Wang *et al. Zoonoses* (2023) 3:1
DOI 10.15212/ZOONOSES-2022-0035

ORIGINAL ARTICLE

RNA-seq Analysis of the BCG Vaccine in a Humanized Mouse Model

Jie Wang[1,2], Jie Mi[2], Yan Liang[2], Xueqiong Wu[2], Junxian Zhang[2], Yinping Liu[2], Lan Wang[2], Yong Xue[2], Yingchang Shi[2], Wenping Gong[2,*] and Xinru Wang[1,3,*]

Abstract

Objective: This study was aimed at screening differentially expressed genes (DEGs) and exploring the potential immune mechanism induced by the Bacillus Calmette-Guerin (BCG) vaccine in a humanized mouse model.

Methods: Candidate DEGs between mice vaccinated with BCG or injected with PBS were identified through transcriptomics, and their biological functions, signaling pathways, and protein interaction networks were analyzed through bioinformatics.

Results: A total of 1035 DEGs were identified by transcriptomics: 398 up-regulated and 637 down-regulated. GO analysis indicated that these DEGs were significantly enriched in cell adhesion, oxygen transport, receptor complex, carbohydrate binding, serine-type endopeptidase activity, and peroxidase activity terms. KEGG analysis indicated that these DEGs were involved in the Rap1 signaling pathway, axon guidance, PI3K-Akt signaling pathway, natural killer cell mediated cytotoxicity, and cytokine-cytokine receptor interaction. Protein interaction network analysis demonstrated that the Myc, Vegfa, and Itgb3 proteins had the highest aggregation degree, aggregation coefficient, and connectivity.

Conclusion: The BCG vaccine induced 1035 DEGs in humanized mice. Among them, the differentially expressed down-regulated genes *myc* and *itgb3* involved in the PI3K-Akt signaling pathway may play essential roles in the immune mechanism of the BCG vaccine.

Key words: BCG, *Mycobacterium tuberculosis*, Transcriptomics, differentially expressed genes (DEGs), Bioinformatics

Edited by:
Zhe Wang, Department of Animal Science, Shanghai Jiaotong University

Reviewed by:
Reviewer 1, Hao Li, College of Veterinary Medicine, China Agricultural University, China
Reviewer 2, Vijay, Soni, Weill Cornell Medicine, USA
Reviewer 3 chose to be anonymous

***Corresponding authors:**
E-mail: wangxinru@126.com (XW), gwp891015@whu.edu.cn (WG)

[1]The Postgraduate Training Base of Jinzhou Medical University (The PLA Rocket Force Characteristic Medical Center), Beijing 100088, China
[2]Tuberculosis Prevention and Control Key Laboratory/Beijing Key Laboratory of New Techniques of Tuberculosis Diagnosis and Treatment, Senior Department of Tuberculosis, The 8th Medical Center of PLA General Hospital, Beijing 100091, China
[3]Department of Clinical Laboratory, The PLA Rocket Force Characteristic Medical Center, Beijing 100088, China

Received: September 09 2022
Revised: October 17 2022
Accepted: December 10 2022
Published Online: January 03 2023

INTRODUCTION

Tuberculosis (TB) is a respiratory infectious disease caused by *Mycobacterium tuberculosis* (*M. tuberculosis*). According to the Global Tuberculosis Report 2021 released by the World Health Organization, 9.87 million new TB cases occurred worldwide in 2020, among which China accounted for 8.5%, ranking second among countries [1]. China has a high burden of TB. Studies have shown that vaccination is one of the most effective ways to prevent and control TB [2,3]. Bacillus Calmette-Guerin (BCG) is a live attenuated vaccine of *Mycobacterium bovis* (*M. bovis*). It is the only vaccine approved for TB prevention and has been used for more than 100 years [4]. The BCG vaccine has a particularly protective effect against severe tuberculosis and miliary tuberculosis in children [5], and has been found to decrease infant mortality by 50% [6]. However, the protective effect of BCG lasts approximately 10–20 years [7], and BCG vaccination

has a defensive efficiency of 0% to 80% in adults [2,8]. Furthermore, growing epidemiological evidence demonstrates that BCG vaccination can induce trained immunity, which can confer non-specific protection against bacterial or viral infections with pathogens other than *M. tuberculosis* [8,9].

The BCG vaccine induces non-specific immunity by triggering innate immune cells (such as monocytes and natural killer cells), thereby enhancing metabolic activity and up-regulating proinflammatory cytokines (IL-1β, TNF-α, and IL-6), thus resulting in mounting of a faster and stronger immune response in the second infection [8,10]. As an old vaccine with a history of 100 years, BCG has saved countless lives worldwide, yet the molecular mechanism of immune protection induced by BCG remains unclear. In a previous study, we established a complete transcriptome analysis method by studying the MP3RT vaccine [11]. We constructed a Chinese population-specific humanized C57BL/6 mouse whose genetic background markedly differs from that of wild-type mice. This animal model enhances the accuracy of prediction and aids in evaluating the immune response of the Chinese population. Moreover, it simulates a human immune response similar to that observed in individuals vaccinated with BCG [12]. To elucidate the molecular mechanism of immune protection induced by the BCG vaccine, we used transcriptomic technology to determine the differential gene expression profiles of humanized mice immunized with the BCG vaccine. The differentially expressed genes (DEGs) screened by transcriptomic analysis were further subjected to Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) signal pathway enrichment, and protein interaction network analyses. This study provides a theoretical basis for exploring the molecular mechanism of immune protection induced by the BCG vaccine.

MATERIALS AND METHODS

Main materials, reagents, and instruments

Seven to eight week old female humanized C57BL/6 mice (HLA-A11$^{+/+}$DRB1*01:01$^{+/+}$H-2-β2m$^{-/-}$/IAβ$^{-/-}$) were donated by Professor Yusen Zhou from the Beijing Institute of Microbiology and Epidemiology (Beijing, China). RPMI 1640 medium was purchased from Gibco (USA). Mouse peripheral blood lymphocyte isolation solution was purchased from Tianjin Haoyang Biological Products Technology Co., Ltd. TRIzol® Reagent was purchased from Shanghai ShineGene Molecular Biotechnology Co., Ltd.

Experimental methods

Medical ethics

All animal-associated experiments were performed according to the principles of the Experimental Animal Regulation Ordinances formulated by the China National Science and Technology Commission. The mice were well cared for during the feeding process. All protocols were approved by the animal ethics committee of the 8th Medical Center of PLA General Hospital (approval No. 309201808171015). All animals were raised in the SPF laboratory of the 8th Medical Center of PLA General Hospital. The *M. tuberculosis* strain challenge test was conducted in a qualified negative pressure biosafety level II laboratory.

BCG vaccine immunization and M. tuberculosis infection in humanized mice

We randomly divided humanized mice into two groups (ten per group). Mice in the BCG vaccine (Danish strain 1331) group were subcutaneously immunized with 30 μg BCG dissolved in 100 μl phosphate buffered saline (PBS), and mice in the PBS negative control group were subcutaneously inoculated with 100 μl PBS. On day 56 after immunization, each mouse was injected with 2×10^5 CFUs of *M. tuberculosis* H37Rv strain through the tail vein.

Peripheral blood mononuclear cell (PBMC) extraction

On the 91st day after immunization, each mouse was anesthetized with diethyl ether, and 2 ml blood was collected by enucleation of the eyeball. Subsequently, the mice were sacrificed by cervical dislocation. Blood samples were placed in purple EDTA anticoagulant tubes. PBMCs were extracted from the collected blood samples with mouse peripheral blood lymphocyte separation solution. Briefly, each mouse blood sample was slowly added into a sterile centrifuge tube containing lymphocyte separation solution at a volume of 1:1 to form a clear interface and centrifuged at 2500 rpm for 25 min at room temperature. Subsequently, the cloudy PBMC layer was aspirated and transferred to a clean centrifuge tube with 5 ml of RMPI 1640 culture solution. The supernatant was aspirated and discarded, and samples were centrifuged at 1500 rpm for 10 min at room temperature. The supernatant was discarded, 5 ml of RMPI 1640 medium preheated to 37°C was added, and the precipitated cells were resuspended and centrifuged at 1500 rpm for 10 min at room temperature. Again, the supernatant was discarded, PBMCs were collected, and the number of PBMCs was quantified at 5×10^6 per mouse for later use.

RNA extraction

One milliliter of TRIzol and 200 μl chloroform were added to the above-prepared mouse PBMCs to extract RNA. The mixture of PBMCs, TRIzol, and chloroform was shaken and mixed well. After 15 minutes, the mixture was centrifuged at 12000 g for 15 minutes at 4°C. Next, the supernatant was transferred to a new centrifuge tube and mixed with 750 μl of isopropanol. Ten minutes later, the mixture was centrifuged at 12000 g and 4°C for 10 min. Subsequently, the supernatant was gently aspirated, and the precipitate was resuspended with 1 ml of 75% ethanol and centrifuged at 7500 g for 10 minutes (4°C). The

supernatant was discarded to obtain an RNA precipitate. Ten microliters of DNase/RNase-free deionized water (DEPC, TIANGEN®, Cat # RT121-02, Beijing, China) was added to dissolve RNA for subsequent experiments.

RNA quantification and library preparation for transcriptome sequencing

RNA concentration and purity were determined with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE). RNA integrity was detected with an RNA Nano 6000 Assay Kit and Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). After the sample was qualified, the library was constructed. Qubit2.0 was used for preliminary quantification, and Agilent 2100 was used to detect the insert size of the library. Subsequent experiments were performed if the insert size met expectations. The effective concentration of the library was accurately quantified by Q-PCR (and was found to be > 2 nm), and the library inspection was completed. The qualified samples (4 and 5 in PBS and BCG groups, respectively) were selected for sequencing on the Illumina NovaSeq platform. Each sample's sequencing output was not less than 6 Gb of clean data, and the percentage of Q30 bases (bases with clean data quality value greater than or equal to 30) reached 85%.

Bioinformatics analysis

Clean data were obtained by filtering the above sequenced raw data, and sequence alignment was performed between the cleaned data and the mouse reference genome (Genome assembly: GRCm38; ftp://ftp.ensembl.org/pub/release-95/fasta/mus_musculus/). The length and randomness tests were used to evaluate the library quality of the obtained mapped data, and the quantitative gene expression analysis was completed. Subsequently, EdgeR software (http://bioconductor.org) was used for differential expression screening, and a volcano plot of DEGs was drawn. The resulting *P* values were adjusted with Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted *P*-value < 0.05 and absolute fold change >1.5, determined in EdgeR were considered differentially expressed. GO and KEGG enrichment analyses were performed on DEGs in clusterProfiler software. Protein interaction network analysis of DEGs was performed in BLAST and Cytoscape software.

Statistical analysis

EdgeR software was used to analyze the gene expression profile of humanized mice. The fold change represents the expression ratio between two sample groups, and the *P* value indicates the significance of differential expression. DEGs were screened according to fold change >1.5 and *P* value < 0.05. The R package clusterProfiler was used to analyze the KEGG pathway enrichment of DEGs. A hypergeometric test was used for enrichment analysis to identify KEGG pathways that were significantly enriched with respect to the whole genome background. The fragments per kilobase per million (FPKM) algorithm was used to normalize the gene expression levels in each sample according to the following formula:

$$\mathrm{KPKM} = \frac{\text{cDNA Fragments}}{\text{Mapped Fragments (Millions)} \times \text{Transcript Length (kb)}}$$

where cDNA Fragments represents the number of fragments aligned to a transcript, that is, the number of double-ended reads; Mapped Fragments (Millions) represents the total number of fragments aligned to a transcript, in millions; and Transcript Length (kb) indicates the length of the transcript.

RESULTS

Identification of DEGs induced by the BCG vaccine

According to the above RNA sample test results, we used four samples in the PBS group and five in the BCG group for subsequent transcriptome sequencing and DEGs analysis in the transcriptome sequencing test. Transcriptomic sequencing data of the two groups were analyzed, and cluster heat maps (Fig 1A) and volcano plots (Fig 1B) were drawn. These four samples in the PBS group showed similar clustering characteristics, whereas the five samples in the BCG group showed similar characteristics. In this cluster heat map, redder color indicates greater gene up-regulation, and greener color indicates greater gene down-regulation. A total of 1035 genes were differentially expressed in the BCG vaccine group compared with the PBS group, of which 398 were significantly up-regulated, and 637 were significantly down-regulated.

GO analysis and KEGG enrichment analysis

GO and KEGG enrichment analyses were performed on DEGs of humanized mice induced by the BCG vaccine. First, DEGs were sorted according to the enrichment in GO function from low to high, and the top ten GO terms according to enrichment score were screened. The results (Fig 2) indicated the following: (1) Biological processes (BP): DEGs induced by the BCG vaccine in humanized mice were significantly enriched in cell adhesion, cytolysis, oxygen transport, platelet aggregation, myofilament gliding, negative regulation of angiogenesis, and negative regulation of viral genome replication. (2) Cellular component (CC): DEGs induced by the BCG vaccine were enriched mainly in cortical cytoskeleton, lateral plasma membrane, receptor complex, platelet α, neuron cell body, cell basement membrane, and membrane anchor component functions. (3) Molecular function (MF): DEGs were enriched mainly in proteins involved in metabolism, such as oxygen transport activity, carbohydrate binding, serine-type endopeptidase activity, peroxidase activity, collagen binding, fibrin binding, oxygen binding, and actin-dependent ATPase activity.

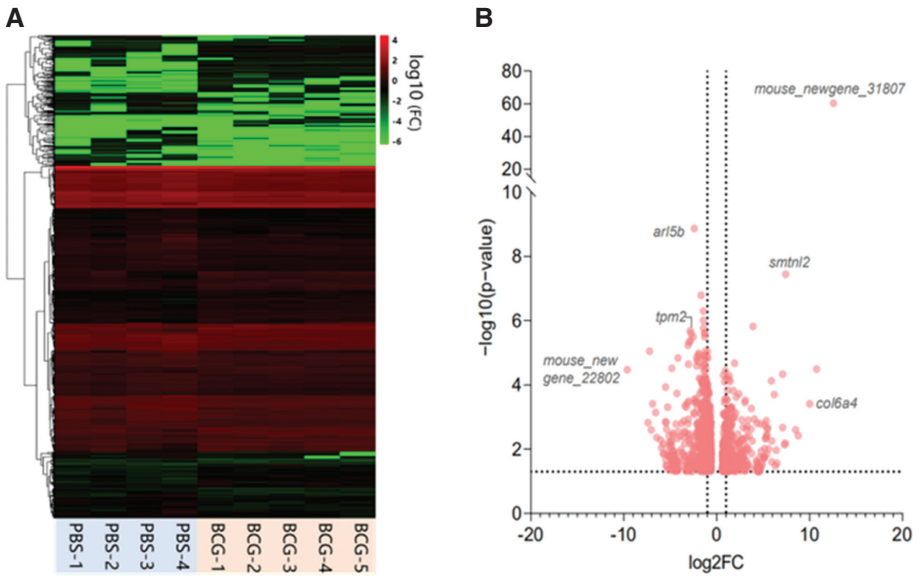

FIGURE 1 | Cluster heat map and volcano plot of DEGs between the BCG and PBS groups. A. The abscissa represents the sample name and the clustering result of the sample, and the ordinate represents the clustering result of DEGs and genes. Different columns in the figure represent different samples, and different rows represent different genes. The color represents the level of gene expression $\log_{10}$ (gene+0.000001) in the sample. Redder color indicates greater up-regulation, and greener color indicates greater down-regulation. PBS-1, PBS-2, PBS-3, and PBS-4, respectively, represent four mice in the control group, and BCG-1, BCG-2, BCG-3, BCG-4, and BCG-5, respectively, represent five mice in the experimental group. B. Each point in the differential expression volcano plot represents a gene, and the abscissa represents the ratio of the expression levels of a gene in two samples, expressed as $\log_2$ (FC); The ordinate represents the negative logarithm of the *P* value. In the figure, genes with $\log_2$ FC greater than 1 were differentially up-regulated, whereas genes with $\log_2$ FC less than -1 were differentially down-regulated.

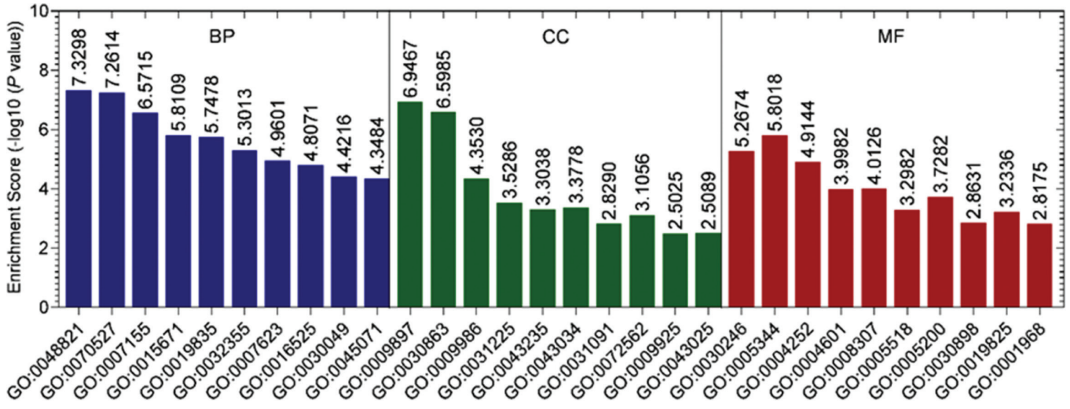

FIGURE 2 | GO function analysis of the DEGs between BCG and PBS groups. The enrichment scores of DEGs were subjected to GO functional enrichment analysis. The abscissa represents the GO node name, and the ordinate represents the DEG enrichment score -$\log_{10}$ (*P* value); *P* value represents the significance of GO term enrichment in the DEG table. The smaller the *P* value, the more significant the GO term ($P \leq 0.05$). BP: biological process; CC: cell component; MF: molecular function.

The DEGs with these significantly enriched GO terms were ranked according to *P* value from low to high, and the top 20 genes with the most significant differential expression are listed in Table 1 (complete information in S1 Table). Furthermore, KEGG pathway enrichment analysis of DEGs induced by the BCG vaccine indicated that DEGs were enriched mainly in the Rap1 signaling pathway, axonal guidance, proteoglycan in cancer, PI3K-Akt signaling pathway, natural killer cell-mediated cytotoxicity, focal adhesions, cytokine/cytokine receptor interaction, platelet activation, hematopoietic cell lineage, and other pathways. Up-regulated and down-regulated genes were enriched in the 20 signaling pathways shown in Fig 3. We identified 29 DEGs enriched in the PI3K-Akt signaling pathway. Ten DEGs were up-regulated, and 19 genes were down-regulated, among which *itgb3*, *itga2b*, and *thbs1* were significantly down-regulated (Fig 3, Table 2; complete information in S2 Table).

TABLE 1 | Characteristics of DEGs in the BCG group compared with the PBS group.

Gene	FPKM		*P* value	log_2FC	Regulation direction
	PBS	BCG			
tpm2	14.524822	2.4156622	2.09E-06	–2.86349	Down
gp1ba	42.315465	19.4307658	2.43E-06	–1.31836	Down
clec1b	196.5148998	97.1250564	2.46E-06	–1.28428	Down
mmp19	19.99045525	11.4687296	9.37E-06	–0.96642	Down
cald1	31.497439	13.7299974	1.22E-05	–1.44641	Down
slc6a4	61.15986275	27.4963346	1.30E-05	–1.36066	Down
atf5	16.11523425	4.3661804	1.38E-05	–2.15761	Down
gp6	16.03238325	10.6185854	1.46E-05	–0.91712	Down
ache	8.609503	3.9799398	1.54E-05	–1.33381	Down
ltga2b	450.203518	209.9541044	2.27E-05	–1.29837	Down
gp1bb	272.5843173	72.6678322	2.54E-05	–2.15732	Down
gp9	370.835926	151.3087954	2.76E-05	–1.50929	Down
egr1	35.76184325	75.5925798	3.58E-05	0.937227	Up
thbs1	142.4034135	70.0341186	6.10E-05	–1.23274	Down
treml1	103.8206915	50.7555154	7.77E-05	–1.23267	Down
proz	0.65986	1.2278496	7.81E-05	1.447331	Up
myl9	424.2745668	207.663745	7.93E-05	–1.24547	Down
vcan	21.8476275	9.4080434	8.65E-05	–1.51915	Down
gata1	18.70625425	10.9283998	9.79E-05	–0.96089	Down
ltgb3	144.037958	74.754024	0.000122	–1.15785	Down

The enrichment scores of DEGs were analyzed with GO functional enrichment; the abscissa shows the GO node name, and the ordinate shows the DEG enrichment score $-\log_{10}$ (*P* value). The *P* value represents the significance of GO term enrichment in the DEG table. The smaller the *P* value, the more significant the GO term ($P \leq 0.05$). BP: biological process; CC: cell component; MF: molecular function.

Each row in the figure represents a KEGG pathway. The abscissa is the enrichment factor, which represents the ratio of the proportion of genes annotated to the pathway among DEGs to the proportion of genes annotated to the pathway among all genes. The greater the enrichment factor, the more significant the DEGs' enrichment in the pathway. The color of the dot represents the Q value, the dot size represents the number of DEGs annotated in the pathway, and the shape of the dot represents the type of DEG in the pathway (for example, ○ contains up-regulated and down-regulated genes).

Protein interaction network analysis of DEGs

To further analyze the relationships among proteins encoded by DEGs, we performed protein interaction network analysis of DEGs. We used BLAST software to sequence and align target genes with proteins in the STRING database to identify homologous proteins. The interaction network was constructed according to the interaction relationships among homologous proteins. As shown in Fig 4, the proteins Myc, Vegfa, and Itgb3 had the largest aggregation degree and the highest aggregation coefficient, thereby indicating the highest connectivity. The genes *myc* and *itgb3* were differentially down-regulated in transcriptome sequencing data, and *vegfa* was differentially up-regulated. The gene *myc* was significantly down-regulated in proteoglycans in cancer and the PI3K-Akt signaling pathway. In addition to these two pathways, the differentially expressed up-regulated gene *vegfa* participated in cytokine receptor interaction, focal adhesion, and the Rap1 signaling pathway. The differentially expressed down-regulated gene *itgb3* is involved in the regulation of ten signal pathways shown in Table 2.

The nodes in the figure are proteins, and the edges are interaction relationships. The size of a node in an interaction network is proportional to the degree of the node. The more edges connected to the node, the greater the degree, and the larger the node. The color of a node is associated with its clustering coefficient, and the color gradient from blue to red corresponds to the clustering

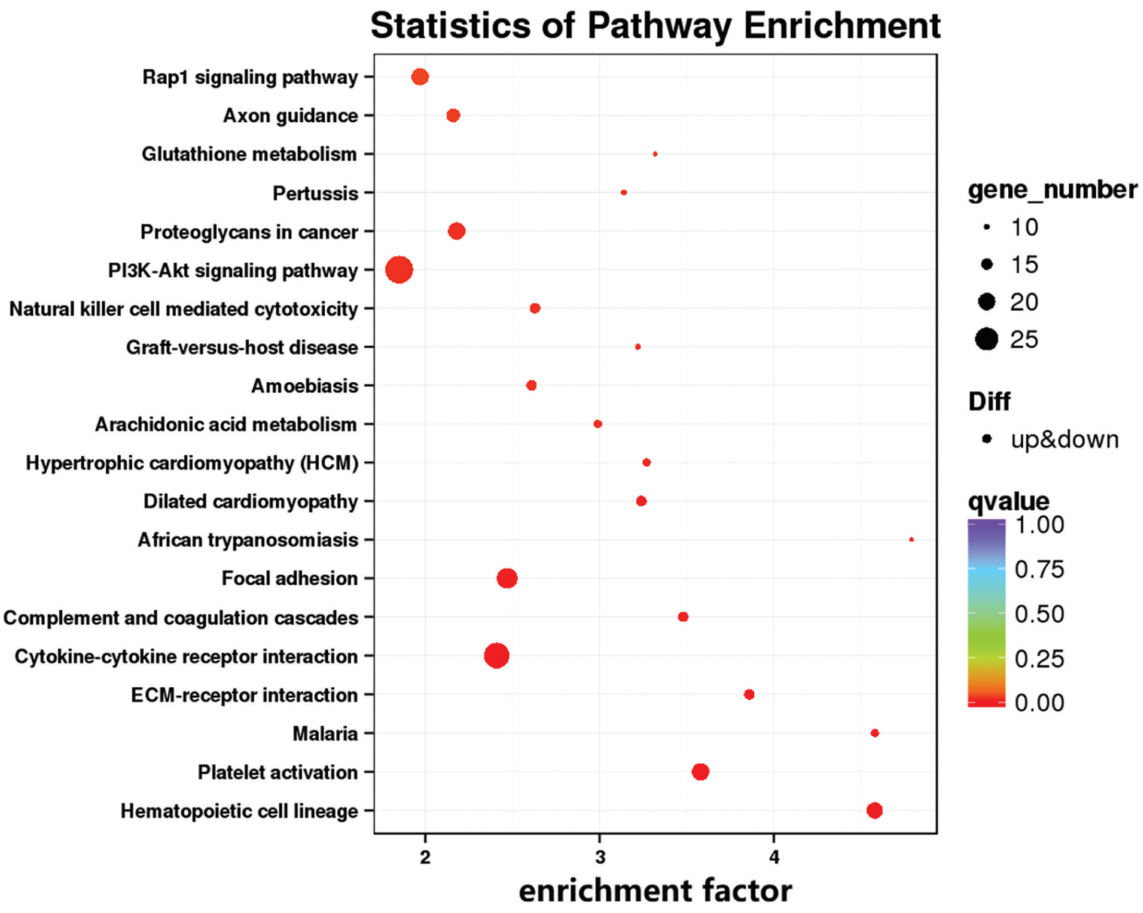

FIGURE 3 | KEGG pathway analysis of the DEGs between the BCG and PBS groups. Each row in the figure represents a KEGG pathway. The abscissa is the enrichment factor, which represents the ratio of the proportion of genes annotated to the pathway in DEGs to the proportion of genes annotated to the pathway in all genes. The larger the enrichment factor, the more significant the enrichment level of DEGs in the pathway. The color of the dot represents the Q value, the dot size represents the number of DEGs annotated in the pathway, and the shape of the dot represents the type of DEG in the pathway (for example, ○ contains up-regulated and down-regulated genes).

coefficient from low to high. The aggregation coefficient indicates the connectivity between the neighboring points of this node. The higher the aggregation coefficient value, the better the connectivity between the neighboring points of the node.

DISCUSSION

To further explore the molecular mechanism of immune protection induced by the BCG vaccine, we used transcriptomics and bioinformatics techniques to establish differential gene expression profiles at the transcriptome level in humanized mice induced by the BCG vaccine. Consequently, 1035 genes were found to be differentially expressed in the BCG vaccine immunized group compared with the PBS group, of which 398 were up-regulated, and 637 were down-regulated. Among them, the differentially expressed down-regulated genes *myc* and *itgb3*, and the differentially expressed up-regulated gene *vegfa* may play essential roles in the molecular immune mechanism of the BCG vaccine.

Studies have shown that *myc* and *egr1* may regulate the response of macrophages to *M. bovis* by regulating the host cell death mechanism. *M. bovis* infection inhibits the expression of *myc* and suppresses the host proinflammatory response, thus leading to the survival of bacteria in the cells [13]. Other studies have found that the gene *myc*, encoding a nuclear phosphoprotein, plays roles in cell cycle progression, apoptosis, and cell transformation, and is highly expressed in most tumors [14]. The up-regulated gene *myc* activates the PI3K-Akt signaling pathway and promotes the proliferation of breast cancer cells [15]. The gene *egr1* encodes a zinc finger transcription factor that regulates gene expression in cell differentiation. Egr1 also promotes the transcription of proinflammatory cytokines in mouse macrophages in response to stimulation by bacterial antigens [13]. In this study, the expression of the *egr1* gene induced by the BCG vaccine was up-regulated, and the expression of the *myc* gene was down-regulated in the PI3K-Akt signaling pathway. The protein Myc showed the greatest aggregation degree, the highest aggregation coefficient, and the highest connectivity in the protein interaction network. The role of the *myc* gene in the molecular mechanism of immune protection of the BCG vaccine requires further study.

TABLE 2 | Signaling pathways associated with DEGs.

Pathway ID	Description	Enrichment factor	*P* value	Gene symbols
ko04640	Hematopoietic cell lineage	4.58	1.93446E-08	*kit, epor, cd34, csf2, itgb3, tfrc, tnf, itga6, il1b, gp9, cd59a, itga2b, il7, cd24a, gp5, gp1ba, gp1bb, cd59b*
ko04611	Platelet activation	3.58	5.89012E-07	*col1a1, vwf, itgb3, p2rx1, mapk12, adcy5, fermt3, p2ry1, tbxas1, gp9, gucy1a1, itga2b, p2ry12, ptgs1, gp5, gp1ba, gp1bb, gp6, pla2g4b*
ko05144	Malaria	4.58	8.38587E-06	*tnf, selp, il1b, il12a, vcam1, thbs1, gypa, hbb-bs, hba-a2, hba-a1, hbb-bt,*
ko04512	ECM-receptor interaction	3.86	1.20688E-05	*col1a1, vwf, itgb3, itgb4, itgb6, itga6, gp9, col6a4, itga2b, thbs1, gp5, gp1ba, gp1bb, gp6*
ko04060	Cytokine-cytokine receptor interaction	2.41	1.88934E-05	*pdgfb, ccl3, kit, epor, mpl, ccl8, il12rb2, csf2, ccl4, tnfrsf10b, vegfa, tnf, il23a, il24, il1b, il12a, pdgfc, ppbp, pf4, cxcl10, ccl5, ccr3, il7, ccr4, cx3cr1, il2rb, gm21970*
ko04610	Complement and coagulation cascades	3.48	4.02278E-05	*vwf, f2rl2, clu, pros1, serping1, f5, itgax, cd59a, c1qb, c1s1, f13a1, c3ar1, cd59b*
ko04510	Focal adhesion	2.47	5.16081E-05	*pdgfb, col1a1, rac1, vwf, pgf, actn1, rac3, igf1, itgb3, itgb4, vcl, parvb, vegfa, itgb6, itga6, pdgfc, ilk, col6a4, itga2b, thbs1, erbb2, myl9*
ko05143	African trypanosomiasis	4.79	7.87567E-05	*tnf, il1b, il12a, vcam1, hbb-bs, hba-a2, hba-a1, hbb-bt*
ko05414	Dilated cardiomyopathy	3.24	9.19381E-05	*igf1, itgb3, cacng4, itgb4, adcy5, tnf, itgb6, itga6, lmna, tpm2, itga2b, tnnc1*
ko05410	Hypertrophic cardiomyopathy (HCM)	3.27	0.000263293	*igf1, itgb3, cacng4, itgb4, tnf, itgb6, itga6, lmna, tpm2, itga2b, tnnc1, mouse_newgene_19955*
ko00590	Arachidonic acid metabolism	2.99	0.000617838	*alox12, ggt1, gpx8, pla2g10, tbxas1, cyp2b10, ptgs2, ptgs1, cyp2j6, gpx1, cyp4a12b, pla2g4b*
ko05205	Proteoglycans in cancer	2.18	0.000829125	*col1a1, rac1, igf1, itgb3, ctsl, myc, mapk12, wnt10b, vegfa, tnf, wnt10a, cttn, ank1, mras, gpc1, thbs1, camk2b, erbb2, ank3, mouse_newgene_19955*
ko04650	Natural killer cell mediated cytotoxicity	2.63	0.00083357	*rac1, gzmb, rac3, csf2, tnfrsf10b, tnf, klrk1, klrd1, klrc1, klrb1c, prf1, klrc2, ncr1, mouse_newgene_3077*
ko05146	Amoebiasis	2.61	0.000904266	*col1a1, actn1, csf2, arg1, nos2, serpinb9b, vcl, tnf, il1b, il12a, serpinb6b, serpinb9, serpinb6a, mouse_newgene_3077*
ko05332	Graft-versus-host disease	3.22	0.00095095	*gzmb, tnf, il1b, klrd1, klrc1, prf1, h2-q10, mouse_newgene_10403, mouse_newgene_10515, mouse_newgene_5721*
ko04151	PI3K-Akt signaling pathway	1.85	0.000960549	*pdgfb, col1a1, rac1, vwf, efna2, pgf, kit, epor, bcl2l1, igf1, fgf22, itgb3, gh, itgb4, myc, nr4a1, vegfa, itgb6, itga6, pdgfc, col6a4, gng11, itga2b, thbs1, il7, gng7, irs1, il2rb, mouse_newgene_3077*
ko05133	Pertussis	3.14	0.001184403	*nos2, mapk12, serping1, tnf, il23a, il1b, il12a, nlrp3, c1qb, c1s1*
ko00480	Glutathione metabolism	3.32	0.0013635	*gstt1, gstm5, ggt1, gpx8, mgst3, gclm, gsta2, gpx1, gpx4*
ko04360	Axon guidance	2.16	0.002190738	*sema6b, rac1, efna2, ephb3, rac3, f2rl2, sema4c, sema6d, sema3a, dpysl5, plxna4, srgap3, ilk, trpc6, trpc1, sema3d, camk2b*
ko04015	Rap1 signaling pathway	1.97	0.002769846	*pdgfb, rac1, efna2, pgf, kit, rac3, igf1, fgf22, itgb3, mapk12, adcy5, vegfa, rassf5, p2ry1, pdgfc, mras, itga2b, thbs1, mouse_newgene_19955, mouse_newgene_5721*

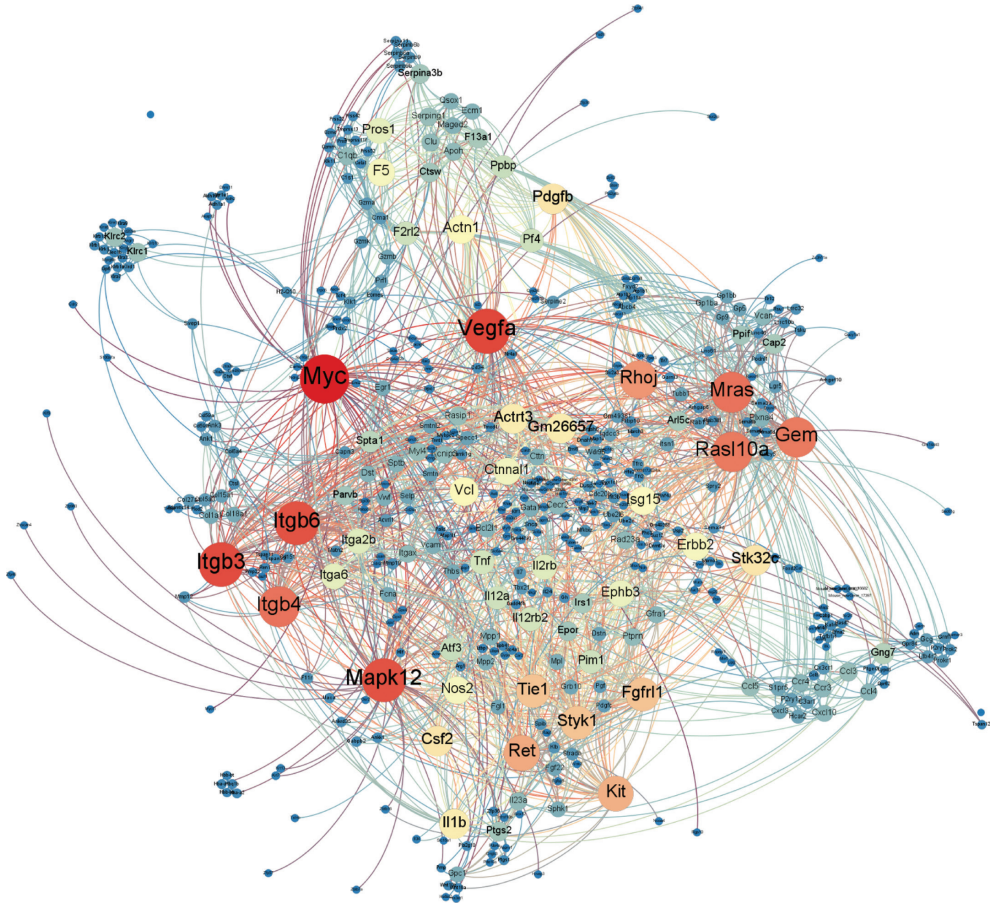

FIGURE 4 | Protein interaction networks of the DEGs. The nodes in the figure are proteins, and the edges are interaction relationships. The size of a node in an interaction network is proportional to the degree of the node. The more edges connected to the node, the greater its degree, and the larger the node. The color of a node is associated with its clustering coefficient, and the color gradient from blue to red corresponds to the clustering coefficient from low to high. The aggregation coefficient indicates the connectivity between the neighboring points of this node. The higher the aggregation coefficient value, the better the connectivity between the neighboring points of the node.

Itgb3 also known as CD61, is a critical member of the integrin family. As an adhesion receptor on the cell surface, it participates in tumor adhesion and angiogenesis [16]. Studies have found that Itgb3 promotes the clearance of *M. tuberculosis* in endothelial cells by altering phagosome transport; moreover, Itgb3 agonist treatment decreases the bacterial load *in vivo* [17], and *itgb3* gene up-regulation activates FAK/PI3K/Akt signaling, thereby promoting the proliferation and invasion of nasopharyngeal carcinoma cells [18]. In this study, the gene *itgb3* in the phagosome pathway and PI3K-Akt signaling pathway showed down-regulation by the BCG vaccine. In the protein interaction network analysis, the protein Itgb3 showed the most considerable aggregation degree, the highest aggregation coefficient, and the highest connectivity. In this study, the expression of the *itgb3* gene in the phagosome pathway induced by the BCG vaccine was down-regulated by 1.1 fold, in contrast to the conclusion by Che et al. that the up-regulation of *itgb3* promotes the elimination of tuberculosis bacteria [16]. Interestingly, our study also indicated that, in the phagosome pathway, the expression of the *itgb6* gene was up-regulated by four fold, and the expression of the *itgb4* gene was down-regulated by 0.9 fold. Therefore, the ITGB branch in the phagosome pathway induced by the BCG vaccine was generally up-regulated. The heterogeneity in the above results might have been due to differences in species and the timing of immunization. The potential influence of the ITGB family on the immunological protective mechanism induced by BCG vaccination should be further investigated in future studies.

For other DEGs with high rank, we identified no tuberculosis-associated research results in the literature. We briefly discuss research on these genes in other diseases. Vascular endothelial growth factor A (VEGFA) is a major player in tumor angiogenesis, and the gene *vegfa* is highly expressed in Th cells in patients with non-small cell lung cancer [19]. In this study, the BCG-induced gene *vegfa* was differentially expressed and up-regulated in several pathways, including proteoglycans in cancer, the Rap1 signaling pathway, and cytokine-cytokine receptor interaction. The matrix metalloproteinase (MMP) family members play crucial roles in the host immune response to tuberculosis. Among many MMPs, MMP-1, MMP-2, MMP-3, MMP-8, and MMP-9 are the most widely studied isoforms.

Some studies have found that inhibiting the expression of MMP-1/9 inhibits *M. tuberculosis* infection [20]. However, among the DEGs induced by BCG screened in this study, we observed no differential expression of the above five isoforms. Instead, the genes *mmp12* and *mmp19* were down-regulated. Phosphoglycerate kinase 1, encoded by the gene *mmp19*, is highly expressed in glioma, and its expression level is positively correlated with the clinical stage of the tumor [21]. In contrast, mmp19 was differentially down-regulated in this study. Tropomyosin 2, encoded by the *tpm2* gene, has low expression in atherosclerosis samples and may serve as a diagnostic marker for atherosclerosis [22]. In this study, the differentially expressed down-regulated gene *tpm2* was enriched in the muscle's structural components, the cytoskeleton, and myofilament sliding entry and was found to participate in signaling pathways such as myocardial contraction and adrenergic signaling in cardiomyocytes. Proteins encoded by genes *gp1ba*, *gp1bb*, *gp9*, and *gp5* are components of the platelet surface glycoprotein Ib-IX-V complex, and their deletion can lead to Bernard-Soulier syndrome [23]. In this study, these four genes were down-regulated and found to be involved in the platelet activation pathway, among which the gene *gp1ba* was enriched in platelet aggregation terms. *Clec1b* is a characteristic gene highly associated with tumor progression, and its expression is down-regulated in the progression stage of hepatocellular carcinoma. It may be a target for tumor clinical diagnosis or immunotherapy [24]. In this study, the differential expression of *clec1b* induced by BCG was down-regulated, and its role in the immune protection mechanism of BCG requires further study. Studies have found that calmodulin-binding protein (Cald1) is associated with a variety of cancers, and its low expression in ovarian cancer tissues enhances the migration rate of cancer cells and promotes the metastasis of cancer cells [25]. The gene *cald1* may serve as a potential target in diagnosing and treating ovarian cancer. The role of the differentially expressed down-regulated gene *cald1* in the molecular immune mechanism of the BCG vaccine must be further studied. Methylation of the serotonin transporter gene *slc6a4* contributes to an increased risk of post-stroke depression [26]. In this study, the differentially down-regulated gene *slc6a4* was enriched in circadian rhythm function.

We searched the literature associated with BCG post-immunization transcriptomic analysis. The chemotactic factors CCL3 and CCL4 have been found to be down-regulated in transcriptomic analysis of human monocytes induced by BCG vaccination; the chemotactic factor CCL2 is up-regulated 3 months later; and interferon-induced antiviral gene *ifitm1* is down-regulated [27]. In contrast, in this study, genes *ccl3* and *ccl4* in humanized mice induced by BCG were up-regulated, *ccl2* was not differentially expressed, and the *ifitm1* gene was down-regulated. These results suggest that BCG-induced immune responses differ between humans and humanized mice, owing to the differences in species and the timing of immunization. Furthermore, other studies have found that BCG-infected human macrophages increase TNF-α [28], in agreement with our results.

This study has several limitations. First, the transcriptomic analysis was conducted in an animal model rather than in samples collected from humans, thus resulting in heterogeneity between our results and previous literature data, owing to species differences. Second, the transcriptome analysis was performed on only samples obtained on day 91 post-infection; therefore, our experimental results are limited to a single time point and cannot fully reflect the full range of immune responses after BCG vaccination.

CONCLUSION

In summary, transcriptomic analysis of 1035 DEGs induced by the BCG vaccine in a humanized mouse model revealed that the differentially down-regulated genes *myc* and *itgb3* are involved in the PI3K-Akt signaling pathway. These genes might play essential roles in the molecular mechanism of immune protection of the BCG vaccine, a possibility that must be further verified. This study may indicate novel biomarkers for the diagnosis and treatment of TB. In addition, our results may provide a theoretical basis for further exploring the immune mechanism induced by BCG vaccination.

ACKNOWLEDGEMENTS

This study was funded by the Beijing Municipal Science & Technology Commission (grant No. L192065), National Natural Science Foundation of China (grant No. 81801643), and the National key research and development program of China (grant No. 2022YFA1303500-003).

CONFLICTS OF INTEREST

The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

REFERENCES

1. WHO. Global tuberculosis report 2021. in Genevapp: World Health Organization. 2021: Geneva.
2. Gong W, Liang Y, Wu X. The current status, challenges, and future developments of new tuberculosis vaccines. Hum Vaccin Immunother. 2018;14(7):1697-1716.
3. Schrager LK, Vekemens J, Drager N, Lewinsohn DM, Olesen OF. The status of tuberculosis vaccine development. Lancet Infect Dis. 2020;20(3):e28-e37.
4. Rümke HC. BCG: an almost 100-year-old vaccine. Ned Tijdschr Geneeskd. 2020;164:D5146.
5. Hussey G, Hawkridge T, Hanekom W. Childhood tuberculosis: old and new vaccines. Paediatr Respir Rev. 2007;8(2):148-154.
6. O'Neill LAJ, Netea MG. BCG-induced trained immunity: can it offer protection against COVID-19? Nat Rev Immunol. 2020;20(6):335-337.
7. Fatima S, Kumari A, Das G, Dwivedi VP. Tuberculosis vaccine: a journey from BCG to present. Life Sci. 2020;252:117594.
8. Wang J, Zhang Q, Wang H, Gong W. The potential roles of BCG vaccine in the prevention or treatment of COVID-19. Front Biosci (Landmark Ed). 2022;27(5):157.

9. Gong W, An H, Wang J, Cheng P, Qi Y. The natural effect of BCG vaccination on COVID-19: the debate continues. Front Immunol. 2022;13:953228.
10. Covián C, Fernández-Fierro A, Retamal-Díaz A, Díaz FE, Vasquez AE, Lay MK, et al. BCG-induced cross-protection and development of trained immunity: implication for vaccine design. Front Immunol. 2019;10:2806.
11. Wang J, Mi J, Liang Y, Wu XQ, Zhang JX, Liu YP. Transcriptomic analysis of tuberculosis peptide-based vaccine MP3RT in humanized mice. Chinese J Tuberc Res Dis. 2022;45(9):894-903.
12. Zeng Y, Gao T, Zhao G, Jiang Y, Yang Y, Yu H, et al. Generation of human MHC (HLA-A11/DR1) transgenic mice for vaccine evaluation. Hum Vaccin Immunother. 2016;12(3):829-836.
13. Killick KE, Magee DA, Park SD, Taraktsoglou M, Browne JA, Conlon KM, et al. Key hub and bottleneck genes differentiate the macrophage response to virulent and attenuated mycobacterium bovis. Front Immunol. 2014;5:422.
14. Baluapuri A, Wolf E, Eilers M. Target gene-independent functions of MYC oncoproteins. Nat Rev Mol Cell Biol. 2020;21(5):255-267.
15. Chen Q, Shen H, Zhu X, Liu Y, Yang H, Chen H, et al. A nuclear lncRNA Linc00839 as a Myc target to promote breast cancer chemoresistance via PI3K/AKT signaling pathway. Cancer Sci. 2020;111(9):3279-3291.
16. Zhu C, Kong Z, Wang B, Cheng W, Wu A, Meng X. ITGB3/CD61: a hub modulator and target in the tumor microenvironment. Am J Transl Res. 2019;11(12):7195-7208.
17. Chen X, Cao X, Lei Y, Reheman A, Zhou W, Yang B, et al. Distinct persistence fate of mycobacterium tuberculosis in various types of cells. mSystems. 2021;6(4):e0078321.
18. Zheng ZQ, Li ZX, Zhou GQ, Lin L, Zhang LL, Lv JW, et al. Long noncoding RNA FAM225A promotes nasopharyngeal carcinoma tumorigenesis and metastasis by acting as ceRNA to sponge miR-590-3p/miR-1275 and upregulate ITGB3. Cancer Res. 2019;79(18):4612-4626.
19. Rakshit S, Sunny JS, George M, Hanna LE, Sarkar K. R-loop modulated epigenetic regulation in T helper cells mechanistically associates coronary artery disease and non-small cell lung cancer. Transl Oncol. 2021;14(10):101189.
20. Zhou X, Lie L, Liang Y, Xu H, Zhu B, Huang Y, et al. GSK-3α/β activity negatively regulates MMP-1/9 expression to suppress mycobacterium tuberculosis infection. Front Immunol. 2021;12:752466.
21. Luo Q, Luo H, Chen X, Yan P, Fu H, Huang H, et al. The expression of MMP19 and its clinical significance in glioma. Int J Clin Exp Pathol. 2018;11(11):5407-5412.
22. Meng LB, Shan MJ, Qiu Y, Qi R, Yu ZM, Guo P, et al. TPM2 as a potential predictive biomarker for atherosclerosis. Aging (Albany NY). 2019;11(17):6960-6982.
23. Özdemir ZC, Düzenli Kar Y, Ceylaner S, Bör Ö. A novel mutation in the GP1BA gene in Bernard-Soulier syndrome. Blood Coagul Fibrinolysis. 2020;31(1):83-86.
24. Hu K, Wang ZM, Li JN, Zhang S, Xiao ZF, Tao YM. CLEC1B expression and PD-L1 expression predict clinical outcome in hepatocellular carcinoma with tumor hemorrhage. Transl Oncol. 2018;11(2):552-558.
25. Boljevic I, Malisic E, Milovic Kovacevic M, Jovanic I, Jankovic R. Expression levels of genes involved in cell adhesion and motility correlate with poor clinicopathological features of epithelial ovarian cancer. J BUON. 2020;25(4):1911-1917.
26. Kang HJ, Lee EH, Kim JW, Kim SW, Shin IS, Kim JT, et al. Association of SLC6A4 methylation with long-term outcomes after stroke: focus on the interaction with suicidal ideation. Sci Rep. 2021;11(1):2710.
27. Kong L, Moorlag SJCFM, Lefkovith A, Li B, Matzaraki V, van Emst L, et al. Single-cell transcriptomic profiles reveal changes associated with BCG-induced trained immunity and protective effects in circulating monocytes. Cell Rep. 2021;37(7):110028.
28. Mahla RS, Kumar A, Tutill HJ, Krishnaji ST, Sathyamoorthy B, Noursadeghi M, et al. NIX-mediated mitophagy regulate metabolic reprogramming in phagocytic cells during mycobacterial infection. Tuberculosis (Edinb). 2021;126:102046.