Medicine®

OPEN

Comprehensive analyses of potential key genes in active tuberculosis

A systematic review

Jiarui Chen, MS[a], Chong Liu, MD[a], Tuo Liang, MD[a], Guoyong Xu, MS[a], Zide Zhang, MS[a], Zhaojun Lu, MS[a], Jie Jiang, MS[a], Tianyou Chen, MS[a], Hao Li, MS[a], Shengsheng Huang, MD[a], Liyi Chen, MD[a], Xihua Sun, MD[a], Jiemei Cen, MS[b], Xinli Zhan, MD[a],*

Abstract

Background: Tuberculosis (TB) is a global health problem that brings us numerous difficulties. Diverse genetic factors play a significant role in the progress of TB disease. However, still no key genes for TB susceptibility have been reported. This study aimed to identify the key genes of TB through comprehensive bioinformatics analysis.

Methods: The series microarray datasets from the gene expression omnibus (GEO) database were analyzed. We used the online tool GEO2R to filtrate differentially expressed genes (DEGs) between TB and health control. Database for annotation can complete gene ontology function analysis as well as Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis. Protein-protein interaction (PPI) networks of DEGs were established by STRING online tool and visualized by Cytoscape software. Molecular Complex Detection can complete the analysis of modules in the PPI networks. Finally, the significant hub genes were confirmed by plug-in Genemania of Cytoscape, and verified by the verification cohort and protein test.

Results: There are a total of 143 genes were confirmed as DEGs, containing 48 up-regulated genes and 50 down-regulated genes. The gene ontology and Kyoto Encyclopedia of Genes and Genomes analysis show that upregulated DEGs were associated with cancer and phylogenetic, whereas downregulated DEGs mainly concentrate on inflammatory immunity. PPI networks show that signal transducer and activator of transcription 1 (STAT1), guanylate binding protein 5 (GBP5), 2′-5′-oligoadenylate synthetase 1 (OAS1), catenin beta 1 (CTNNB1), and guanylate binding protein 1 (GBP1) were identified as significantly different hub genes.

Conclusion: We conclude that these genes, including TAT1, GBP5, OAS1, CTNNB1, GBP1 are a candidate as potential core genes in TB and treatment of TB in the future.

Abbreviations: CTNNB1 = catenin beta 1, DEG = differentially expressed gene, GBP1 = guanylate binding protein 1, GBP5 = guanylate binding protein 5, GEO = gene expression omnibus, GO = gene ontology, IFI44 = interferon-induced protein 44 like, KEGG = Kyoto Encyclopedia of Genes and Genomes, MTB = *Mycobacterium tuberculosis*, OAS1 = 2′-5′-oligoadenylate synthetase 1, OAS2 = 2′-5′-oligoadenylate synthetase 2, OAS3 = 2′-5′-oligoadenylate synthetase 3, PARP9 = poly(ADP-ribose) polymerase family member 9, PPI = protein-protein interaction , SOCS3 = suppressor of cytokine signaling 3, STAT1 = signal transducer and activator of transcription 1, TB = tuberculosis, XAF1 = XIAP associated factor 1.

Keywords: expression profiling data, GEO, hub genes, tuberculosis

1. Introduction

Tuberculosis (TB) is a disease caused by *Mycobacterium tuberculosis* (MTB) infection and is still a cause of global death. The World Health Organization reported that about 10.0 million people suffered from TB disease in 2017, including males and females, 5.8 million and 3.2 million, respectively. And about 1.3 million among human immunodeficiency virus-negative and 300,000 among human immunodeficiency virus-positive deaths were caused by TB in 2017.[1]

Only 1 in 10 people infected with TB will become infected, suggesting that TB links to some factors. We have recognized that TB is caused by infection with MTB, while it is still unclear why some people develop active TB after infection while others are latent. The incidence of TB after infection varies among individuals. Host genetic factors have been shown to contribute to MTB infection outcomes. With the development and spread of science and technology, some high-throughput technology platforms for gene expression analysis have rapidly developed. Microarray analysis, as an important tool in the field of tumor medicine, has important clinical application value in the key genes, molecular diagnosis of some diseases, molecular diagnosis, and the discovery of new drug targets.[2,3] There are already many disease databases, such as gene expression omnibus (GEO)[4] and the cancer genome atlas. Also, more and more researchers are using bioinformatics analysis. Expression profiling by array plays an important role in the field of molecular biology and it can simultaneously detect thousands of genes.[5] Nowadays, many researchers have studied the field of TB pathogenesis at the genetic level, such as single nucleotide polymorphisms.[6,7] However, key genes associated with TB have not been reported until now. Thus, our study uses online bioinformatics data to explore differentially expressed genes (DEGs) between TB and healthy controls. The significant DEGs were analyzed by gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment. The whole-genome analysis will reveal new genetic mutations related to TB. And after analyzing the biological functions and pathways of the DEGs, this study provides ideas for the molecular level research of TB and identifies core candidate genes for the future diagnosis, treatment as well as prognosis of patients infected by TB.

2. Materials and methods

2.1. Data source

The gene expression datasets profiling of TB and health control analyzed in our study were got from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). After a careful review, we select 2 series about TB as an observational cohort and 1 series as validation cohort from the database, all of them are from the Illumina GPL6947 platform (HumanHT-12 V3.0 expression bead chip). All of the data were freely available online, and we did not do any experiments on humans or animals.

2.2. Data processing of DEGs

The DEGs were conducted by a free online tool, GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r/), between TB and health control, and the *P* value and |log FC| were calculated at the same time. We set the DEGs at a level with *P* value <.05 and |log FC| > 1.0. After a statistical analysis of each dataset, the Venn diagram drawn by R was used to identify the intersecting parts.

2.3. GO and KEGG pathway analysis of DEGs

Database for Annotation, Visualization, and Integrated Discovery[8] tools (https://david.ncifcrf.gov/) will be used to completed GO annotation[9] and KEGG pathway enrichment[10] analysis of DEGs. GO analysis divided gene function into biological process (BP), molecular function (MF), and cellular component (CC). KEGG, storing large amounts of data on biological pathways and diseases, is a widely used database. We set the cutoff criterion at a level with the *P* value <.05.

2.4. Construction of PPI network and identification of key gene

We use the Search Tool for the Retrieval of Interacting Genes database (http://string-db.org/) to complete protein-protein interaction (PPI) information analysis. To assess the potential PPI relationship, we uploaded all the identified DEGs to the STRING database. The PPI pairs were extracted using the combined score of 0.4. Then, the entire giant network was entered into the Cytoscape software[11] (version 3.7.1) and using the plug-in Molecular Complex Detection to evaluate the main modules of the PPI subnet. The standard is set to the default value. GeneMANIA[12] has collected hundreds of datasets and organism-specific functional genomics data. And we re-predict top hub genes with plug-in GeneMANIA of Cytoscape.

2.5. Variation of the key genes

GSE19439, containing 12 health people and 13 TB patients, acted as a validation cohort and was analyzed by GEO2R. And the volcanos of the datasets were drawn by a free online tool Sangerbox (http://sangerbox.com/AllTools?tool_id=9699135). Furthermore, TB with spinal TB (n = 5) and degenerative intervertebral disc tissues (n = 5) were collected for validation of the key genes using the label-free quantification proteomics technique. The informed consent of all participants was obtained, and it was approved by the Ethics Committee of the First Affiliated Hospital of Guangxi Medical University.

3. Results

3.1. Identification of DEGs

Our study contains 3 series of expression profiles, and the details can be seen in Table 1. We set the criteria at a level of *P* value <.05 and |log FC| > 1. A total of 1006 DEGs were identified from GSE19443, including 642 upregulated genes and 379 downregulated genes. And 1125 DEGs were identified from GSE19444, including 564 upregulated genes and 570 downregulated genes. All DEGs were determined by comparing active TB with healthy control samples. The intersection of the DEGs profiles was performed by Venn analysis (Fig. 1). A total of 143

Table 1

Statistics of the 3 microarray databases derived from the GEO database.

Dataset ID	ATB	HC	Total number
GSE19443	28	16	44
GSE19444	12	12	24
GSE19439	13	12	25

ATB = active tuberculosis, GEO = gene expression omnibus, HC = health control.
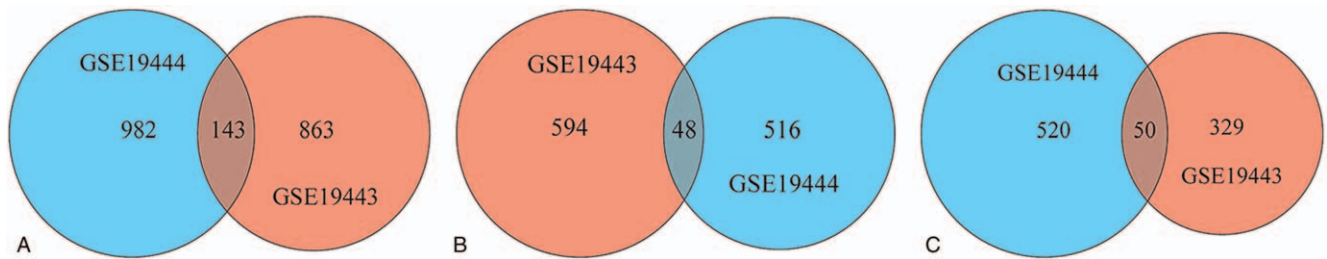
Figure 1. Venn diagram of differentially expressed genes common to 2 GEO datasets. (A) to (C) Represent the Venn diagrams of total, up-regulated DEGs, and down-regulated differentially expressed genes, respectively. GEO = gene expression omnibus.

Table 2

Gene ontology analysis of differentially expressed genes associated with tuberculosis.

Sort	Category	Term	Count	%	P value	Genes
Upregulate	GOTERM_BP_FAT	GO:0072006~nephron development	5	10.41666667	2.97E-04	MAGI2, LAMA5, HOXA11, TEK, CTNNB1
	GOTERM_BP_FAT	GO:0001822~kidney development	6	12.5	3.76E-04	MAGI2, LAMA5, HOXA11, TEK, ID3, CTNNB1
	GOTERM_BP_FAT	GO:0072001~renal system development	6	12.5	4.91E-04	MAGI2, LAMA5, HOXA11, TEK, ID3, CTNNB1
	GOTERM_BP_FAT	GO:0001655~urogenital system development	6	12.5	8.56E-04	MAGI2, LAMA5, HOXA11, TEK, ID3, CTNNB1
	GOTERM_BP_FAT	GO:0072009~nephron epithelium development	4	8.333333333	0.00221307	MAGI2, LAMA5, HOXA11, CTNNB1
	GOTERM_CC_FAT	GO:0045177~apical part of cell	4	8.333333333	0.044370181	HPN, SORBS2, TEK, CTNNB1
	GOTERM_MF_FAT	GO:0005198~structural molecule activity	6	12.5	0.045208057	RPS26, MAGI2, LAMA5, SORBS2, RPS26P11, COL4A5
	GOTERM_MF_FAT	GO:0003712~transcription cofactor activity	5	10.41666667	0.049206991	RBPMS, USP21, RUNX1T1, ID3, CTNNB1
Downregulate	GOTERM_BP_FAT	GO:0045087~innate immune response	17	34.69388	4.48E-11	GBP5, SOCS3, OAS3, OAS1, OAS2, SLAMF7, STAT1, C1QC, AIM2, GCH1, C1QB, IFI27, PARP9, FCGR1B, XAF1, SSC5D, GBP1
	GOTERM_BP_FAT	GO:0006952~defense response	19	38.77551	4.50E-09	GBP5, SOCS3, OAS3, OAS1, IFI44L, OAS2, SLAMF7, STAT1, C1QC, SP140, AIM2, GCH1, C1QB, IFI27, PARP9, FCGR1B, XAF1, SSC5D, GBP1
	GOTERM_BP_FAT	GO:0034341~response to interferon-gamma	9	18.36735	4.68E-09	GBP5, SOCS3, FCGR1B, OAS3, OAS1, OAS2, STAT1, GBP1, GCH1
	GOTERM_BP_FAT	GO:0034097~response to cytokine	14	28.57143	3.43E-08	IFI27, GBP5, PARP9, SOCS3, FCGR1B, OAS3, OAS1, OAS2, XAF1, STAT1, GAS6, AIM2, GBP1, GCH1
	GOTERM_BP_FAT	GO:0006955~immune response	18	36.73469	4.74E-08	GBP5, SOCS3, OAS3, OAS1, OAS2, SLAMF7, STAT1, C1QC, AIM2, GAS6, GCH1, C1QB, IFI27, PARP9, FCGR1B, XAF1, SSC5D, GBP1
	GOTERM_MF_FAT	GO:0001730~2′-5′-oligoadenylate synthetase activity	3	6.122449	3.51E-05	OAS3, OAS1, OAS2
3	GOTERM_MF_FAT	GO:0070566~adenylyltransferase activity	3	6.122449	0.001566	OAS3, OAS1, OAS2
	GOTERM_MF_FAT	GO:0003725~double-stranded RNA binding	3	6.122449	0.011076	OAS3, OAS1, OAS2
	GOTERM_MF_FAT	GO:0005525~GTP binding	5	10.20408	0.014158	GBP5, IFI44L, GBP4, GBP1, GCH1
	GOTERM_MF_FAT	GO:0032561~guanyl ribonucleotide binding	5	10.20408	0.016904031	GBP5, IFI44L, GBP4, GBP1, GCH1

CTNNB1 = catenin beta 1, GBP1 = guanylate binding protein 1, GBP5 = guanylate binding protein 5, GO = gene ontology, IFI44L = interferon-induced protein 44 like, OAS1 = 2′-5′-oligoadenylate synthetase 1, OAS2 = 2′-5′-oligoadenylate synthetase 2, PARP9 = poly(ADP-ribose) polymerase family member 9, SOCS3 = suppressor of cytokine signaling 3, STAT1 = signal transducer and activator of transcription 1.

KEGG pathway analysis of differentially expressed genes associated with tuberculosis.

Sort	Category	Term	Count	%	P value	Genes
Upregulate	KEGG_PATHWAY	hsa05200:Pathways in cancer	6	12.5	0.006960693	LAMA5, RUNX1T1, PRKACB, MAPK10, CTNNB1, COL4A5
	KEGG_PATHWAY	hsa04510:Focal adhesion	4	8.333333	0.026794599	LAMA5, MAPK10, CTNNB1, COL4A5
	KEGG_PATHWAY	hsa04014:Ras signaling pathway	4	8.333333	0.033997769	KSR2, TEK, PRKACB, MAPK10
	KEGG_PATHWAY	hsa05146:Amoebiasis	3	6.25	0.044473305	LAMA5, PRKACB, COL4A5
Downregulate	KEGG_PATHWAY	hsa05160:Hepatitis C	5	10.20408163	2.61E-04	SOCS3, OAS3, OAS1, OAS2, STAT1
	KEGG_PATHWAY	hsa05164:Influenza A	5	10.20408163	7.27E-04	SOCS3, OAS3, OAS1, OAS2, STAT1
	KEGG_PATHWAY	hsa05168:Herpes simplex infection	5	10.20408163	8.79E-04	SOCS3, OAS3, OAS1, OAS2, STAT1
	KEGG_PATHWAY	hsa05162:Measles	4	8.163265306	0.003939807	OAS3, OAS1, OAS2, STAT1

CTNNB1 = catenin beta 1, KEGG = Kyoto Encyclopedia of Genes and Genomes, OAS1 = 2′-5′-oligoadenylate synthetase 1, OAS2 = 2′-5′-oligoadenylate synthetase 2, OAS3 = 2′-5′-oligoadenylate synthetase 3, SOCS3 = suppressor of cytokine signaling 3, STAT1 = signal transducer and activator of transcription 1, XAF1 = XIAP associated factor 1.
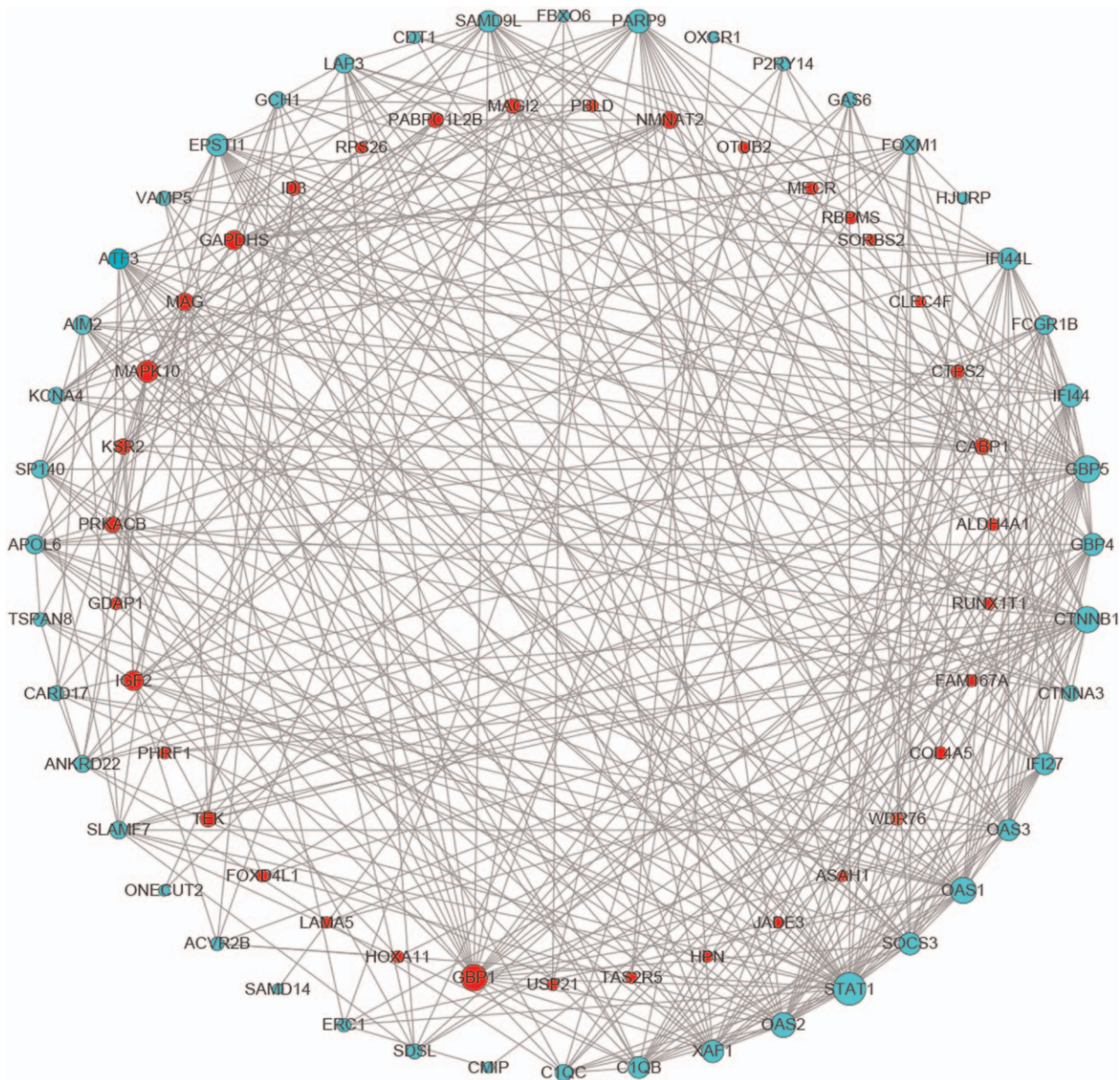
Figure 2. Protein-protein interaction network constructed by the differentially expressed genes. Red nodes belong to upregulated genes and blue nodes belong to downregulated genes.

Table 4

Top 12 hub genes with higher degree of connectivity.

Gene symbol	Gene description	Degree	Up/dowm
STAT1	Signal transducer and activator of transcription 1	35	Downregulated
GBP5	Guanylate binding protein 5	26	Downregulated
OAS1	2′-5′-Oligoadenylate synthetase 1	25	Downregulated
CTNNB1	Catenin beta 1	25	Downregulated
GBP1	guanylate binding protein 1	25	Downregulated
OAS2	2′-5′-Oligoadenylate synthetase 2	23	Downregulated
PARP9	Poly (ADP-ribose) polymerase family member 9	21	Downregulated
IFI44	Interferon induced protein 44 like	20	Downregulated
XAF1	XIAP associated factor 1	19	Downregulated
SOCS3	Suppressor of cytokine signaling 3	19	Downregulated
GBP4[*]	Guanylate binding protein 4	19	Downregulated
EPSTI1[*]	Epithelial stromal interaction 1 (breast)	19	Downregulated

[*] GBP4 and EPSTI1 have the same degree, so we showed them in the table too.

DEGs were significantly differentially expressed between 2 groups, including 48 upregulated and 50 downregulated genes.

3.2. GO term enrichment analysis

To search the functions of the DEGs, we use the online tool Database for Annotation, Visualization, and Integrated Discovery to identify GO categories and KEGG pathways. We uploaded all the DEGs to the gene list separately. After GO analysis we know that the upregulated DEGs were mainly relevant to BP, CC, and MF. BP includes nephron development, kidney development, renal system development, urogenital system development, and nephron epithelium development. CC only includes the apical part of the cell. MF includes structural molecule activity and transcription cofactor activity with significance. And GO analysis showed that the downregulated DEGs were mainly enriched in BP and MF, while there is no significance in CC. BPs contain an innate immune response, defense response, response to interferon-gamma, response to cytokine, immune response, and so on. And MF includes oligoadenylate synthetase activity, adenylyl-transferase activity, double-stranded RNA binding, GTP binding,
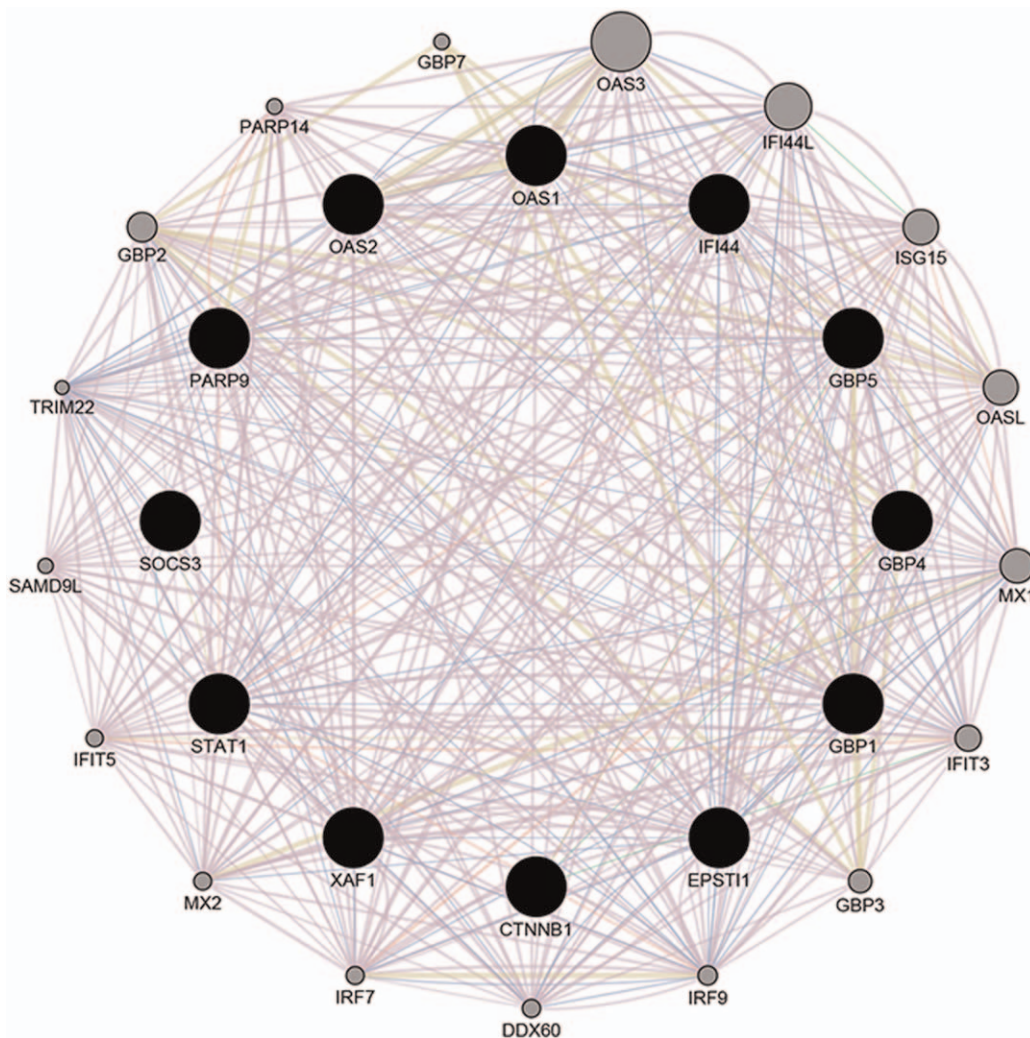
Figure 3. GeneMANIA analysis protein-protein interaction network of top 12 hub genes with high value of degree and betweenness. Black notes belong to the top 12 differentially expressed genes, and others belong to the GeneMania predicted genes.

Table 5

Gene functional enrichment in GeneMANIA analysis.

Function	FDR	Genes in network
Cellular response to type I interferon	5.02E-17	17
Type I interferon signaling pathway	5.02E-17	17
Response to type I interferon	5.02E-17	17
Response to virus	1.56E-09	17
Response to interferon-gamma	5.74E-09	13

FDR = false discovery rate.

guanyl ribonucleotide binding, and so on. There is no significance in CC (Table 2).

3.3. KEGG pathway analysis

The enrichment of all upregulated DEGs and downregulated DEGs was identified separately with KEGG analysis. From the result, we found that upregulated DEGs enrich pathways in cancer, focal adhesion, Ras signaling pathway, and amoebiasis. And 5 of 50 downregulated DEGs were enriched in the same 3 pathways. Genes include suppressor of cytokine signaling 3 (SOCS3), OAS3, 2′-5′-oligoadenylate synthetase 1 (OAS1), 2′-5′-oligoadenylate synthetase 2 (OAS2), and signal transducer and activator of transcription 1 (STAT1). And OAS3, OAS1, OAS2, and STAT1 also enrich in Measles pathway (Table 3).

3.4. PPI network construction and identification of hub gene

After uploading the DEGs into the STRING database, we got a PPI network with 90 nodes and 391 edges (Fig. 2). The top genes for PPI network connectivity assessment were identified (Table 4). The results showed that STAT1 was the most prominent gene with connectivity degree = 35, followed by guanylate binding protein 5 (GBP5; degree = 26), OAS1 (degree = 25), catenin beta 1 (CTNNB1; degree = 25), guanylate binding protein 1 (GBP1; degree = 25), OAS2 (degree = 23), poly(ADP-ribose) polymerase family member 9 (PARP9; degree = 21), interferon-induced protein 44 like (IFI44; degree=20), XIAP associated factor 1 (XAF1; degree = 19), and SOCS3 (degree = 19). All of the hub genes are downregulated. The GeneMANIA Cytoscape plugin

can bring fast gene function prediction. GeneMANIA uses the intrinsic association method to identify the most relevant genes in the query gene set. The plugin uses a large database of functional interaction networks from multiple organisms, and each related gene can be traced back to the source network used for prediction. After GeneMania analysis of top 12 hub genes (Fig. 3), we found that the functions of top 12 mainly enriched in the cellular response to type I interferon, type I interferon signaling pathway, response to type I interferon, response to virus, and response to interferon-gamma (Table 5).

3.5. The verify of hub genes

To further evaluate the key genes, the GSE19439 was analyzed by GEO2R, and the volcano draw by the Sangerbox showed that the key genes were all lower expressed with a significance in 3 datasets (Fig. 4). Additionally, STAT1, GBP1, GBP4, and OAS1 were randomly selected to detect their protein expression in spinal TB and degenerative intervertebral disc tissues. The result showed that all of them were lower expressed in spinal TB than degenerative intervertebral disc tissues. Though, there was no significant difference in the GBP4 (Fig. 5).

4. Discussion

As a lung infection disease caused by MTB, TB has been a health threat for hundreds of years and is 1 of the top 10 causes of death around the world. MTB is a kind of bacteria, while its pathogenesis differs from ordinary bacteria, and it has very strong resistance to the outside environment. It is not sensitive to commonly used sulfonamides and various antibiotics. It is sensitive to streptomycin, isoniazid, rifampicin, and ethambutol. With the long-term use of these drugs, the treatment of TB is easy to produce drug-resistant strains. It is significant to identify key genes and new TB-specific targeted therapies.

In our study, public databases were used for gene expression and protein expression analysis to identify potential key genes associated with TB. Finally,143 differentially expressed genes were identified in TB compared to healthy controls, including 48 upregulated DEGs and 50 downregulated DEGs. GO analysis shows that both upregulated DEGs and downregulated DEGs were mainly enriched in BP, while downregulated DEGs were
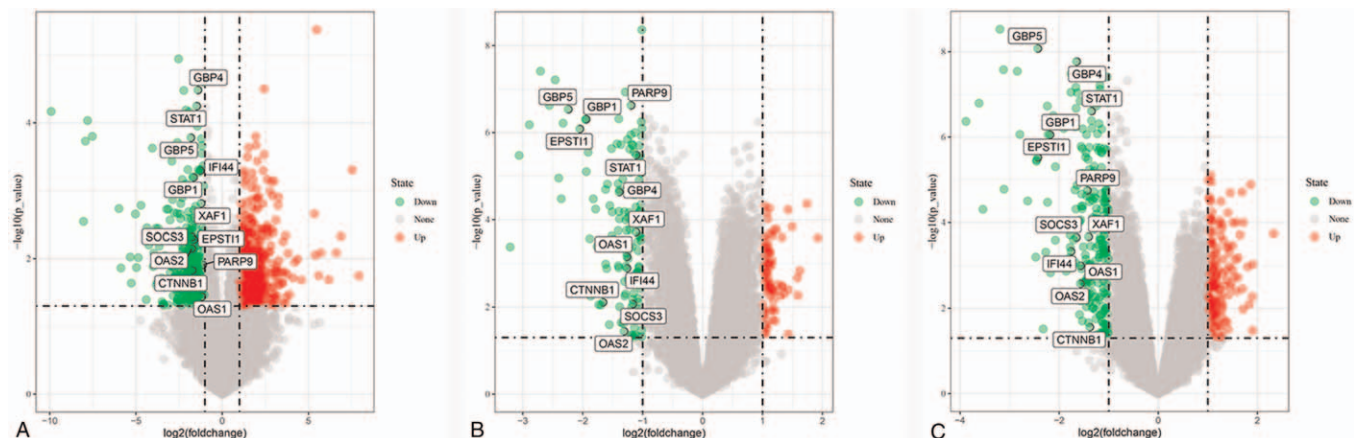
Figure 4. The volcanos of the differentially expressed gene. (A) to (C) Represent the volcano of DEGs from GSE19443, GSE19444, and GSE19439, respectively. Red nodes belong to upregulated genes, and green nodes belong to downregulated genes. DEG = differentially expressed gene.
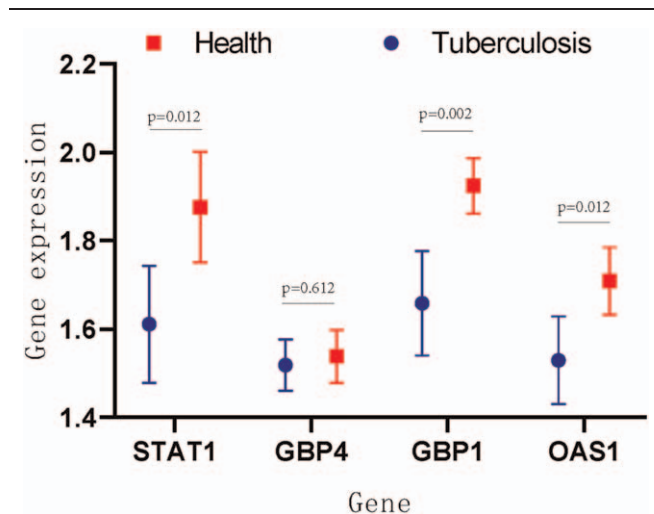
Figure 5. The proteins expression level of the randomly selected gene.

also enriched in MF. And all DEGs were almost no enrichment enriched in CC with significance. BP indicated that upregulated DEGs were mainly related to system development and cancer, while downregulated DEGs were mainly related to terms of inflammation and immunity. And KEGG analysis reveals that downregulated DEGs were mainly concentrated in the Virus pathway, including Hepatitis C, Influenza A, Herpes simplex infection, and Measles. Up-regulated DEGs were enriched in pathways in cancer, and other pathways related to the term of cell movement, adhesion, transfer, were also closely associated with the development of tumors. Maybe that is why MTB, with different pathogenesis, is different from ordinary bacteria. The PPI network was constructed to better study the interrelationship of DEGs and the top central genes were determined by connectivity, including STAT1, GBP5, OAS1, CTNNB1, GBP1, OAS2, PARP9, IFI44, XAF1, and SOCS3. All these genes were down-regulated.

STAT1, belonging to the STAT family,[13] mediates the effects of many cytokines involved in the innate and adaptive immune responses to viruses and intracellular bacteria.[14,15] It transduces activities of many cytokines, including type I-III interferons (IFNs), interleukin (IL)-21, IL-27 as well as IL-35.[16–18] The phosphorylated STAT1 has been known as a pathogen suppressor because its function serves as an immunosurveillance mediator.[17,19] STAT1 expression is associated with a good prognosis in several types of cancer, such as colon cancer,[20,21] hepatocellular carcinoma,[22] esophageal cancer,[23] pancreatic cancer,[24] soft tissue sarcoma,[25] and metastatic melanomas.[26] It has been proved that STAT1 gene mutations on autosomes are associated with susceptibility to TB.[16] It means that STAT1 plays a very important role in TB. In our study, STAT1 was downregulated in TB compared to healthy control. Therefore, STAT1 may be a key gene in TB.

Guanylate-binding proteins (GBPs), including 7 isoforms (GBP1–GBP7), are GTPases induced by IFN, which have central roles in cell-autonomous immunity,[27–30] regulates antibacterial immunity and intrinsic antiviral, and predicts the progress of several infectious diseases.[31] Among the top 10 key genes we studied, GBP1[32] and GBP5[33,34] are among the family members. Qiu et al had confirmed that the relationship between mitochondrial dysfunction and macrophage cell senescence is

associated with the down-regulation of GBP1.[35] GBP5 splice variant is associated with the tumor.[36] In our study, GBP5 was downregulated in TB, while upregulated in other diseases.[37]

$2'$-$5'$-oligoadenylate synthetases (OASs), contained OAS1, OAS2, and OAS3, are enzymes induced by IFN that play significant roles in the innate immune response when against viruses.[38,39] Some researchers believe that polymorphisms of OAS1 affect the expression of IFN-γ, the elimination of MTB, and ultimately affect the development of TB.[40,41] OAS1 mainly affects susceptibility to viral infections.[42,43] All the OASs family members are downregulated in TB, while some researchers observe OASs upregulation in some genes expression traits that distinguish active from latent TB infection.[44–47] This is different from the findings of most people, probably because of differences in expressions at different stages after TB infection.

The CTNNB1 gene, also known as the β-catenin gene, is activated in the WNT signaling pathway of cancer cells and acts to inhibit the activity of GSK-3β, thereby stabilizing the CTNNB1 protein.[48,49] In recent years, many literature have reported that the increase of CTNNB1 gene expression plays an important role in the formation of tumors such as colon cancer[50] and ovarian cancer.[51] Once the activity of CTNNB1 is restricted in ovarian cancer, the cell cycle will soon stop in the G1 phase and cells will begin to apoptosis.[52,53] Recent evidence has indicated that the WNT pathway has an etiological role in TB disease.[54–56] One study showed that Wnt/β-catenin acts as a negative feedback loop, inhibiting inflammation stimulated y M.bovis BCG (BCG). Thereby, attenuating toll-like receptor-induced proinflammatory responses in alveolar epithelial cells.[55] Chinese researchers reported that the distribution of CTNNB1 gene polymorphism differs between Chinese Tibetan and Han populations with TB.[57]

From the PPI network, we know that the downregulated genes are more aggregated and interact more strongly, meaning that these genes play a more important role in the pathogenesis of TB. The top 10 genes were identified by evaluating with connectivity degree in the PPI network. After re-predicted the hub genes by GeneMania, we found that the functions of them mainly enriched in the immune aspect, for example, cellular response to interferon. The verification cohort of GSE19439 supported that the key genes were down-regulated in TB. Furthermore, the proteins expression of the randomly selected gene also supported our conclusion. Though, it was not distinctively different in GBP4. The top 12 genes we analyzed were mainly downregulated genes, and the reduction of their expression levels may lead to more susceptibility to TB and got TB.

5. Conclusion

After compared TB and heath control, we identified 143 DEGs, including 48 upregulated DEGs and 50 downregulated DEGs. Top 12 hub genes were selected, and they were mainly enriched in the immune aspect. The reduction of their expression maybe is the main cause of TB. Top 5 genes, including STAT1, GBP5, OAS1, CTNNB1, and GBP1, are a candidate as potential core genes in TB. The verification cohort and the detection of proteins supported the results. Further study is needed to invested how the key genes perform and how to use these hub genes to help us prevent, early diagnose, and targeted treatment of TB.

Author contributions

Conceptualization: Tuo Liang, Jie Jiang.

References

[1] Global Tuberculosis Report 2018, 2018.

[2] Chai Y, Tan F, Ye S, Liu F, Fan Q. Identification of core genes and prediction of miRNAs associated with osteoporosis using a bioinformatics approach. Oncol Lett 2019;17:468–81.

[3] Xia M, Zhang Q, Luo M, et al. Regulatory network analysis reveals the oncogenesis roles of feed-forward loops and therapeutic target in T-cell acute lymphoblastic leukemia. BMC Med Genomics 2019;12:8.

[4] Kazezian Z, Gawri R, Haglund L, et al. Gene expression profiling identifies interferon signalling molecules and IGFBP3 in human degenerative annulus fibrosus. Sci Rep 2015;5(2045–2322 (Electronic)):15662.

[5] Kononen J, Bubendorf L, Kallioniemi A, et al. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 1998;4:844–7.

[6] Silva-Ramirez B, Saenz-Saenz CA, Bracho-Vela LA, et al. Association between vitamin D receptor gene polymorphisms and pulmonary tuberculosis in a Mexican population. Indian J Tuberc 2019;66:70–5. (0019-5707 (Print)).

[7] Naz A, Aslam MA, Khan AUH, et al. Genetic polymorphism in association with susceptibility to tuberculosis: a study in a Pakistani population. Braz J Microbiol 2019;50:429–34. doi: 10.1007/s42770-019-00048-8 (1678-4405 (Electronic)).

[8] Huang DW, Sherman BT, Tan Q, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol 2007;8:R183.

[9] Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol 2010;11:R14.

[10] Mao X, Cai T, Olyarchuk JG, Wei L. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics 2005;21:3787–93.

[11] Su G, Morris JH, Demchak B, Bader GD. Biological network exploration with Cytoscape 3. Curr Protoc Bioinformatics 2014;47: 8.13.1-24 (1934–340X (Electronic)).

[12] Franz M, Rodriguez H, Lopes C, et al. GeneMANIA update 2018. Web Server Issue 2018;46:W60–4.

[13] Copeland NG, Gilbert DJ, Schindler C, et al. Distribution of the mammalian Stat gene family in mouse chromosomes. Genomics 1995;29:225–8.

[14] Casanova JL, Holland SM, Notarangelo LD. Inborn errors of human JAKs and STATs. Immunity 2012;36:515–28.

[15] Decker T, Stockinger S, Karaghiosoff M, Müller M, Kovarik P. IFNs and STATs in innate immunity to microorganisms. J Clin Invest 2002;109:1271–7.

[16] Boisson-Dupuis S, Kong XF, Okada S, et al. Inborn errors of human STAT1: allelic heterogeneity governs the diversity of immunological and infectious phenotypes. Curr Opin Immunol 2012;24:364–78.

[17] Najjar I, Fagard R. STAT1 and pathogens, not a friendly relationship. Biochimie 2010;92:425–44. (1638-6183 (Electronic)).

[18] Vignali DA, Kuchroo VK. IL-12 family cytokines: immunological playmakers. Nat Immunol 2012;13:722–8.

[19] Lim YJ, Yi MH, Choi JA, et al. Roles of endoplasmic reticulum stress-mediated apoptosis in M1-polarized macrophages during mycobacterial infections. Sci Rep 2016;6:37211(2045–2322 (Electronic)).

[20] Gordziel C, Bratsch J, Moriggl R, Knösel T, Friedrich K. Both STAT1 and STAT3 are favourable prognostic determinants in colorectal carcinoma. Br J Cancer 2013;109:138–46.

[21] Slattery ML, Lundgreen A, Kadlubar SA, Bondurant KL, Wolff RK. JAK/STAT/SOCS-signaling pathway and colon and rectal cancer. Mol Carcinog 2013;52:155–66.

[22] Chen G, Wang H, Xie S, Ma J, Wang G. STAT1 negatively regulates hepatocellular carcinoma cell proliferation. Oncol Rep 2013;29:2303–10.

[23] Zhang Y, Molavi O, Su M, Lai R. The clinical and biological significance of STAT1 in esophageal squamous cell carcinoma. BMC Cancer 2014; (1471-2407 (Electronic)).

[24] Sun Y, Yang S, Sun N, Chen J. Differential expression of STAT1 and p21 proteins predicts pancreatic cancer progression and prognosis. Pancreas 2014;43:619–23.

[25] Takahashi A, Nakayama R, Ishibashi N, et al. Analysis of gene expression profiles of soft tissue sarcoma using a combination of knowledge-based filtering with integration of multiple statistics. PloS One 2014;9:e106801.

[26] Osborn JL, Greer SF. Metastatic melanoma cells evade immune detection by silencing STAT1. Int J Mol Sci 2015;16:4343–61.

[27] Pilla-Moffett D, Barber MF, Taylor GA, Coers J. Interferon-inducible GTPases in host resistance, inflammation and disease. J Mol Biol 2016;428:3495–513.

[28] Meunier E, Broz P. Interferon-inducible GTPases in cell autonomous and innate immunity. Cell Microbiol 2016;18:168–80.

[29] Kim BH, Chee JD, Bradfield CJ, Park ES, Kumar P, MacMicking JD. Interferon-induced guanylate-binding proteins in inflammasome activation and host defense. Nat Immunol 2016;17:481–9.

[30] Man SM, Place DE, Kuriakose T, Kanneganti TD. Interferon-inducible guanylate-binding proteins at the interface of cell-autonomous immunity and inflammasome activation. J Leukoc Biol 2017;101:143–50.

[31] Piro AS, Hernandez D, Luoma S, et al. Detection of cytosolic Shigella flexneri via a C-terminal triple-arginine motif of GBP1 inhibits actin-based motility. mBio 2017;8:e01979-17doi: 10.1128/mBio.01979-17 (2150-7511 (Electronic)).

[32] Pan W, Zuo X, Feng T, Shi X, Dai J. Guanylate-binding protein 1 participates in cellular antiviral response to dengue virus. Virol J 2012;9:292(1743-422X (Electronic)).

[33] Kim BH, Shenoy AR, Kumar P, Bradfield CJ, MacMicking JD. IFN-inducible GTPases in host cell defense. Cell Host Microbe 2012;12:432–44.

[34] Hotter D, Sauter D, Kirchhoff F. Guanylate binding protein 5: impairing virion infectivity by targeting retroviral envelope glycoproteins. Small GTPases 2017;8:31–7.

[35] Qiu X, Guo H, Yang J, Ji Y, Wu CS, Chen X. Down-regulation of guanylate binding protein 1 causes mitochondrial dysfunction and cellular senescence in macrophages. Sci Rep 2018;8:1679.

[36] Wehner M, Herrmann C. Biochemical properties of the human guanylate binding protein 5 and a tumor-specific truncated splice variant. FEBS J 2010;277:1597–605.

[37] Wang Q, Wang X, Liang Q, et al. Distinct prognostic value of mRNA expression of guanylate-binding protein genes in skin cutaneous melanoma. Oncol Lett 2018;15:7914–22.

[38] Lohofener J, Steinke N, Kay-Fedorov P, et al. The activation mechanism of 2′-5′-oligoadenylate synthetase gives new insights into OAS/cGAS triggers of innate immunity. Structure 2015;23:851–62.

[39] Li Y, Banerjee S, Wang Y, et al. Activation of RNase L is dependent on OAS3 expression during infection with diverse human viruses. Proc Natl Acad Sci U S A 2016;113:2241–6.

[40] Leisching G, Wiid I, Baker B. OAS1, 2, and 3: significance during active tuberculosis? J Infect Dis 2018;217:1517–21.

[41] Leisching G, Wiid I, Baker B. The association of OASL and type I interferons in the pathogenesis and survival of intracellular replicating bacterial species. Front Cell Infect Microbiol 2017;7:196(2235–2988 (Electronic)).

[42] Tsutsui-Takeuchi M, Ushio H, Fukuda M, et al. Roles of retinoic acid-inducible gene-I-like receptors (RLRs), toll-like receptor (TLR) 3 and 2′-5′ oligoadenylate synthetase as viral recognition receptors on human mast cells in response to viral infection. Immunol Res 2015;61:240–9.

[43] Liu X, Xing H, Gao W, et al. A functional variant in the OAS1 gene is associated with Sjogren's syndrome complicated with HBV infection. Sci Rep 2017;7: (2045-2322 (Electronic)).

[44] Leisching G, Cole V, Ali AT, Baker B. OAS1, OAS2 and OAS3 restrict intracellular M.tb replication and enhance cytokine secretion. Int J Infect Dis 2019;80S:S77–84. doi: 10.1016/j.ijid.2019.02.029 (1878-3511 (Electronic)).

[45] Berry MP, Graham CM, McNab FW, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. Nature 2010;466:973–7.

[46] Maertzdorf J, Weiner J3rd, Mollenkopf HJ, et al. Common patterns and disease-related signatures in tuberculosis and sarcoidosis. Proc Natl Acad Sci U S A 2012;109:7853–8.

[47] Ottenhoff TH, Dass RH, Yang N, et al. Genome-wide expression profiling identifies type 1 interferon response pathways in active tuberculosis. PloS one 2012;7:e45839.

[48] Yan X, Lyu T, Jia N, Yu Y, Hua K, Feng W. Huaier aqueous extract inhibits ovarian cancer cell motility via the AKT/GSK3beta/beta-catenin pathway. PloS One 2013;8:e63731.

[49] Thakur R, Mishra DP. Pharmacological modulation of beta-catenin and its applications in cancer therapy. J Cell Mol Med 2013;17:449–56.

[50] Feng Y, Sakamoto N, Wu R, et al. Tissue-specific effects of reduced beta-catenin expression on adenomatous polyposis coli mutation-instigated tumorigenesis in mouse colon and ovarian epithelium. PLoS Genet 2015;11:e1005638.

[51] Yoshioka S, King ML, Ran S, et al. WNT7A regulates tumor growth and progression in ovarian cancer through the WNT/beta-catenin pathway. Mol Cancer Res 2012;10:469–82.

[52] Ford CE, Jary E, Ma SSS Nixdorf##S, Heinzelmann-Schwarz VA, Ward RL. The Wnt gatekeeper SFRP4 modulates EMT, cell migration and downstream Wnt signalling in serous ovarian cancer cells. PloS One 2013;8:e54362.

[53] Kwon M, Lee SJ, Wang Y, et al. Filamin A interacting protein 1-like inhibits WNT signaling and MMP expression to suppress cancer cell invasion and metastasis. Int J Cancer 2014;135:48–60.

[54] Wu X, Deng G, Hao X, et al. A caspase-dependent pathway is involved in Wnt/beta-catenin signaling promoted apoptosis in Bacillus Calmette-Guerin infected RAW264.7 macrophages. Int J Mol Sci 2014;15:5045–62.

[55] Li Y, Shi J, Yang J, et al. A Wnt/beta-catenin negative feedback loop represses TLR-triggered inflammatory responses in alveolar epithelial cells. Mol Immunol 2014;59:128–35.

[56] Schaale K, Brandenburg J, Kispert A, Leitges M, Ehlers S, Reiling N. Wnt6 is expressed in granulomatous lesions of *Mycobacterium tuberculosis*-infected mice and is involved in macrophage differentiation and proliferation. J Immunol 2013;191:5182–95.

[57] Zhou WJ, Hu XJ, Zhang JY, et al. [Association of gene polymorphisms in Wnt signal pathway with tuberculosis in Chinese Tibetan population]. Sichuan Da Xue Xue Bao Yi Xue Ban 2016;47:920–5.