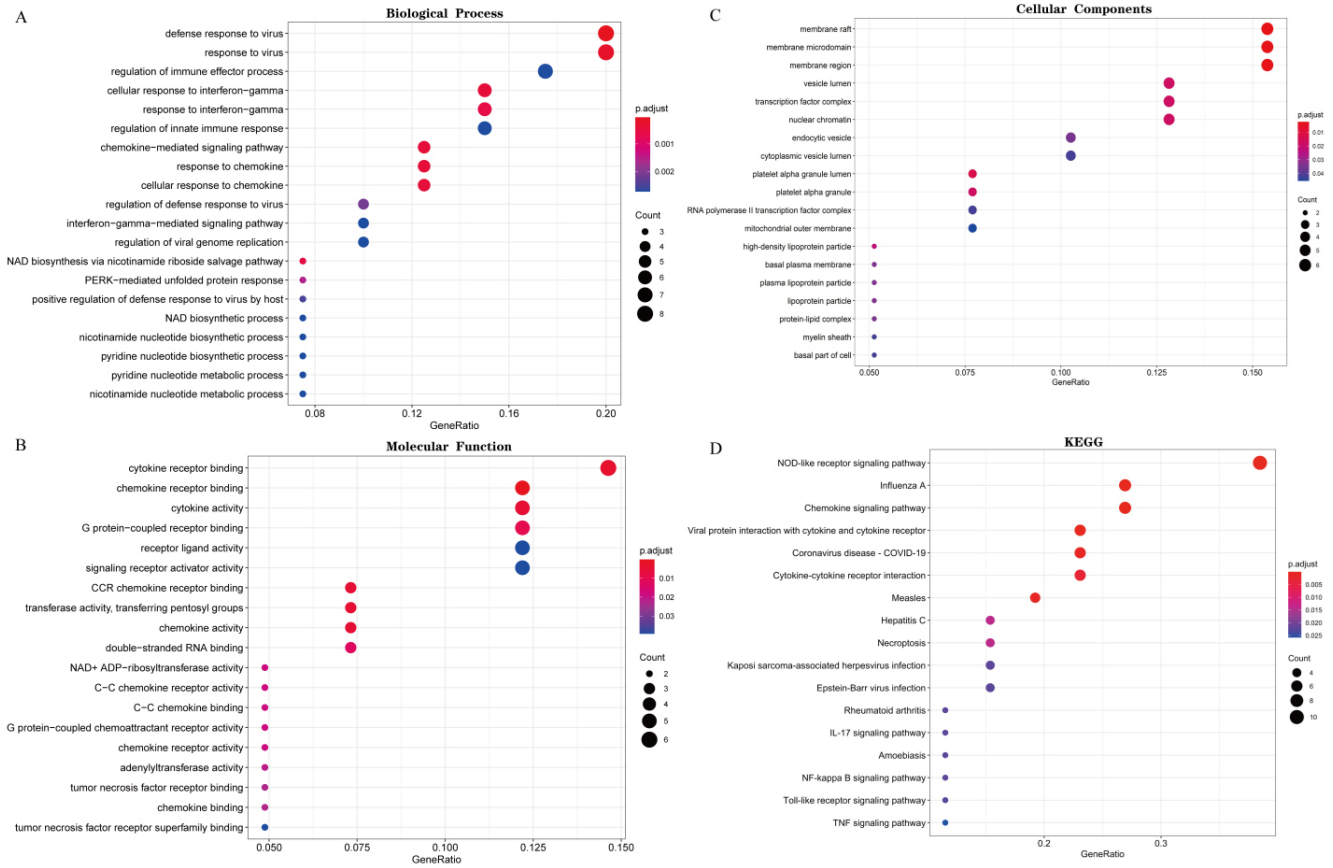

Figure 4. *Bubble map for GO and KEGG pathway analyses of DEGs.*
(A) Top 20 biological process (BP) terms, (B) Top 20 molecular functions (MF) terms, (C) Top 20 cellular components (CC) terms, and (D) Top 20 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The horizontal axis represents the significance of enrichment (denoted by Gene Ratio; the larger the value, the more significant the enrichment), and the vertical axis represents the GO terms enriched. The dot size and depth indicate the number of DEGs contained in the GO terms and the degree of the rich factor enrichment, respectively. We selected the top 20 GO terms according to Gene Ratio.

For KEGG pathway enrichment analysis, the first several important KEGG pathways of DEGs include the NOD-like receptor signaling pathway, influenza A, chemokine signaling pathway, COVID-19, Toll-like receptor signaling pathway, and NF-kappa B signaling pathway. To illustrate the position and significance of DEGs in the pathway in detail, based on KEGG enrichment analysis, the pathview software package in R Studio was utilized to visualize the 46 DEGs identified. The pathway diagrams are stored as the attached files.

PPI Network Construction Analysis and Host Hub Gene Selection

PPI network analysis has been regarded as a useful tool for exploring biological responses in health and disease.[33] According to the STRING database and Cytoscape software, 46 DEGs are located in the PPI network complex, including 46 nodes and 83 edges. The PPI enrichment p-value is less than 1.0e-16, and the confidence score is greater than 0.4. Finally, Cytoscape is used to visualize the PPI network of these DEGs [34] (Figure 5). After deleting isolated and partially connected nodes, a complex PPI network was successfully constructed (Figure 5A).

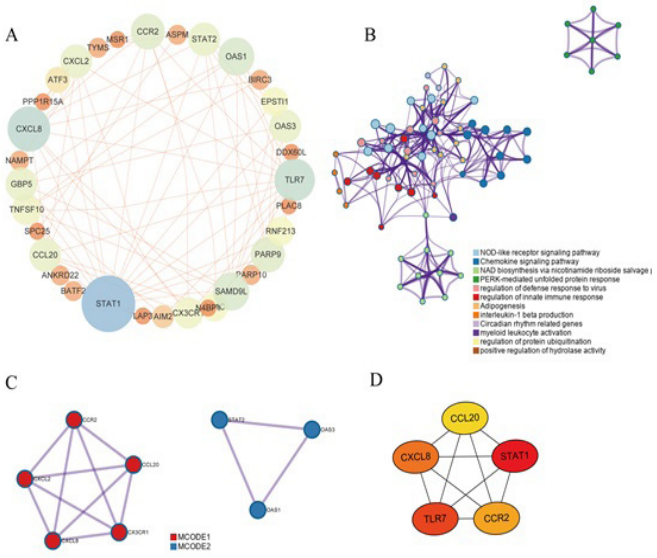

Figure 5. *Protein-protein interaction (PPI) network of differentially expressed genes and identification of hub genes.*

(A) PPI network of differentially expressed genes (DEGs). (B) Network of enriched terms of differentially expressed proteins colored by cluster ID. Each node represents an enriched term and is colored by its cluster ID. Nodes that share the same cluster ID are typically close to each other. (C) MCODE components identified in the PPI interaction network. Each node represents a gene, and different colors represent different MCODE components. (D) Subnetwork of the top five hub genes from the PPI network. Node color reflects the degree of connectivity (Red color represents a higher degree, and yellow color represents a lower degree).

The online software Metascape applies the mature complex recognition algorithm MCODE, which can automatically extract protein complexes embedded in large-scale networks.[35] It is worth mentioning that Metascape has obvious advantages in PPI network module analysis and GO enrichment analysis[36] and can directly obtain relevant data. We use Cytoscape to visualize the module analysis results generated by Metascape. The MCODE plugin of the latter filters the module network within the PPI network again (Figure 5B-C), and the options are set as default parameters. Host hub genes are a series of genes that play a vital role in diversified BP, and other genes are usually regulated by these hub genes.[37] As another plugin in Cytoscape, Cytohubba utilizes maximal clique centrality (MCC) to identify key nodes in the network and select the most relevant genes as target hub genes.[38] In this study, the top 5 hub genes with the highest degree of interaction were identified (Figure 5D), including STAT1 (Signal transducer and activator of transcription 1-alpha/beta), TLR7 (Toll-like receptor 7), CXCL8 (C-X-C motif chemokine ligand 8), CCR2 (C-C chemokine receptor type 2), CCL20 (C-C motif chemokine 20).

Construction and Analysis of the miRNA-Hub Gene Network and TF-Hub Gene Network

NetworkAnalyst was used to screen targeted miRNA and TF for hub genes.[39] For these 5 identified hub genes, the top 3 miRNA-targeted DEGs are *TLR7*, regulated by 39 miRNAs, *CXCL8*, regulated by 36 miRNAs, and *STAT1*, regulated by 19 miRNAs. The miRNA that controls the largest number of hub genes (3 genes) is hsa-mir-335-5p (Figure 6A), and other important miRNAs are shown in Table 3. Unfortunately, when the selected 5 core genes were imported into the ENCODE of the TF database, only 2 of them could be analyzed: *CCL20*, regulated by 39 TFs, and *CCR2*, regulated by 2 TFs (Figure 6B).

Table 3.

MiRNAs and its target genes.

miRNA	Hub genes targeted by miRNA	Gene count	Betweenness Centrality
hsa-mir-335-5p	*CXCL8, CCR2, CCL20*	3	0.21933622
hsa-mir-93-5p	*CXCL8, TLR7*	2	0.0620336
hsa-mir-106a-5p	*CXCL8, TLR7*	2	0.0620336
hsa-mir-203a-3p	*CXCL8, STAT1*	2	0.05025596
hsa-mir-146a-5p	*CXCL8, STAT1*	2	0.05025596
hsa-mir-150-5p	*TLR7, STAT1*	2	0.11976912
hsa-mir-155-5p	*CXCL8, STAT1*	2	0.05025596
hsa-mir-302c-3p	*CXCL8, TLR7*	2	0.0620336
hsa-mir-302d-3p	*CXCL8, TLR7*	2	0.0620336
hsa-mir-520b	*CXCL8, TLR7*	2	0.0620336

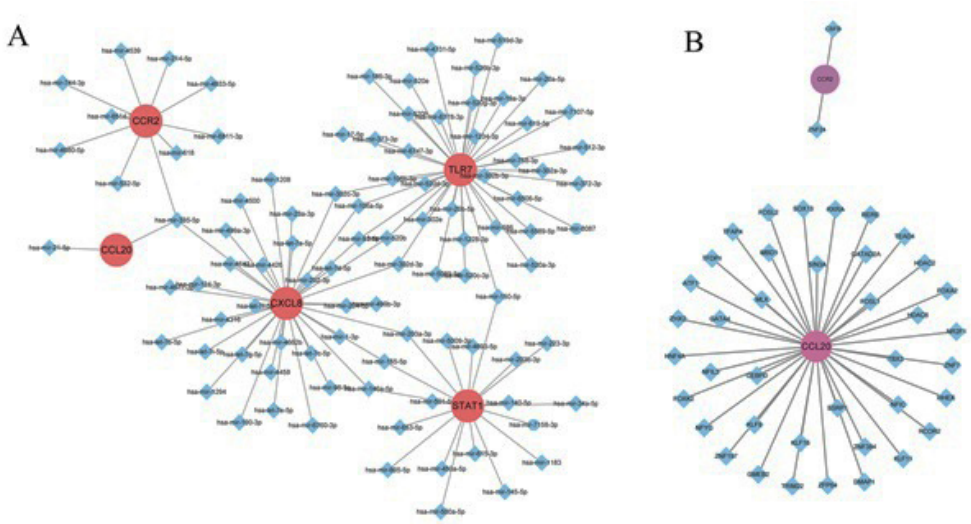

Figure 6. *Regulatory networks of miRNA-hub gene and TF-hub gene.*

The networks of (A) miRNA-hub gene and (B) TF-hub gene. The red or purple circle nodes are the hub genes, and the blue diamond nodes are the miRNAs and the TF.

Verifying the Expression Levels of Hub Genes by Real-Time PCR

To verify the accuracy of the prediction results, this study utilized real-time PCR to detect the mRNA expression levels of the 5 host hub core genes at 24-, 48-, and 72 hours after MTB infection of the cells. At least 3 biological replicates were performed for each experiment. The results showed that the expressions of core genes are generally consistent with the heat map results derived from the gene expression matrix (GSE51029_series_matrix) data. Specifically, compared with the PBS group, the mRNA expression levels of *STAT1*, *CCL20*, and *CXCL8* in the MTB group were all up-regulated at 3 time points (Figure 7A-C), while the mRNA expression levels of *TLR7* and *CCR2* at all 3 time points showed down-regulation (Figure 7D-E).

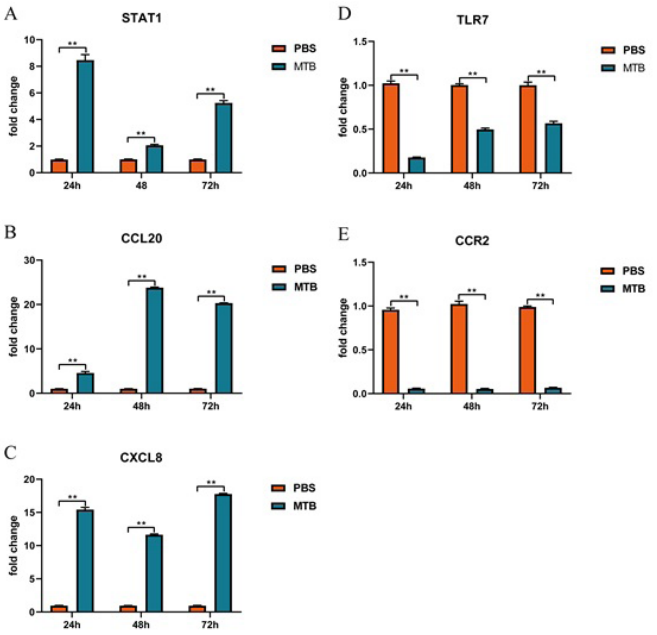

Figure 7. *The expression levels of relevant gene mRNA analyzed by Real-Time PCR, comparing the MTB infection group with the control group.*

*(A-C) STAT1, CCL20, and CXCL8 were demonstrated to be upregulated in the MTB group compared to the PBS group at 24, 48, 72 hours. (D-E) TLR7 and CCR2 were shown to be downregulated in the MTB group compared to the PBS group at 24, 48, and 72 hours. Data represent mean±SD (n=3). Asterisks denote: **P<0.01indicates significance. The yellow color represents the MTB group, and the blue color represents the PBS group.*

Discussion

TB is an infectious disease that seriously harms human health. It is caused by MTB, a bacterium that is parasitic in macrophages.[40] As an intracellular pathogen, MTB largely depends on its ability to compromise the host's innate immune defense system, specifically the macrophage.[41] As the human body's first line of immune defense, macrophages can kill MTB through phagocytosis, oxidative stress, acidification, and antigen presentation.[42] Therefore, the bioinformatics research on macrophages stimulated by MTB is particularly important.

The method of bioinformatics helps analyze the expression of key genes to reveal the potential molecular mechanisms underlying the biological behavior of TB, providing novel insights for elucidating the pathogenesis of the disease. Microarray technology enables us to investigate host genetic changes and gene expression associated with TB and has proven to be a valuable method for identifying new biomarkers in other diseases.[43] In this study, 4 GEO microarray data sets (GSE51029, GSE52819, GSE54992, and GSE65517) were integrated to identify DEGs between PBMC of TB patients and healthy people to offset the false positive rate in the analysis of independent data sets. According to the criteria of |log FC|> 1 and *P*-values <0.05,

this study used FunRich software to display the intersection of DEGs in the four-microarray expression data and found a total of 46 overlapping DEGs that can be considered as candidates. The GO function and KEGG pathway enrichment analysis, PPI network analysis, and core gene selection were carried out successively, followed by the construction of target gene-miRNA networks and target gene-TF networks, and the verification of core gene expression at the mRNA level.

According to the results of GO enrichment analysis, in the BP category, DEGs are primarily involved in regulating immune effect processes, IFN-γ-mediated signaling pathways, nicotinamide nucleotide biosynthesis processes, NAD biosynthesis processes, and chemokine-mediated signaling pathways. Studies have shown that in the case of initial infection, macrophages will encounter MTB before being stimulated by the Th1 cytokine IFN-γ.[44] However, only after stimulation with IFN-γ can the antibacterial ability and antigen presentation function of macrophages be fully activated.[45] For the MF category, DEGs are associated with cytokine/chemokine receptor binding, cytokine/chemokine activity, G protein-coupled receptor binding, receptor ligand activity, and signal receptor activator activity. MTB induces host pro-inflammatory mediators that play a crucial role in disease control;[46] among these, chemokines are small molecular-weight proteins involved in immune regulation and inflammation.[40] For the CC category, DEGs are significantly enriched in membrane rafts, mitochondrial outer membrane, cytoplasmic vesicle cavity, endocytic vesicles, and nuclear chromatin. KEGG enrichment analysis showed that the NOD-like receptor signaling pathway and the Toll-like receptor signaling pathway are 2 important pathways associated with TB. Innate immune cells utilize various pattern recognition receptors, including Toll-like receptors, C-type lectin receptors, and NOD-like receptors, to respond to pathogen components and perform a range of biological functions.[47,48] Studies have shown that experimental models of TB have highlighted the importance of TLRs in preventing MTB.[49] Additionally, antigen recognition by NOD-2 (nucleotide-binding oligomerization domain 2), a member of the NOD-like receptor family, is also crucial in conferring immunity against bacteria and viruses.[50,51] This suggests that the coordinated triggering of TLRs and NOD-2 may lead to a stronger and lasting immune response, thereby limiting the growth of MTB.[52]

By constructing a PPI network and analyzing it with MCODE and Cytohubba in Cytoscape, 5 core genes were identified. Studies have reported that in the early stage of TB infection, *STAT1* can promote the activation of downstream apoptotic factors through phosphorylation,[53] thereby activating transcription. At the same time, *STAT1* is also crucial in promoting the polarization of macrophages into the M1 phenotype. Polarized M1 macrophages can eliminate MTB infection more effectively than polarized M2 macrophages.[54] It is reported that after ssRNA upregulates TLR7, the number of MTB in macrophages is significantly reduced, and the macrophage viability is significantly increased, indicating that TLR7 can effectively inhibit the growth of MTB and improve the viability of macrophages.[42] Kane et al.[55] found that fibroblasts, through their CXCL8-dependent contribution to cell recruitment and mycobacterial killing in granulomas, have a previously unrecognized role in regulating TB inflammation. In the study by Dunlap et al.[56] the mouse model provided evidence that the *CCR2* axis is essential for protective immunity against the emerging MTB lineage infection. Another report showed that *CCL20* is overexpressed in monocytes infected by MTB and inhibits the production of reactive oxygen species (ROS).[57]

To study the molecular mechanisms of potential core gene disorders, it is necessary to identify potential miRNAs using bioinformatics methods. The miRNA is an endogenous, non-coding RNA molecule with a length of 18-22 nucleotides that targets the 3' untranslated region of a gene. It can regulate gene expression at the post-transcriptional level by degrading or inhibiting the translation of target genes.[58] miRNAs are known to regulate gene expression post-transcriptionally by either inhibiting protein synthesis or targeting mRNA for degradation.[59] Increasing evidence suggests that miRNAs are closely linked to the development and progression of cancer and other major diseases. In this study, the top 3 targeted DEGs in the miRNA gene network were *CXCL8*, *TLR7*, and *STAT1*. At the same time, we observed 10 miRNAs and found that their targeting involves at least 2 core genes. Among them, the miRNA that controls the largest number of core genes is hsa-mir-335-5p; however, there are currently few studies related to it. One study found that the gain and loss of hsa-miR-335-3p function can lead to changes in the expression of GATA4 and NKX2-5, 2 cardiac differentiation markers, during the cardiac differentiation of human embryonic stem cells.[60] In addition, hsa-miR-335-3p has been identified as an upstream regulator of 2 modules related to the recurrence of osteosarcoma patients.[61] The results of bioinformatics analysis found that hsa-mir-335-5p has high potential as a new biomarker. This study also established a hot TF-hub gene regulatory network to further explore the molecular mechanism of TB.[62] TF is the primary regulator of gene expression and is associated with the pathogenesis of TB. In our study, we also found some TFs that interact closely with hub genes (*MLX*, *TFDP1*, *RXRA*, *ZNF197*, *GMEB2*, *TRIM22*). The complex interaction between TF and hub genes has made a huge contribution to the development of the disease. Our study analyzed the miRNA/TF interactions of hub genes, providing a possibility to predict potential therapeutic agents and explore the potential mechanisms of TB. Therefore, our results lay the foundation for further research.

Conclusion

In summary, our study identified DEGs through bioinformatics analysis to discover potential biomarkers that may be related to the progression of TB. The study provides a set of candidate DEGs and 5 important hub genes. These results can provide a basis for screening candidate drugs and biomarkers for the diagnosis of TB and can be used for future research on the molecular mechanism. However, further studies, including cell and animal experiments, are

needed to elucidate the functions of DEGs in physiological and pathological processes. In the future, in-depth exploration of molecular mechanisms of new hub genes is required in TB, and related experimental models can be constructed based on these genes for early detection, prognosis judgment, risk assessment, and targeted TB therapy.

Sources of Funding

The research project was funded by grants from the National Natural Science Foundation of China (Nos. 82560396 and 82160304).

Competing Interests

The authors declare that they have no financial/commercial conflicts of interest concerning this article.

References

1. WHO. Global tuberculosis report 2020. Geneva, 2020. https://www.who.int/publications/i/item/9789240013131
2. Fan S, Zhou G, Shang P, et al. Clinical Study of 660 Cases of Pulmonary Tuberculosis. Harbin Medical Journal, 2014, 34(1):1-11.
3. Siddiqi K, Lambert ML, Walley J. Clinical diagnosis of smear-negative pulmonary tuberculosis in low-income countries: the current evidence. Lancet Infect Dis. 2003 May;3(5):288-96. doi: 10.1016/s1473-3099(03)00609-1.
4. Won EJ, Choi JH, Cho YN, Jin HM, Kee HJ, Park YW, et al. Biomarkers for discrimination between latent tuberculosis infection and active tuberculosis disease. J Infect. 2017 Mar;74(3):281-293. doi: 10.1016/j.jinf.2016.11.010. Epub 2016 Nov 19. PMID: 27871809.
5. Sharma SK, Vashishtha R, Chauhan LS, Sreenivas V, Seth D. Comparison of TST and IGRA in Diagnosis of Latent Tuberculosis Infection in a High TB-Burden Setting. PLoS One. 2017 Jan 6;12(1):e0169539. doi: 10.1371/journal.pone.0169539. PMID: 28060926; PMCID: PMC5218498.
6. Salem H, Attiya G, El-Fishawy N. Classification of human cancer diseases by gene expression profiles. Applied Soft Computing, 2017, 50:124-34.
7. Ramaswamyreddy SH, Smitha T. Microarray-based gene expression profiling for early detection of oral squamous cell carcinoma. J Oral Maxillofac Pathol. 2018 Sep-Dec;22(3):293-295. doi: 10.4103/jomfp.JOMFP_270_18. PMID: 30651668; PMCID: PMC6306598.
8. Yang X, Zhu S, Li L, Zhang L, Xian S, Wang Y, Cheng Y. Identification of differentially expressed genes and signaling pathways in ovarian cancer by integrated bioinformatics analysis. Onco Targets Ther. 2018 Mar 15;11:1457-1474. doi: 10.2147/OTT.S152238. PMID: 29588600; PMCID: PMC5858852.
9. Xie L, Chao X, Teng T, Li Q, Xie J. Identification of Potential Biomarkers and Related Transcription Factors in Peripheral Blood of Tuberculosis Patients. Int J Environ Res Public Health. 2020 Sep 24;17(19):6993. doi: 10.3390/ijerph17196993. PMID: 32987825; PMCID: PMC7579196.
10. Qin XB, Zhang WJ, Zou L, Huang PJ, Sun BJ. Identification potential biomarkers in pulmonary tuberculosis and latent infection based on bioinformatics analysis. BMC Infect Dis. 2016 Sep 21;16(1):500. doi: 10.1186/s12879-016-1822-6. PMID: 27655333; PMCID: PMC5031349.
11. Dumas J, Gargano M, Dancik GM. An online tool for biomarker analysis in Gene Expression Omnibus (GEO) datasets. Bioinform. 2016: 5292.
12. Xu Z, Zhou Y, Cao Y, Dinh TL, Wan J, Zhao M. Identification of candidate biomarkers and analysis of prognostic values in ovarian cancer by integrated bioinformatics analysis. Med Oncol. 2016 Nov;33(11):130. doi: 10.1007/s12032-016-0840-y. Epub 2016 Oct 18. PMID: 27757782.
13. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019 Apr 3;10(1):1523. doi: 10.1038/s41467-019-09234-6. PMID: 30944313; PMCID: PMC6447622.
14. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 2016 Jan 4;44(D1):D457-62. doi: 10.1093/nar/gkv1070. Epub 2015 Oct 17. PMID: 26476454; PMCID: PMC4702792.
15. Gene Ontology Consortium. The Gene Ontology (GO) project in 2006. Nucleic Acids Res. 2006 Jan 1;34(Database issue):D322-6. doi: 10.1093/nar/gkj021. PMID: 16381878; PMCID: PMC1347384.
16. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000 Jan 1;28(1):27-30. doi: 10.1093/nar/28.1.27. PMID: 10592173; PMCID: PMC102409.
17. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 2013 Jan;41(Database issue):D808-15. doi: 10.1093/nar/gks1094. Epub 2012 Nov 29. PMID: 23203871; PMCID: PMC3531103.
18. Wang H, Zhu H, Zhu W, Xu Y, Wang N, Han B, Song H, Qiao J. Bioinformatic Analysis Identifies Potential Key Genes in the Pathogenesis of Turner Syndrome. Front Endocrinol (Lausanne). 2020 Mar 6;11:104. doi: 10.3389/fendo.2020.00104. PMID: 32210915; PMCID: PMC7069359.
19. Pizzuti C, Rombo SE. Algorithms and tools for protein-protein interaction networks clustering, with a special focus on population-based stochastic methods. Bioinformatics. 2014 May 15;30(10):1343-52. doi: 10.1093/bioinformatics/btu034. Epub 2014 Jan 22. PMID: 24458952.
20. Bandettini WP, Kellman P, Mancini C, Booker OJ, Vasu S, Leung SW, et al. MultiContrast Delayed Enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study. J Cardiovasc Magn Reson. 2012 Nov 30;14(1):83. doi: 10.1186/1532-429X-14-83. PMID: 23199362; PMCID: PMC3552709.
21. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014;8 Suppl 4(Suppl 4):S11. doi: 10.1186/1752-0509-8-S4-S11. Epub 2014 Dec 8. PMID: 25521941; PMCID: PMC4290687.
22. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. Nucleic Acids Res. 2019 Jul 2;47(W1):W234-W241.

doi: 10.1093/nar/gkz240. PMID: 30931480; PMCID: PMC6602507.
23. Soifer HS, Rossi JJ, Saetrom P. MicroRNAs in disease and potential therapeutic applications. Mol Ther. 2007 Dec;15(12):2070-9. doi: 10.1038/sj.mt.6300311. Epub 2007 Sep 18. PMID: 17878899.
24. Baldwin AS Jr. Series introduction: the transcription factor NF-kappaB and human disease. J Clin Invest. 2001 Jan;107(1):3-6. doi: 10.1172/JCI11891. PMID: 11134170; PMCID: PMC198555.
25. Xia J, Gill EE, Hancock RE. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. Nat Protoc. 2015 Jun;10(6):823-44. doi: 10.1038/nprot.2015.052. Epub 2015 May 7. PMID: 25950236.
26. Yang D, He Y, Wu B, Deng Y, Wang N, Li M, Liu Y. Integrated bioinformatics analysis for the screening of hub genes and therapeutic drugs in ovarian cancer. J Ovarian Res. 2020 Jan 27;13(1):10. doi: 10.1186/s13048-020-0613-2. PMID: 31987036; PMCID: PMC6986075.
27. Yang W, Zhao X, Han Y, Duan L, Lu X, Wang X, et al. Identification of hub genes and therapeutic drugs in esophageal squamous cell carcinoma based on integrated bioinformatics strategy. Cancer Cell Int. 2019 May 22;19:142. doi: 10.1186/s12935-019-0854-6. PMID: 31139019; PMCID: PMC6530124.
28. Zhang YW, Lin Y, Yu HY, Tian RN, Li F. Characteristic genes in THP1 derived macrophages infected with Mycobacterium tuberculosis H37Rv strain identified by integrating bioinformatics methods. Int J Mol Med. 2019 Oct;44(4):1243-1254. doi: 10.3892/ijmm.2019.4293. Epub 2019 Jul 30. PMID: 31364746; PMCID: PMC6713430.
29. Feng Z, Bai X, Wang T, Garcia C, Bai A, Li L, et al. Differential Responses by Human Macrophages to Infection With *Mycobacterium tuberculosis* and Non-tuberculous Mycobacteria. Front Microbiol. 2020 Feb 7;11:116. doi: 10.3389/fmicb.2020.00116. PMID: 32117140; PMCID: PMC7018682.
30. Ding Z, Sun L, Bi Y, Zhang Y, Yue P, Xu X, et al. Integrative Transcriptome and Proteome Analyses Provide New Insights Into the Interaction Between Live Borrelia burgdorferi and Frontal Cortex Explants of the Rhesus Brain. J Neuropathol Exp Neurol. 2020 May 1;79(5):518-529. doi: 10.1093/jnen/nlaa015. PMID: 32196082.
31. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001 Dec;25(4):402-8. doi: 10.1006/meth.2001.1262. PMID: 11846609.
32. Li H, Long J, Xie F, Kang K, Shi Y, Xu W, et al. Transcriptomic analysis and identification of prognostic biomarkers in cholangiocarcinoma. Oncol Rep. 2019 Nov;42(5):1833-1842. doi: 10.3892/or.2019.7318. Epub 2019 Sep 17. PMID: 31545466; PMCID: PMC6787946.
33. Vella D, Marini S, Vitali F, Di Silvestre D, Mauri G, Bellazzi R. MTGO: PPI Network Analysis Via Topological and Functional Module Identification. Sci Rep. 2018 Apr 3;8(1):5499. doi: 10.1038/s41598-018-23672-0. PMID: 29615773; PMCID: PMC5882952.
34. Feng H, Gu ZY, Li Q, Liu QH, Yang XY, Zhang JJ. Identification of significant genes with poor prognosis in ovarian cancer via bioinformatical analysis. J Ovarian Res. 2019 Apr 22;12(1):35. doi: 10.1186/s13048-019-0508-2.
35. Liang J, Wu M, Bai C, Ma C, Fang P, Hou W, et al. Network Pharmacology Approach to Explore the Potential Mechanisms of Jieduan-Niwan Formula Treating Acute-on-Chronic Liver Failure. Evid Based Complement Alternat Med. 2020 Dec 30;2020:1041307. doi: 10.1155/2020/1041307. PMID: 33456481; PMCID: PMC7787753.
36. Li W, Wang S, Qiu C, Liu Z, Zhou Q, Kong D, Ma X, Jiang J. Comprehensive bioinformatics analysis of acquired progesterone resistance in endometrial cancer cell line. J Transl Med. 2019 Feb 27;17(1):58. doi: 10.1186/s12967-019-1814-6. PMID: 30813939; PMCID: PMC6391799.
37. Zhang YM, Meng LB, Yu SJ, Ma DX. Identification of potential crucial genes in monocytes for atherosclerosis using bioinformatics analysis. J Int Med Res. 2020 Apr;48(4):300060520909277. doi: 10.1177/0300060520909277. PMID: 32314637; PMCID: PMC7175059.
38. Guo C, Li Z. Bioinformatics Analysis of Key Genes and Pathways Associated with Thrombosis in Essential Thrombocythemia. Med Sci Monit. 2019 Dec 5;25:9262-9271. doi: 10.12659/MSM.918719. PMID: 31801935; PMCID: PMC6911306..
39. Zhou R, Liu D, Zhu J, Zhang T. Common gene signatures and key pathways in hypopharyngeal and esophageal squamous cell carcinoma: Evidence from bioinformatic analysis. Medicine (Baltimore). 2020 Oct 16;99(42):e22434. doi: 10.1097/MD.0000000000022434. PMID: 33080677; PMCID: PMC7571924.
40. Lyon SM, Rossman MD. Pulmonary tuberculosis. Tuberculosis and Nontuberculous Mycobacterial Infections. 2017,5(1):283-98.
41. Kumar M, Sahu SK, Kumar R, Subuddhi A, Maji RK, Jana K, et al.. MicroRNA let-7 modulates the immune response to ycobacterium tuberculosis infection via control of A20, an inhibitor of the NF-κB pathway. Cell Host Microbe. 2015 Mar 11;17(3):345-356. doi: 10.1016/j.chom.2015.01.007. Epub 2015 Feb 12. PMID: 25683052.
42. Bao M, Yi Z, Fu Y. Activation of TLR7 Inhibition of Mycobacterium Tuberculosis Survival by Autophagy in RAW 264.7 Macrophages. J Cell Biochem. 2017 Dec;118(12):4222-4229. doi: 10.1002/jcb.26072. Epub 2017 May 23. PMID: 28419514.
43. Li L, Lei Q, Zhang S, Kong L, Qin B. Screening and identification of key biomarkers in hepatocellular carcinoma: Evidence from bioinformatic analysis. Oncol Rep. 2017 Nov;38(5):2607-2618. doi: 10.3892/or.2017.5946. Epub 2017 Sep 7. PMID: 28901457; PMCID: PMC5780015.
44. Brzezinska M, Szulc I, Brzostek A, Klink M, Kielbik M, Sulowska Z, et al. The role of 3-ketosteroid 1(2)-dehydrogenase in the pathogenicity of Mycobacterium tuberculosis. BMC Microbiol. 2013 Feb 20;13:43. doi: 10.1186/1471-2180-13-43. PMID: 23425360; PMCID: PMC3599626.
45. Raja A. Immunology of tuberculosis. Indian J Med Res. 2004 Oct;120(4):213-32. PMID: 15520479.
46. Ansari AW, Kamarulzaman A, Schmidt RE. Multifaceted Impact of Host C-C Chemokine CCL2 in the Immuno-Pathogenesis of HIV-1/M. tuberculosis Co-Infection. Front Immunol. 2013 Oct 4;4:312. doi: 10.3389/fimmu.2013.00312. PMID: 24109479; PMCID: PMC3790230.
47. Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. Cell. 2006 Feb 24;124(4):783-801. doi:

10.1016/j.cell.2006.02.015. PMID: 16497588.
48. Akira S, Takeda K, Kaisho T. Toll-like receptors: critical proteins linking innate and acquired immunity. Nat Immunol. 2001 Aug;2(8):675-80. doi: 10.1038/90609. PMID: 11477402.
49. Fremond CM, Yeremeev V, Nicolle DM, Jacobs M, Quesniaux VF, Ryffel B. Fatal Mycobacterium tuberculosis infection despite adaptive immune response in the absence of MyD88. J Clin Invest. 2004 Dec;114(12):1790-9. doi: 10.1172/JCI21027. PMID: 15599404; PMCID: PMC535064.
50. PANDEY A K, YANG Y, JIANG Z, et al. NOD2, RIP2 and IRF5 play a critical role in the type I interferon response to *Mycobacterium tuberculosis* [J]. Public Library of Science Pathogens, 2009, 5(7):e1000500.
51. Lupfer C, Thomas PG, Kanneganti TD. Nucleotide oligomerization and binding domain 2-dependent dendritic cell activation is necessary for innate immunity and optimal CD8+ T Cell responses to influenza A virus infection. J Virol. 2014 Aug;88(16):8946-55. doi: 10.1128/JVI.01110-14. Epub 2014 May 28. PMID: 24872587; PMCID: PMC4136245.
52. Khan N, Pahari S, Vidyarthi A, Aqdas M, Agrewala JN. NOD-2 and TLR-4 Signaling Reinforces the Efficacy of Dendritic Cells and Reduces the Dose of TB Drugs against Mycobacterium tuberculosis. J Innate Immun. 2016;8(3):228-42. doi: 10.1159/000439591. Epub 2015 Nov 28. PMID: 26613532; PMCID: PMC6738777.
53. Yao K, Chen Q, Wu Y, Liu F, Chen X, Zhang Y. Unphosphorylated STAT1 represses apoptosis in macrophages during *Mycobacteriumtuberculosis* infection. J Cell Sci. 2017 May 15;130(10):1740-1751. doi: 10.1242/jcs.200659. Epub 2017 Mar 27. PMID: 28348106.
54. Lim YJ, Yi MH, Choi JA, Lee J, Han JY, Jo SH, et al. Roles of endoplasmic reticulum stress-mediated apoptosis in M1-polarized macrophages during mycobacterial infections. Sci Rep. 2016 Nov 15;6:37211. doi: 10.1038/srep37211. PMID: 27845414; PMCID: PMC5109032.
55. O'Kane CM, Boyle JJ, Horncastle DE, Elkington PT, Friedland JS. Monocyte-dependent fibroblast CXCL8 secretion occurs in tuberculosis and limits survival of mycobacteria within macrophages. J Immunol. 2007 Mar 15;178(6):3767-76. doi: 10.4049/jimmunol.178.6.3767. PMID: 17339475.
56. Dunlap MD, Howard N, Das S, Scott N, Ahmed M, Prince O, et al. A novel role for C-C motif chemokine receptor 2 during infection with hypervirulent Mycobacterium tuberculosis. Mucosal Immunol. 2018 Nov;11(6):1727-1742. doi: 10.1038/s41385-018-0071-y. Epub 2018 Aug 16. PMID: 30115997; PMCID: PMC6279476.
57. Rivero-Lezcano OM, González-Cortés C, Reyes-Ruvalcaba D, Diez-Tascón C. CCL20 is overexpressed in Mycobacterium tuberculosis-infected monocytes and inhibits the production of reactive oxygen species (ROS). Clin Exp Immunol. 2010 Nov;162(2):289-97. doi: 10.1111/j.1365-2249.2010.04168.x. Epub 2010 Sep 1. PMID: 20819093; PMCID: PMC2996596.
58. Sun KT, Chen MY, Tu MG, Wang IK, Chang SS, Li CY. MicroRNA-20a regulates autophagy related protein-ATG16L1 in hypoxia-induced osteoclast differentiation. Bone. 2015 Apr;73:145-53. doi: 10.1016/j.bone.2014.11.026. Epub 2014 Dec 5. PMID: 25485521.
59. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell. 2009 Jan 23;136(2):215-33. doi: 10.1016/j.cell.2009.01.002. PMID: 19167326; PMCID: PMC3794896.
60. Kay M, Soltani BM, Aghdaei FH, Ansari H, Baharvand H. Hsa-miR-335 regulates cardiac mesoderm and progenitor cell differentiation. Stem Cell Res Ther. 2019 Jun 27;10(1):191. doi: 10.1186/s13287-019-1249-2. PMID: 31248450; PMCID: PMC6595595.
61. Chen Y, Chen Q, Zou J, Zhang Y, Bi Z. Construction and analysis of a ceRNAceRNA network reveals two potential prognostic modules regulated by hsamiR3355p in osteosarcoma. Int J Mol Med. 2018 Sep;42(3):1237-1246. doi: 10.3892/ijmm.2018.3709. Epub 2018 May 29. PMID: 29845268; PMCID: PMC6089708.
62. Li T, Gao X, Han L, Yu J, Li H. Identification of hub genes with prognostic values in gastric cancer by bioinformatics analysis. World J Surg Oncol. 2018 Jun 19;16(1):114. doi: 10.1186/s12957-018-1409-3. PMID: 29921304; PMCID: PMC6009060.

Corresponding authors

Prof. Aihua Liu, Faculty of Basic Medicine, Kunming Medical University, Kunming 650500, China.liuaihua@kmmu.edu.cn; ORCID: 0000-0001-5726-6211

Prof. Fukai Bao, Faculty of Basic Medicine, Kunming Medical University, Kunming 650500, China. baofukai@kmmu.edu.cn; ORCID: 0000-0003-2652-6660