# Epigenomics

# Integrated miRNA and mRNA expression profiling to identify mRNA targets of dysregulated miRNAs in pulmonary tuberculosis

Xing Zhang[1], Min Zhu[2] & Xiaolong Hu*,[2]

[1]Department of Infectious Disease, First Affiliated Hospital of Soochow University, Suzhou, Jiangsu Province, 215006, PR China
[2]Department of Applied Biology, School of Biology & Basic Medical Science, Medical College, Soochow University, Suzhou, Jiangsu Province, 215123, PR China
*Author for correspondence: xlhu2013@suda.edu.cn

**Aim:** The aim of this study was to identify mRNA targets of dysregulated miRNAs through the integrated analysis of miRNA and mRNA expression profiling in pulmonary tuberculosis (PTB) patients versus healthy individuals. **Materials & methods:** Expression profiles in blood obtained from PTB patients and healthy individuals were analyzed using high-throughput sequencing. **Results:** Forty-one differentially expressed miRNAs and 2565 mRNAs were obtained. A large number of the differentially expressed mRNAs and miRNAs were related to immune-related pathways, particularly the tuberculosis, phagosome and MAPK signaling pathway. Three hundred and fifty-nine potential target genes were identified for 41 differentially expressed miRNAs. Many of target genes were enriched to phagosome, calcium and insulin signaling pathway. **Conclusion:** The mRNA–miRNA regulatory networks described here provide new insights for further elucidation of PTB pathogenesis.

Tuberculosis (TB) is a disease with a history interwoven with the evolution and migration of human, as well as with the origins of microbiology. The main etiologic agent of TB is *Mycobacterium tuberculosis* (Mtb) [1]. Mtb infects 2 billion people around the world, causing nearly 10.4 million new active TB cases and 1.8 million deaths annually [2]. In TB pathogenesis, the host cellular immune response determines whether an infection becomes a latent TB infection or progresses to infectious active TB or extrapulmonary TB [3]. Approximately 90% of infected individuals will remain asymptomatic with latent TB infection. Only 10% of individuals infected with Mtb will develop active disease, suggesting that host immunity is important in regulating progression of TB infection [4,5].

miRNAs, a type of noncoding RNA between 22 and 24 nucleotide (nt), could be important regulators of gene expression at the post-transcriptional level and influence many biological systems including mammalian immune systems [6]. Studies found that miRNAs extensively regulate cell differentiation, development and disease [7]. In mammals, over 60% of mRNAs are thought to be regulated by miRNAs [8]. Studies have found altered gene expression profiles in macrophages and natural killer (NK) cells from individuals with active or latent TB, and individuals with TB infection or healthy individuals. This alteration of cellular composition and related gene expression in patients with TB is likely regulated by miRNAs [9,10]. Several miRNAs regulate T-cell differentiation and function in patients [11,12]. In addition, miRNAs are important for regulating the innate function of macrophages, dendritic cells and NK cells [13]. Some miRNAs in CD4$^+$ T cells are altered in latent and active TB [14]. Novel miRNA combinations can discriminate between individuals with a TB infection and healthy individuals [15,16]. Furthermore, the functions of miRNAs are closely correlated with pulmonary TB pathogenesis, their molecular regulatory mechanism has not been investigated by integrated mRNA and miRNA transcriptome analysis.

In this study, high-throughput sequencing strategy was employed to screen differentially expressed mRNA and miRNAs between individuals with pulmonary tuberculosis (PTB) and healthy individuals with DESeq analysis [17].

Future Medicine

## Table 1. The summary information of mRNA.

| Sample | Sample_C1 (%) | Sample_C2 (%) | Sample_C3 (%) | Sample_F1 (%) | Sample_F4 (%) | Sample_F5 (%) |
|---|---|---|---|---|---|---|
| Total reads | 97,902,050 | 98,099,296 | 98,047,510 | 97,193,002 | 97,321,140 | 98,293,852 |
| Total mapped | 92,770,395 (94.76) | 92,722,766 (94.52) | 92,901,009 (94.75) | 91,899,482 (94.55) | 92,173,068 (94.71) | 93,146,292 (94.76) |
| Multiple mapped | 9,592,910 (9.80) | 9,238,859 (9.42) | 10,185,832 (10.39) | 9,038,996 (9.30) | 9,248,784 (9.50) | 8,669,668 (8.82) |
| Uniquely mapped | 83,177,485 (84.96) | 83,483,907 (85.10) | 82,715,177 (84.36) | 82,860,486 (85.25) | 82,924,284 (85.21) | 84,476 624 (85.94) |
| Read-1 | 42,001,844 (42.90) | 42,266,211 (43.09) | 41,799,707 (42.63) | 41,950,109 (43.16) | 41,933,796 (43.09) | 42,668,772 (43.41) |
| Read-2 | 41,175,641 (42.06) | 41,217,696 (42.02) | 40,915,470 (41.73) | 40,910,377 (42.09) | 40,990,488 (42.12) | 41,807,852 (42.53) |
| Reads map to '+' | 41,586,110 (42.48) | 41,729,231 (42.54) | 41,351,816 (42.18) | 41,421,771 (42.62) | 41,463,938 (42.61) | 42,250,726 (42.98) |
| Reads map to '−' | 41,591,375 (42.48) | 41,754,676 (42.56) | 41,363,361 (42.19) | 41,438,715 (42.64) | 41,460,346 (42.60) | 42,225,898 (42.96) |
| Nonsplice reads | 70,230,815 (71.74) | 72,192,640 (73.59) | 71,200,659 (72.62) | 68,734,539 (70.72) | 69,134,393 (71.04) | 72,125,014 (73.38) |
| Splice reads | 12,946,670 (13.22) | 11,291,267 (11.51) | 11,514,518 (11.74) | 14,125947 (14.53) | 13,789,891 (14.17) | 12,351,610 (12.57) |
| Reads mapped in proper pairs | 79,978,208 (81.69) | 80,026,634 (81.58) | 79,700,700 (81.29) | 79,043,540 (81.33) | 79,291,108 (81.47) | 81,115,196 (82.52) |

Total reads: the number of clean reads is obtained by the sequencing data filter; total mapped: the number of sequencing data can be mapped into the genome; multiple mapped: the number of sequencing data have multiple alignment positions on the reference genome; uniquely mapped: the number of sequencing data have unique alignment position on the reference genome; read-1: the number of left reads can be aligned on the genome; read-2: the number of right reads can be aligned on the genome; reads map to '+': the statistics of the sequencing data can be aligned on the plus strand; reads map to '−': the statistics of the sequencing data can be aligned on the minus strand; nonsplice reads: the statistics of the whole fragment can be aligned on the exon; splice reads: the statistics of junction reads can be aligned on two exons. Reads mapped in proper pairs: the statistics of two ends can be aligned on reference sequence.

Potential target genes of differentially expressed miRNAs were predicted using TargetScan [18] and miRanda [19] algorithms. Enrichment analysis of Gene Ontology (GO) term and KEGG pathways for predicted and differentially expressed genes was conducted to get insights into the functions of mRNAs and miRNAs in PTB pathogenesis. These analyses will contribute to a more comprehensive understanding of the regulatory mechanisms involved in the development and pathogenesis of PTB.

## Materials & methods

### Patients with PTB & healthy individuals

PTB patients were diagnosed based on clinical manifestations, bacterial culture and radiographic findings. Patients had no major complications such as chronic obstructive pulmonary disease, asthma, lung cancer, pneumonia, diabetes or hypertension. All of patients (aged 26–35 years) had not received any medicines treatment before blood extraction. Healthy individuals had no family history of hereditary diseases or low immune function. Three paired samples were selected for detecting mRNAs with whole transcriptome sequencing and miRNAs with small RNA sequencing. Patients with PTB (two male and one female) and healthy individuals (two male and one female) were recruited from the First People's Hospital of Zhangjiagang, Jiangsu Province, China.

### Blood sample collection & RNA extraction

Fasting early morning blood samples were collected from participants in 3.0-ml tubes with heparin lithium anticoagulant. Within 4 h of collection, leukocytes were isolated from whole blood using extraction kits according to the manufacturer's instructions (TIANGEN, Beijing, China). Blood cells were transferred into microcentrifuge tubes with TRIzol reagent according to the manufacturer's protocol (Invitrogen, CA, USA). Total RNAs were quantified by NanoDrop™ ND-2000 (Thermo Fisher Scientific, CA, USA) and RNA integrity was assessed using an Agilent Bioanalyzer 2100 (Agilent Technologies, CA, USA).

### Whole transcriptome sequencing for mRNA & bioinformatics analysis

Total RNA was extracted from PTB and healthy samples using TRIzol reagent following the manufacturer's protocol. RNA integrity was evaluated using an Agilent 2100 Bioanalyzer. Samples with RNA integrity number greater than or equal to nine were used for subsequent analysis. Libraries were constructed using (TruSeq Stranded Total RNA with Ribo-Zero Gold, Illumina, CA, USA) according to the manufacturer's instructions. Libraries were sequenced on an Illumina sequencing platform (HiSeqTM 4000, Illumina, CA, USA) and 150 bp/125 bp paired-end reads were generated. Sequencing was carried out by the ShangHai Oebiotech Co. (Shanghai, China). Reads from the six samples were mapped to assembled transcripts using Bowtie2, and gene expression was estimated as fragments per kb per million reads [32]. Differentially expressed genes were identified using DESeq software (http:
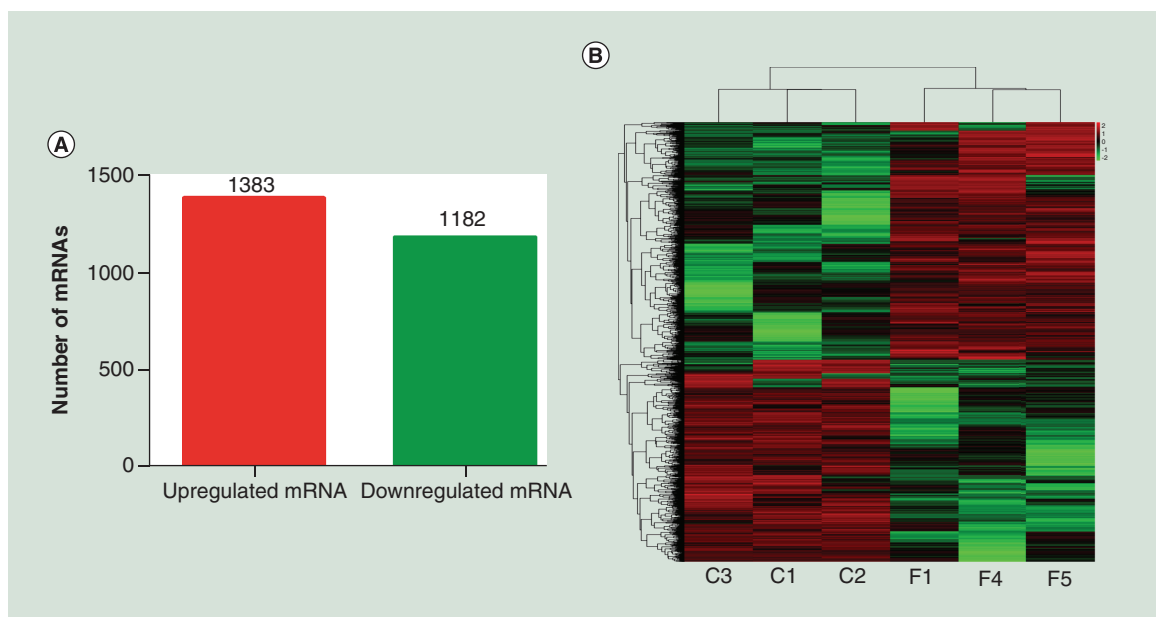
**Figure 1.    Identification of differentially expressed mRNAs in pulmonary tuberculosis. (A)** Differential expression of mRNAs identified in PTB and healthy individuals. **(B)** Heat map showing differentially expressed mRNAs comparing individuals with PTB versus healthy individuals. Each row represents one mRNA, and each column represents a sample. Red, upregulation; green, downregulation. C1, C2 and C3, healthy individuals, F1, F4 and F5, individuals with PTB. Six PTB patients (F1, F2, F3, F4, F5 and F6) and six healthy individuals (C1, C2, C3, C4, C5 and C6) were selected for total RNA extraction, but the low quality of RNA samples (F2, F3, F6, C4, C5 and C6) were discarded, so the original number were used in the Figures. For full color figures, please see online at https://www.futuremedicine.com/doi/full/10.2217/epi-2018-0028.
PTB: Pulmonary tuberculosis.

//bioconductor.org/packages/release/bioc/html/DESeq.html) [17]. mRNAs showing more than twofold changes with adjusted p < 0.05 were considered differentially expressed. Hierarchical clustering was performed and heat maps created using the R platform (http://www.rproject.org/). GO functional and KEGG pathway enrichment analysis were used to determine major biological processes and pathways of differentially expressed mRNAs with DAVID bioinformatics resources (https://david.ncifcrf.gov/home.jsp). GO terms and KEGG pathways with corrected p-values less than 0.05 were considered significantly enriched.

### Small RNA sequencing & bioinformatic analysis

Total RNA was extracted from PTB and healthy blood samples using mirVana RNA Isolation Kits (Applied Biosystems, p/n AM1556, CA, USA) and purified using QIAGEN RNeasy® Kits (QIAGEN, Mainz, Germany). An Agilent 2100 Bioanalyzer was used for quality testing after purification. A total of 5 μg RNA was used for small RNA sequencing by the ShangHai Oebiotech Co. (Shanghai, China).

Basic reads were converted into sequence data (called raw data/reads) by base calling. Low-quality reads were filtered, and reads with 5′-primer contaminants and poly (A) were removed. Reads without 3′-adapter and insert tags, and reads shorter than 15 nt and longer than 41 nt from raw data were filtered to obtain clean reads. For primary analysis, the length distribution of clean sequences in the reference genome was determined. Noncoding RNAs were annotated as rRNAs, tRNAs, snRNAs and snoRNAs. RNAs were aligned and subjected to BLAST (v2.2.28+) search against Rfam (v.10.1) (http://www.sanger.ac.uk/software/Rfam) and GenBank databases (http://www.ncbi.nlm.nih.gov/genbank/). Known miRNAs were identified by aligning against the miRBase (v.21) database (http://www.mirbase.org/), and known miRNA expression patterns in different samples were analyzed. Unannotated small RNAs were analyzed by MiRDeep2 (v.2.0.0.8) to predict novel miRNAs [33]. Based on the hairpin structure of pre-miRNAs and the miRBase database, corresponding miRNA star sequences were identified. Differentially expressed miRNAs were identified with a threshold fold change greater than two and adjusted p-value <0.05. P-values were calculated with the DESeq algorithm in the R package for experiments with biological
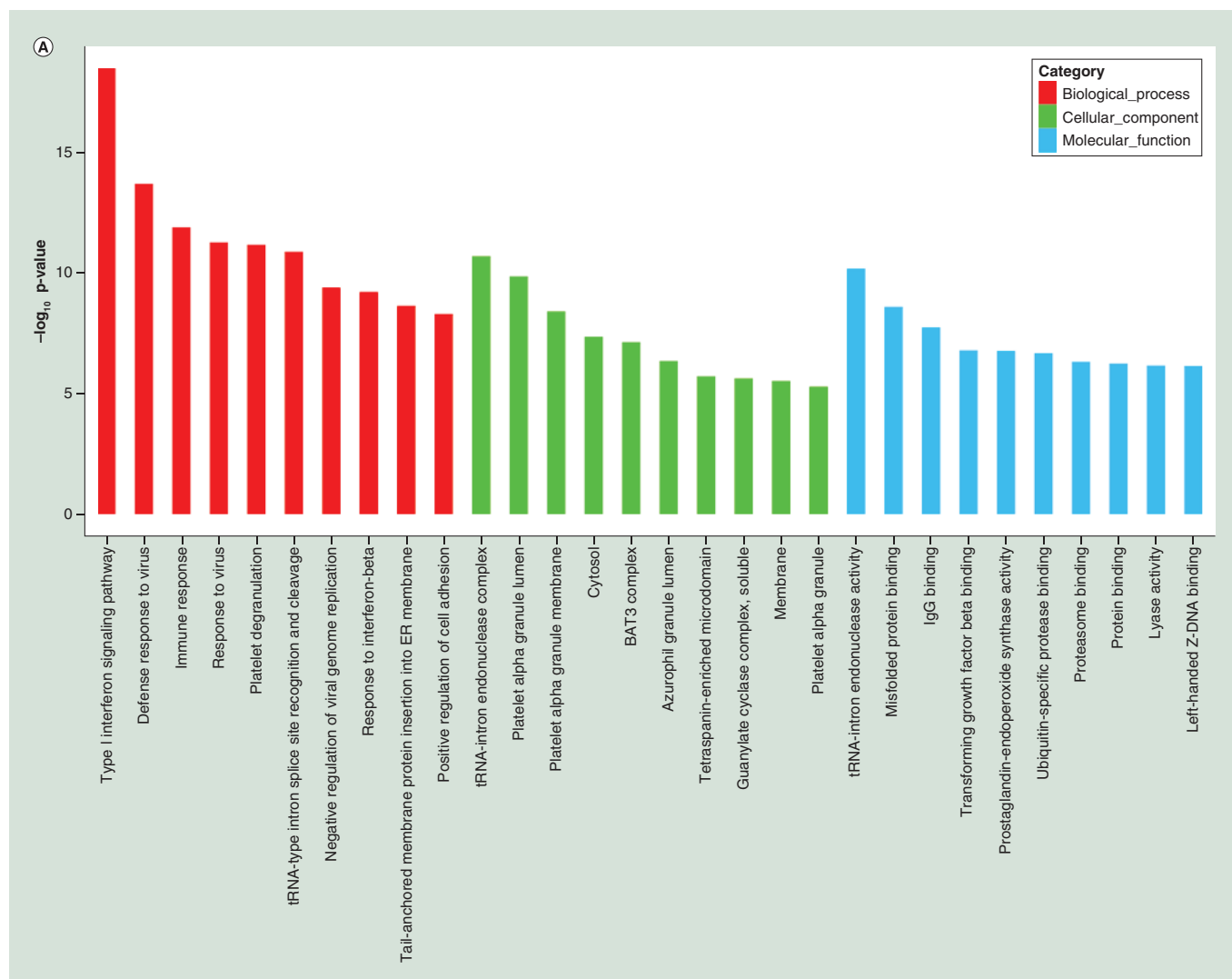
**Figure 2.    Top 20 Gene Ontology and KEGG terms for differentially expressed mRNAs between pulmonary tuberculosis and healthy individuals. (A)** GO terms for upregulated differentially expressed mRNAs between PTB and healthy individuals. **(B)** GO terms for downregulated differentially expressed mRNAs between PTB and healthy individuals. **(C)** Top 20 pathways for upregulated differentially expressed mRNAs between PTB and healthy individuals. **(D)** Top 20 pathways for downregulated differentially expressed mRNAs between PTB and healthy individuals.
PTB: Pulmonary tuberculosis.

replicates [17]. GO enrichment and KEGG pathway enrichment analysis of differentially expressed miRNA target genes used R (v3.2.0) based on hypergeometric distribution.

### Integration analysis

Targets of differentially expressed miRNAs were predicted TargetScan [18] and miRanda [19] algorithms, with parameter: $S \geq 150 \, \Delta G \leq -30$ kcal/mol and demanding strict 5′-seed pairing. According to differentially expression miRNAs profiles, differentially expression mRNA profiles and miRNA–mRNA interaction mechanisms, Pearson correlation coefficients were computed using R (http://www.R-project.org) to determine the negative correlated between the expression levels of each miRNA and its mRNA targets (correlation was <0 and adjusted p-value was <0.05) [34]. The negative correlated between the expression levels of each miRNA and its mRNA targets were screened out to construct the regulatory network with Cytoscape (version 3.0.1; http://www.cytoscape.org/) according to differentially expressed miRNAs and corresponding target genes.
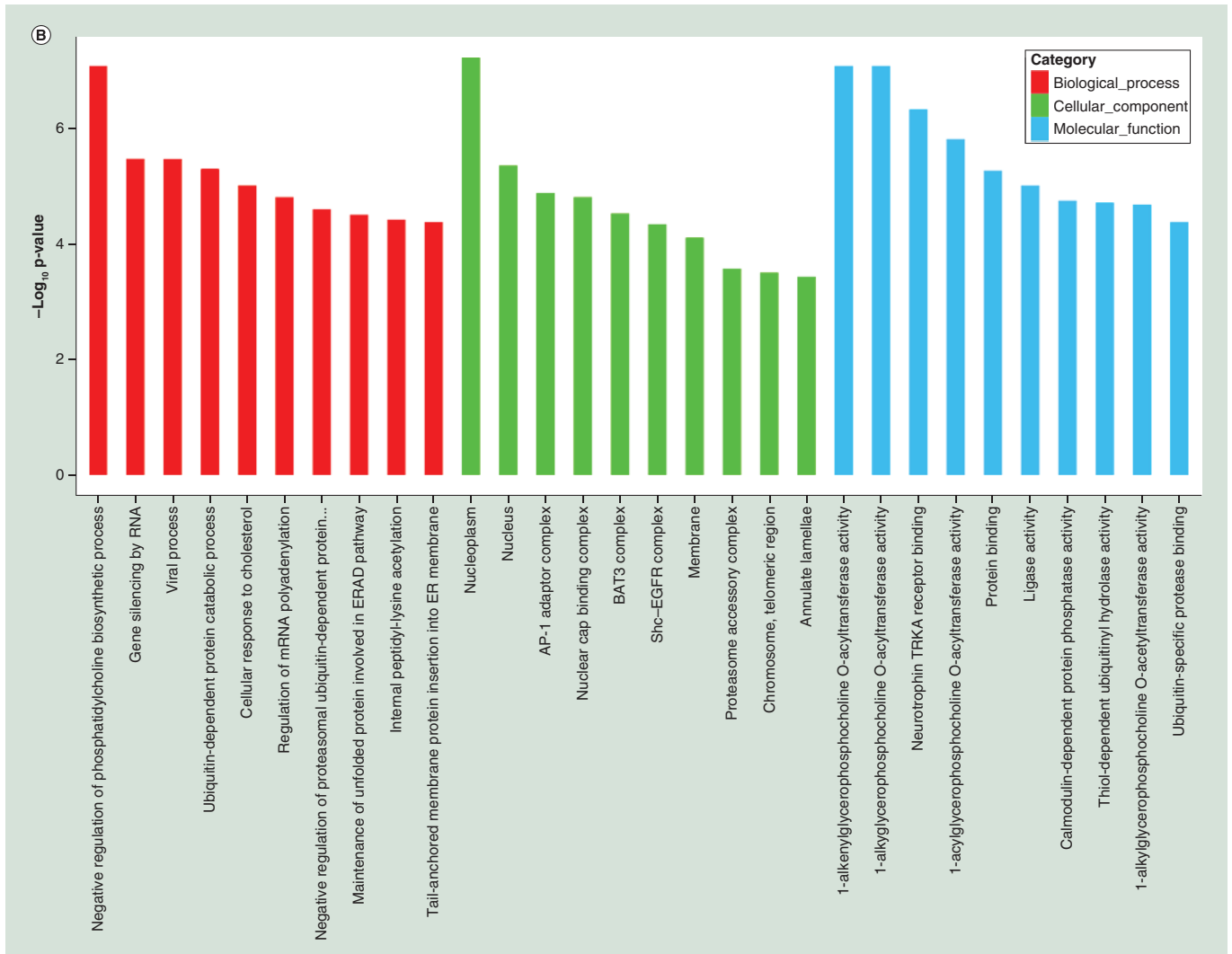
**Figure 2.   Top 20 Gene Ontology and KEGG terms for differentially expressed mRNAs between pulmonary tuberculosis and healthy individuals (cont.). (A)** GO terms for upregulated differentially expressed mRNAs between PTB and healthy individuals. **(B)** GO terms for downregulated differentially expressed mRNAs between PTB and healthy individuals. **(C)** Top 20 pathways for upregulated differentially expressed mRNAs between PTB and healthy individuals. **(D)** Top 20 pathways for downregulated differentially expressed mRNAs between PTB and healthy individuals.
PTB: Pulmonary tuberculosis.

## Results

### mRNA sequencing & analysis

To better understand the pathogenic mechanism of PTB, we conducted a comparative transcriptomic analysis of three people with PTB and three healthy individuals. A total of six cDNA libraries, named F1, F4, F5, C1, C2 and C3 were constructed and sequenced. From the libraries, 97,193,002; 97,321,140; 98,293,852; 97,902,050; 98,099,296 and 98,047,510 clean reads were obtained, and 94.55, 94.71, 94.76, 94.76, 94.52 and 94.75% reads were mapped to the reference genome (Table 1). Original data were normalized and overall characteristics of data distributions for the six samples are shown as box plots (Supplementary Figure 1A). Raw sequencing data from the six samples were normalized and transformed into log2 values. A scatter plot was created and overall distribution of the two datasets was evaluated in a 2D coordinate system (Supplementary Figure 1B).
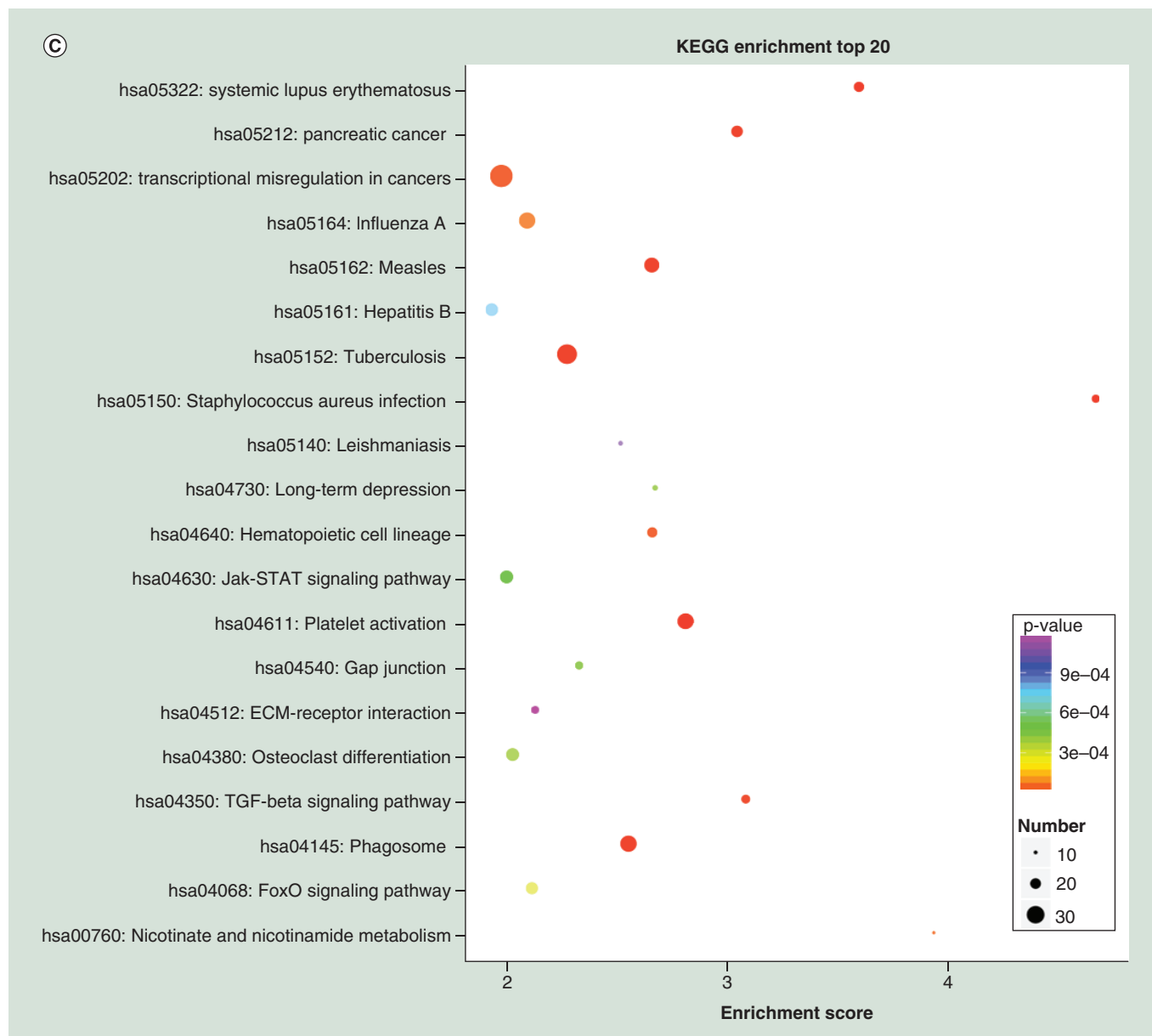
**Figure 2.    Top 20 Gene Ontology and KEGG terms for differentially expressed mRNAs between pulmonary tuberculosis and healthy individuals (cont.). (A)** GO terms for upregulated differentially expressed mRNAs between PTB and healthy individuals. **(B)** GO terms for downregulated differentially expressed mRNAs between PTB and healthy individuals. **(C)** Top 20 pathways for upregulated differentially expressed mRNAs between PTB and healthy individuals. **(D)** Top 20 pathways for downregulated differentially expressed mRNAs between PTB and healthy individuals.
PTB: Pulmonary tuberculosis.

## Differentially expressed gene identification

Significant differences in mRNAs between individuals with PTB and healthy individuals were screened for fold change (greater than or equal to twofold change) and adjusted p-value (p < 0.05) determined by DESeq analysis. The number and distribution of mRNAs in the same plane were displayed in volcano plots satisfying both conditions (Supplementary Figure 2). Two thousand five hundred and sixty-five mRNAs were screened out with significant levels of differentially expression comparing PTB patients with healthy individuals (greater than or equal to twofold change; adjusted p < 0.05), with 1383 significantly upregulated and 1182 significantly downregulated (Figure 1A). Using unsupervised hierarchical clustering analysis, a heat map was generated with differentially expressed mRNAs,
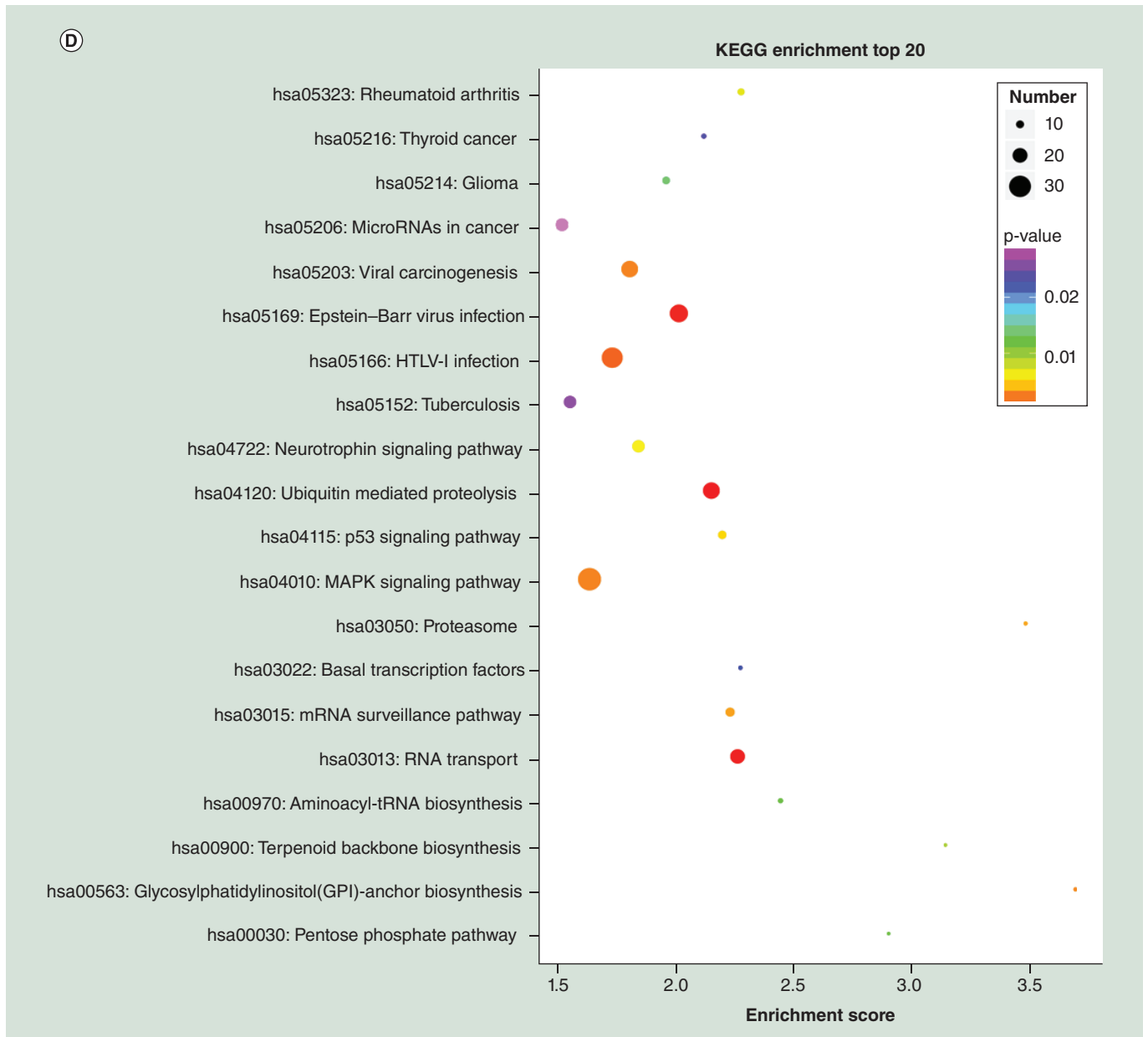
**Figure 2.  Top 20 Gene Ontology and KEGG terms for differentially expressed mRNAs between pulmonary tuberculosis and healthy individuals (cont.). (A)** GO terms for upregulated differentially expressed mRNAs between PTB and healthy individuals. **(B)** GO terms for downregulated differentially expressed mRNAs between PTB and healthy individuals. **(C)** Top 20 pathways for upregulated differentially expressed mRNAs between PTB and healthy individuals. **(D)** Top 20 pathways for downregulated differentially expressed mRNAs between PTB and healthy individuals.
PTB: Pulmonary tuberculosis.

and the results showed that mRNA expression level in individuals of PTB can be robustly separated from that in healthy individuals (Figure 1B).

## GO term & KEGG pathway enrichment analysis of differentially expressed mRNA

To further understand the function of differentially expressed genes, GO term and KEGG pathway analyses were carried out. Top 20 dysregulated GO processes for each subgroup (biological process, cellular component and molecular function) were analyzed according to enriched, dysregulated mRNAs derived from the gene annotation. Prediction terms with p-value <0.05 were selected and ranked by p-value. Based on routine GO classification
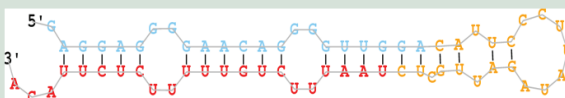
```
Provisional ID          : NC_000001.11_2005
Score total             :    5.3
Score for star read(s)  :   -1.3
Score for read counts   :    0
Score for mfe           :    2
Score for randfold      :    1.6
Score for cons. seed    :    3
Total read count        :   14
Mature read count       :   14
Loop read count         :    0
Star read count         :    0
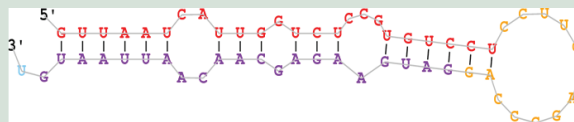```
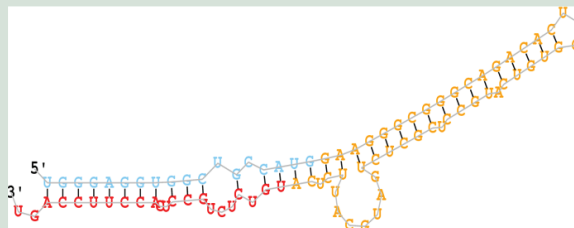


```
Provisional ID          : NC_000001.11_61
Score total             :    0.7
Score for star read(s)  :   -1.3
Score for read counts   :   -4.9
Score for mfe           :    2.4
Score for randfold      :    1.6
Score for cons. seed    :    3
Total read count        :    2
Mature read count       :    2
Loop read count         :    0
Star read count         :    0
```
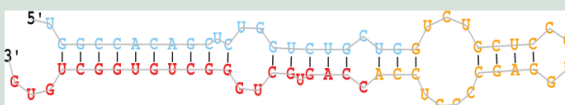
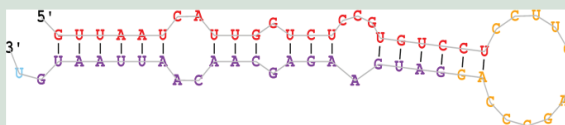

```
Provisional ID          : NC_000001.11_597
Score total             :    4.8
Score for star read(s)  :   -1.3
Score for read counts   :    0
Score for mfe           :    1.5
Score for randfold      :    1.6
Score for cons. seed    :    3
Total read count        :   17
Mature read count       :   17
Loop read count         :    0
Star read count         :    0
```



```
Provisional ID          : NC_000001.11_2269
Score total             :    6
Score for star read(s)  :    3.9
Score for read counts   :   -4.4
Score for mfe           :    1.9
Score for randfold      :    1.6
Score for cons. seed    :    3
Total read count        :    3
Mature read count       :    2
Loop read count         :    0
Star read count         :    1
```



```
Provisional ID          : NC_000001.11_2269
Score total             :    6
Score for star read(s)  :    3.9
Score for read counts   :   -4.4
Score for mfe           :    1.9
Score for randfold      :    1.6
Score for cons. seed    :    3
Total read count        :    3
Mature read count       :    2
Loop read count         :    0
Star read count         :    1
```



```
Provisional ID          : NC_000001.11_173
Score total             :    0.5
Score for star read(s)  :   -1.3
Score for read counts   :   -4.9
Score for mfe           :    2.2
Score for randfold      :    1.6
Score for cons. seed    :    3
Total read count        :    2
Mature read count       :    2
Loop read count         :    0
Star read count         :    0
```
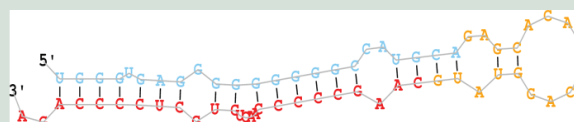


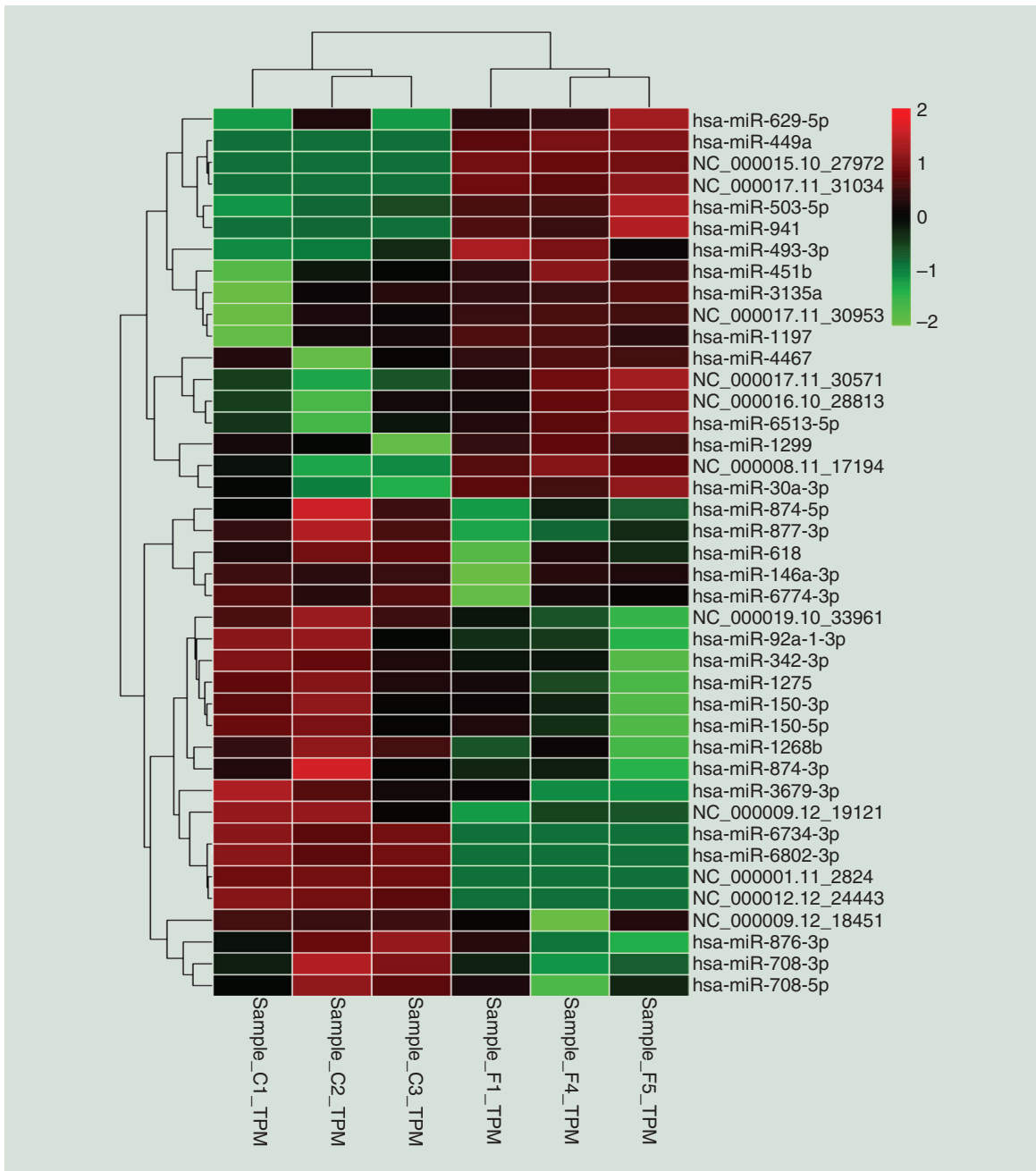**Figure 3.    Six novel miRNAs in pulmonary tuberculosis.**

**Figure 4.   Clustering of expression patterns of 41 differentially expressed miRNAs.** Each row represents one miRNA, and each column represents a sample. Red, upregulation; green, downregulation. C1, C2 and C3, healthy individuals, F1, F4 and F5, individuals with PTB.

algorithms, an enrichment score was used to enrich significant GO terms of differentially expressed genes. GO term enrichment of the upregulated mRNAs included type I interferon signaling pathway, defense response to virus and immune response to virus in biological process; tRNA–intron endonuclease complex, platelet α granule lumen and platelet α granule membrane in cellular component; and tRNA-intron endonuclease activity, misfolded protein binding and IgG binding in molecular function (Figure 2A). GO term enrichment of the downregulated mRNAs included negative regulation of phospholipid biosynthetic process, gene silencing by RNA and viral process in biological process; nucleoplasm, nucleus and AP-1 adaptor complex in cellular component; and 1-alkenylglycerophosphocholine O-acyltransferase activity, 1-alkylglycerophosphocholine O-acyltransferase activity

## Table 2. The summary information of miRNA.

| Sample | Raw_before_deadaptor | Raw_reads | Reads_trimmed_length | Reads_trimmed_Q20 | Reads_trimmed_N | Clean_reads | Clean_reads_uniq |
|---|---|---|---|---|---|---|---|
| Sample_C1 | 23,447,966 | 23,414,044 | 22,314,440 | 22,305,553 | 22,263,514 | 22,263,514 | 527,134 |
| Sample_C2 | 26,400,719 | 26,357,539 | 25,247,090 | 25,232,856 | 25,185,500 | 25,185,500 | 687,431 |
| Sample_C3 | 22,701,428 | 22,672,995 | 21,819,033 | 21,809,034 | 21,767,699 | 21,767,699 | 566,080 |
| Sample_F1 | 25,367,221 | 25,342,877 | 24,107,468 | 24,094,514 | 24,049,448 | 24,049,448 | 592,356 |
| Sample_F4 | 24,699,744 | 24,678,408 | 23,917,689 | 23,906,109 | 23,861,054 | 23,861,054 | 593,053 |
| Sample_F5 | 26,671,292 | 26,645,820 | 25,948,445 | 25,932,545 | 25,883,380 | 25,883,380 | 786,938 |

Sample: sample names; raw_reads: the statistics of the raw reads with trimming down adaptor; reads_trimmed_length: reserve reads with lengths of 15–40 bp; reads_trimmed_Q20: reserve the reads of the percent of Q20 >80%; reads_trimmed_N: trim down the sequences with N base; clean_reads: remove the contaminated and low quality reads; clean_reads_uniq: clean reads, remove the redundant reads.

and neurotrophin tyrosine kinase A (TRKA) receptor binding in molecular function (Figure 2B). Using KEGG pathway analysis, the upregulated mRNAs were enriched in transcriptional misregulation in cancers, TB and phagosome (Figure 2C), and the downregulated mRNAs were enriched in MAPK signaling pathway, HTLV-I infection and Epstein–Barr virus infection (Figure 2D). The results suggested that these pathways might contribute significantly to PTB pathogenesis.

### miRNA library construction & identification
To identify miRNAs involved in PTB pathogenesis, six small RNA libraries, F1, F4, F5, C1, C2 and C3, were constructed and sequenced. Acquired raw reads were 25,342,877; 24,678,408; 26,645,820; 23,414,044; 26,357,539 and 22,672,995. After removing low-quality sequencing data, 24,049,448; 23,861,054; 25,883,380; 22,263,514; 25,185,500 and 21,767,699 clean reads were obtained from total reads of F1, F4, F5, C1, C2 and C3 (Table 2).

### Identification of putative novel miRNAs
Small RNA sequencing data were aligned with databases, and tRNAs, snRNAs and rRNAs were removed. The obtained sequences were used to predict potential novel miRNAs. Hundreds of novel miRNAs were identified with Mirdeep2 software (Supplementary Table 1) and hairpin structures formed by precursor sequences of four novel miRNAs are illustrated by RNAfold software (Figure 3).

### miRNA differential expression profiles
Combining known miRNAs and novel miRNAs identified 2832 miRNAs. Differential expression revealed 41 significantly differentially expressed miRNAs in PTB versus healthy samples (greater than or equal to twofold change; adjusted p < 0.05): 18 upregulated miRNAs and 23 downregulated (Table 3). Using unsupervised hierarchical clustering analysis, a heat map was constructed based on differentially expressed miRNAs (Figure 4).

### Target prediction for significantly differentially expressed miRNAs & functional analysis
To provide information about the possible functions of 41 differentially expressed miRNAs, target prediction analysis was conducted by TargetScan and miRanda algorithms. These algorithms identified 359 putative target genes, including genes that participate in positive regulation of cytokine production in the immune response such as interferon and free fatty acid receptor 3.

To interpret possible physiological processes and pathways regulated by the identified miRNAs, their putative target genes were subjected to GO term and KEGG pathway analysis. The top 30 enriched GO terms of the down-regulated miRNAs were mainly in biological processes (chromosome organization, chromatin modification and histone modification), cellular components (nuclear body and nuclear speck) and molecular function (chromatin binding and phosphoric ester hydrolase activity) (Figure 5A). The top 30 enriched GO terms of the upregulated miRNAs were mainly in biological processes (regulation of small GTPase-mediated signal transduction, platelet activation and negative regulation of locomotion), cellular components (phagocytic vesicle, actin filament bundle and cortical cytoskeleton) and molecular function (enzyme activator activity, GTPase activator activity and GTPase regulator activity) (Figure 5B).

Using KEGG pathway analysis, the upregulated miRNAs were enriched in focal adhesion, tight junction and phagosome pathways (Figure 5C), and the downregulated miRNAs were enriched in calcium signaling pathway,

| Table 3. Differential expression of miRNA from the comparative pulmonary tuberculosis with healthy individuals. | | | | | | |
|---|---|---|---|---|---|---|
| miRNA_id | baseMean_Group_C | baseMean_Group_F | foldChange | log2FoldChange | adjusted pval | up_down |
| NC_000001.11_2824 | 3.293118593 | 0 | 0 | – | 0.008609 | Down |
| NC_000008.11_17194 | 1.653214244 | 7.937837235 | 4.801457 | 2.26347228 | 0.014603 | Up |
| NC_000009.12_18451 | 5.313649934 | 0.970885613 | 0.182715 | -2.4523299 | 0.045129 | Down |
| NC_000009.12_19121 | 55.80079408 | 22.97934618 | 0.41181 | -1.2799479 | 0.011141 | Down |
| NC_000012.12_24443 | 2.748797896 | 0 | 0 | – | 0.024486 | Down |
| NC_000015.10_27972 | 0 | 2.51951425 | Inf. | Inf. | 0.048426 | Up |
| NC_000016.10_28813 | 3.12248263 | 9.774152856 | 3.130251 | 1.64627812 | 0.02964 | Up |
| NC_000017.11_30571 | 5.011893215 | 14.11435049 | 2.816171 | 1.49373516 | 0.024349 | Up |
| NC_000017.11_30953 | 0.970830534 | 5.860731352 | 6.036822 | 2.59378932 | 0.034683 | Up |
| NC_000017.11_31034 | 0 | 2.470248473 | Inf. | Inf. | 0.037992 | Up |
| NC_000019.10_33961 | 12.18196631 | 4.003244869 | 0.328621 | -1.6055052 | 0.014473 | Down |
| hsa-miR-1197 | 1.33814763 | 11.18981589 | 8.362168 | 3.06387711 | 0.007865 | Up |
| hsa-miR-1268b | 49.14717846 | 6.42462627 | 0.130722 | -2.9354242 | 0.006871 | Down |
| hsa-miR-1275 | 419.2961936 | 170.3601747 | 0.4063 | -1.2993816 | 0.003894 | Down |
| hsa-miR-1299 | 2.079724081 | 28.63688778 | 13.76956 | 3.78341067 | 0.008727 | Up |
| hsa-miR-146a-3p | 22.56017442 | 6.870204203 | 0.304528 | -1.7153533 | 0.011597 | Down |
| hsa-miR-150-3p | 658.1221256 | 367.3712096 | 0.558211 | -0.8411168 | 0.018789 | Down |
| hsa-miR-150-5p | 136080.5326 | 80982.07062 | 0.595104 | -0.7487863 | 0.034813 | Down |
| hsa-miR-30a-3p | 106.5967314 | 208.6482221 | 1.95736 | 0.96890943 | 0.011905 | Up |
| hsa-miR-3135a | 1.036390912 | 5.026565826 | 4.850067 | 2.27800481 | 0.047487 | Up |
| hsa-miR-342-3p | 27185.13578 | 14597.87006 | 0.53698 | -0.8970601 | 0.003355 | Down |
| hsa-miR-3679-5p | 27.61215746 | 11.13876598 | 0.403401 | -1.3097142 | 0.028487 | Down |
| hsa-miR-4467 | 1.351457526 | 8.295960858 | 6.138529 | 2.61789292 | 0.004472 | Up |
| hsa-miR-449a | 0 | 2.556317353 | Inf. | Inf. | 0.036862 | Up |
| hsa-miR-451b | 7.182808366 | 25.0356799 | 3.4855 | 1.80136571 | 0.017658 | Up |
| hsa-miR-493-3p | 122.6772606 | 257.2427135 | 2.096906 | 1.06826236 | 0.016758 | Up |
| hsa-miR-503-5p | 71.37212831 | 172.9753578 | 2.42357 | 1.27713383 | 0.00256 | Up |
| hsa-miR-618 | 2733.249783 | 1653.286546 | 0.604879 | -0.7252805 | 0.020802 | Down |
| hsa-miR-629-5p | 630.9940314 | 1090.927771 | 1.728903 | 0.78985732 | 0.032824 | Up |
| hsa-miR-6513-5p | 3.916309564 | 10.88729718 | 2.779989 | 1.47507913 | 0.03754 | Up |
| hsa-miR-6734-3p | 2.512601555 | 0 | 0 | – | 0.041909 | Down |
| hsa-miR-6774-3p | 4.775696875 | 0.528477246 | 0.11066 | -3.175798 | 0.023333 | Down |
| hsa-miR-6802-3p | 2.512601555 | 0 | 0 | – | 0.041909 | Down |
| hsa-miR-708-3p | 14.05112474 | 4.303394528 | 0.306267 | -1.7071386 | 0.01557 | Down |
| hsa-miR-708-5p | 95.20256325 | 23.66608681 | 0.248587 | -2.0081792 | 0.038023 | Down |
| hsa-miR-874-3p | 47.16004903 | 24.74442465 | 0.52469 | -0.9304617 | 0.047116 | Down |
| hsa-miR-874-5p | 8.206464002 | 2.812739856 | 0.342747 | -1.5447846 | 0.04086 | Down |
| hsa-miR-876-3p | 12.37170518 | 4.730170654 | 0.382338 | -1.3870802 | 0.035913 | Down |
| hsa-miR-877-3p | 11.68989609 | 3.990782195 | 0.341387 | -1.5505187 | 0.031932 | Down |
| hsa-miR-92a-1-5p | 119.0211675 | 52.73528934 | 0.443075 | -1.1743776 | 0.004364 | Down |
| hsa-miR-941 | 3823.249903 | 10516.22374 | 2.750598 | 1.45974533 | 1.36E-06 | Up |

Differentially expressed miRNAs were screened out a threshold (fold change >2 and adjusted p-value <0.05) with the DESeq algorithm in this study.
–: Insignificance; Inf.: Infinite.

insulin signaling pathway and axon guidance (Figure 5D). The results suggested that these pathways might contribute significantly to PTB pathogenesis.

## mRNA–miRNA interaction network
To understand the miRNA regulatory mechanisms underlying the pathogenesis of PTB, interactions between miRNAs and their target mRNAs were investigated. Based on the regulatory mechanism of miRNAs and target
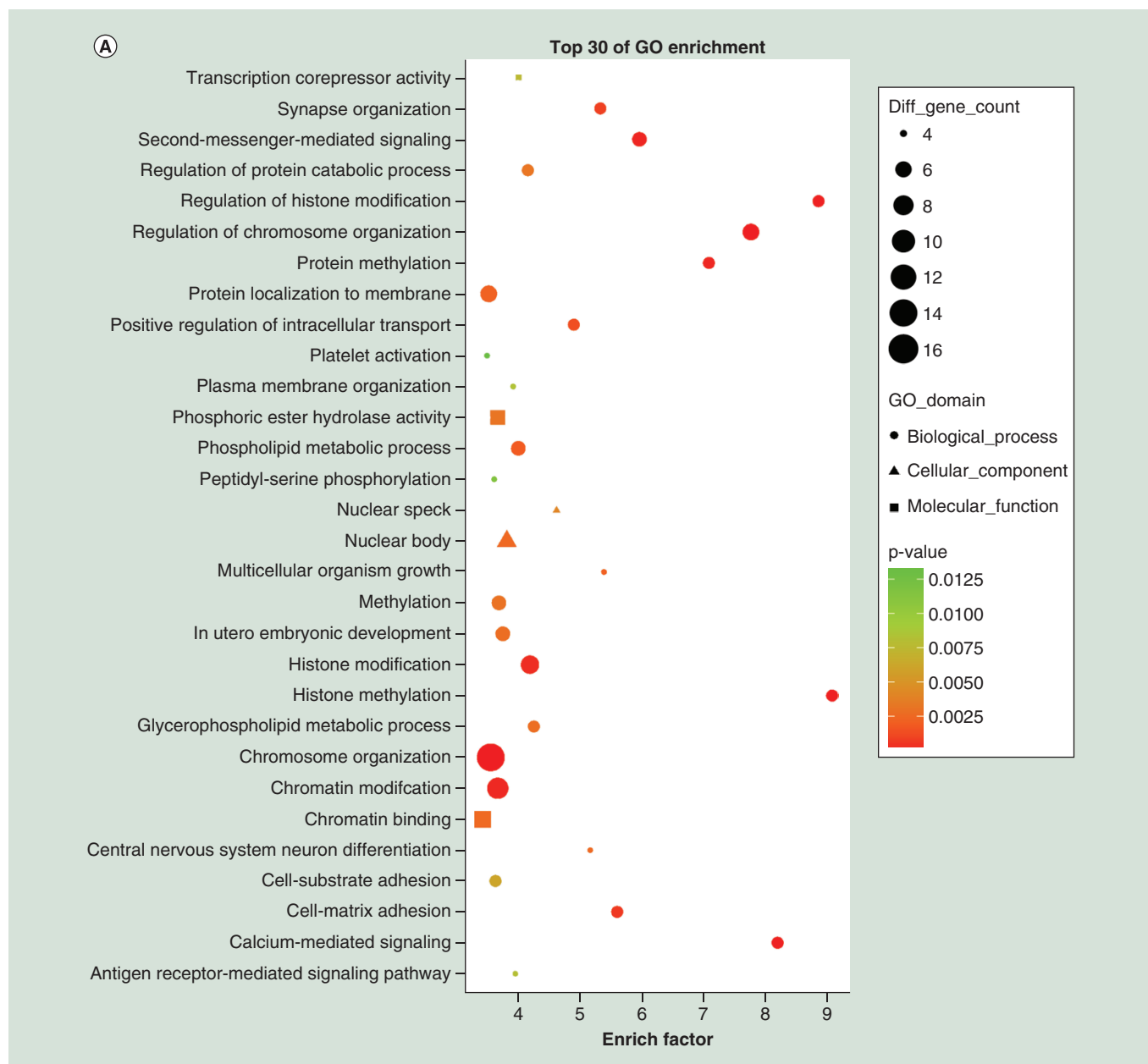
**Figure 5.   Top 30 enriched Gene Ontology terms and KEGG pathways for differentially expressed miRNAs between pulmonary tuberculosis and healthy individuals. (A)** Top 30 Gene Ontology terms for upregulated differentially expressed miRNAs (downregulation of target mRNAs) between PTB and healthy individuals. **(B)** Top 30 Gene Ontology terms for downregulated differentially expressed miRNAs (upregulation of target mRNAs) between PTB and healthy individuals. **(C)** Top 30 pathways for upregulated differentially expressed miRNAs (downregulation of target mRNAs) between PTB and healthy individuals. **(D)** Top 30 pathways for downregulated differentially expressed miRNAs (upregulation of target mRNAs) between PTB and healthy individuals. Gene number: number of target genes in a term or pathway. Rich factor: ratio of number of target genes divided by number of all the genes in a term or pathway. PTB: Pulmonary tuberculosis.

genes (mRNAs), we determined a group of negatively correlated upregulated miRNAs and downregulated mRNAs (Figure 6A) and another group of positively correlated downregulated miRNAs and upregulated mRNAs (Figure 6B) to construct regulatory networks. Three hundred and fifty-nine potential target genes were identified for 41 miRNAs. A series of immunity-related genes including leukocyte immunoglobulin-like receptor, Fc fragment of IgG binding protein and BCL2-like 11 were identified.
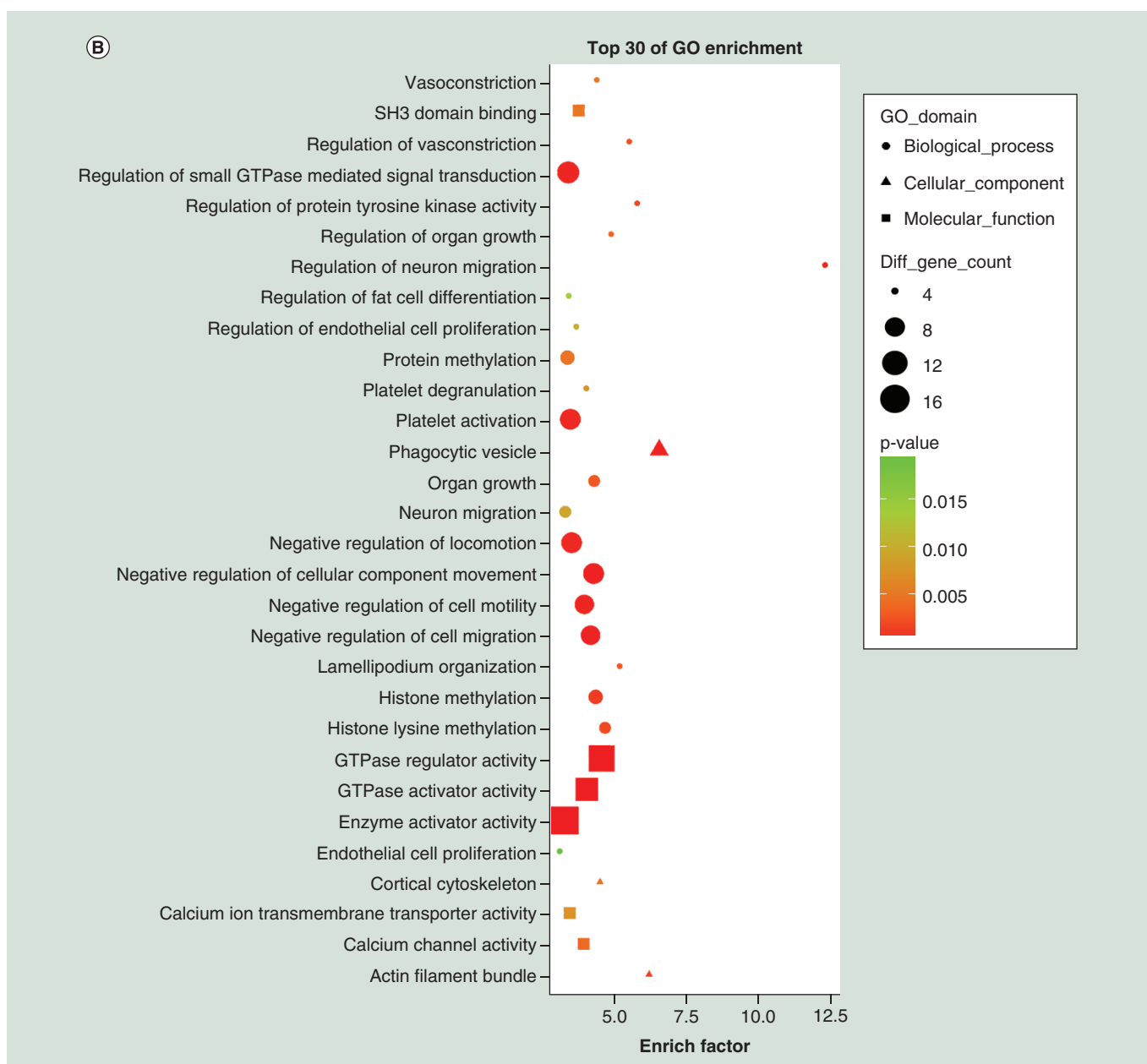
**Figure 5.    Top 30 enriched Gene Ontology terms and KEGG pathways for differentially expressed miRNAs between pulmonary tuberculosis and healthy individuals (cont.). (A)** Top 30 Gene Ontology terms for upregulated differentially expressed miRNAs (downregulation of target mRNAs) between PTB and healthy individuals. **(B)** Top 30 Gene Ontology terms for downregulated differentially expressed miRNAs (upregulation of target mRNAs) between PTB and healthy individuals. **(C)** Top 30 pathways for upregulated differentially expressed miRNAs (downregulation of target mRNAs) between PTB and healthy individuals. **(D)** Top 30 pathways for downregulated differentially expressed miRNAs (upregulation of target mRNAs) between PTB and healthy individuals. Gene number: number of target genes in a term or pathway. Rich factor: ratio of number of target genes divided by number of all the genes in a term or pathway.
PTB: Pulmonary tuberculosis.

## Discussion

miRNAs are epigenetic modulators that post-transcriptionally regulate the expression of protein-coding genes. They have emerged as novel regulators in various biological processes and pathogenic conditions such as inflammation, cancer and infectious diseases [20]. They are important gene regulatory factors in multicellular genomes. More than 50–60% of cellular mRNAs are thought to be regulated by miRNAs. One miRNA can target multiple mRNAs,
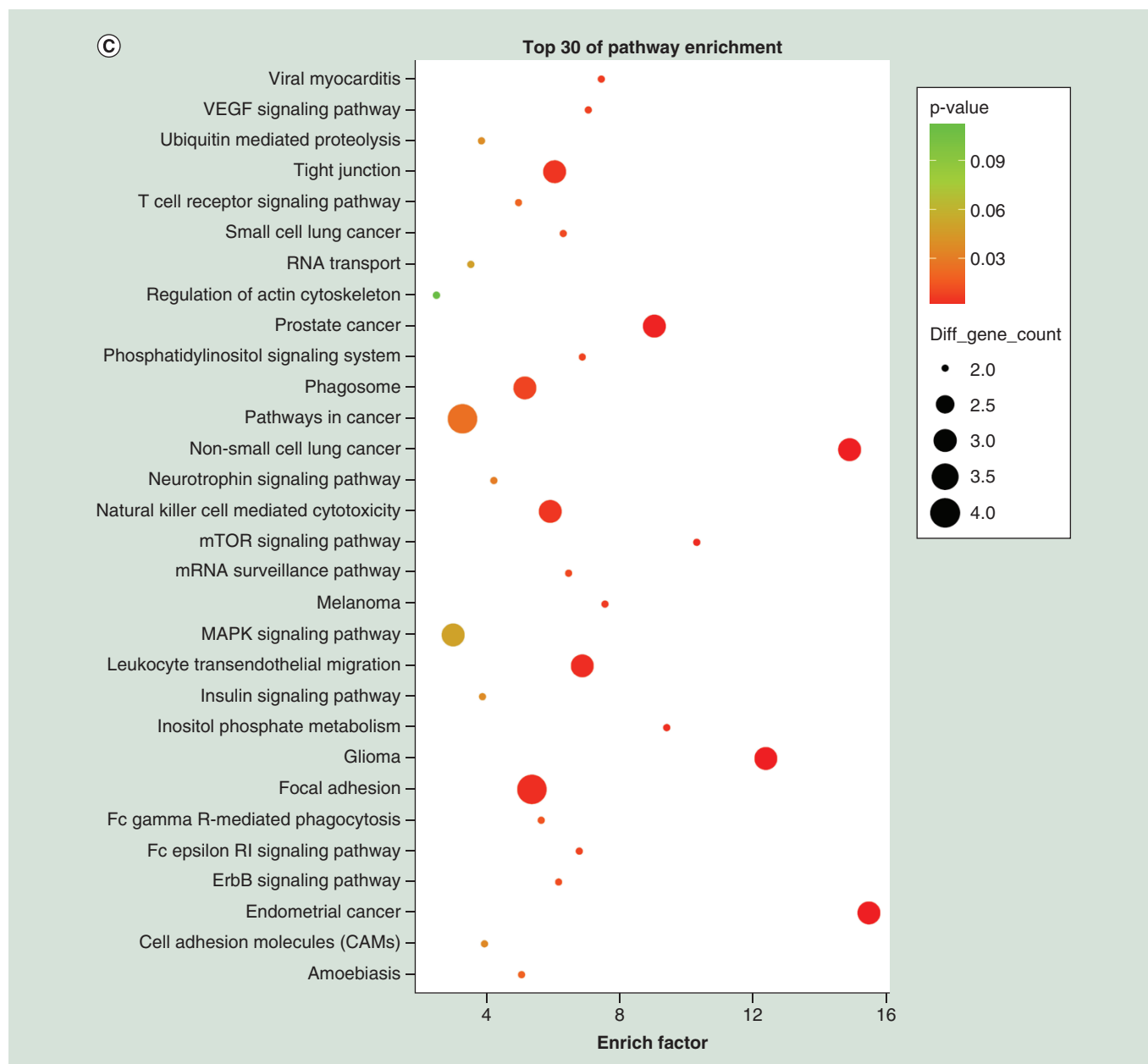
**Figure 5.    Top 30 enriched Gene Ontology terms and KEGG pathways for differentially expressed miRNAs between pulmonary tuberculosis and healthy individuals (cont.). (A)** Top 30 Gene Ontology terms for upregulated differentially expressed miRNAs (downregulation of target mRNAs) between PTB and healthy individuals. **(B)** Top 30 Gene Ontology terms for downregulated differentially expressed miRNAs (upregulation of target mRNAs) between PTB and healthy individuals. **(C)** Top 30 pathways for upregulated differentially expressed miRNAs (downregulation of target mRNAs) between PTB and healthy individuals. **(D)** Top 30 pathways for downregulated differentially expressed miRNAs (upregulation of target mRNAs) between PTB and healthy individuals. Gene number: number of target genes in a term or pathway. Rich factor: ratio of number of target genes divided by number of all the genes in a term or pathway.
PTB: Pulmonary tuberculosis.

and in turn, one mRNA can be targeted by multiple miRNAs [8,21]. Understanding the complexity of the miRNA network in host–pathogen interactions may open new avenues for identifying biomarkers and improving the efficacy of therapies against TB in humans.

Many miRNAs are altered in peripheral blood cells during active TB, identified by high-throughput sequencing. Numerous candidate miRNAs could be used as biomarkers for human TB [16,22]. Although a large number of
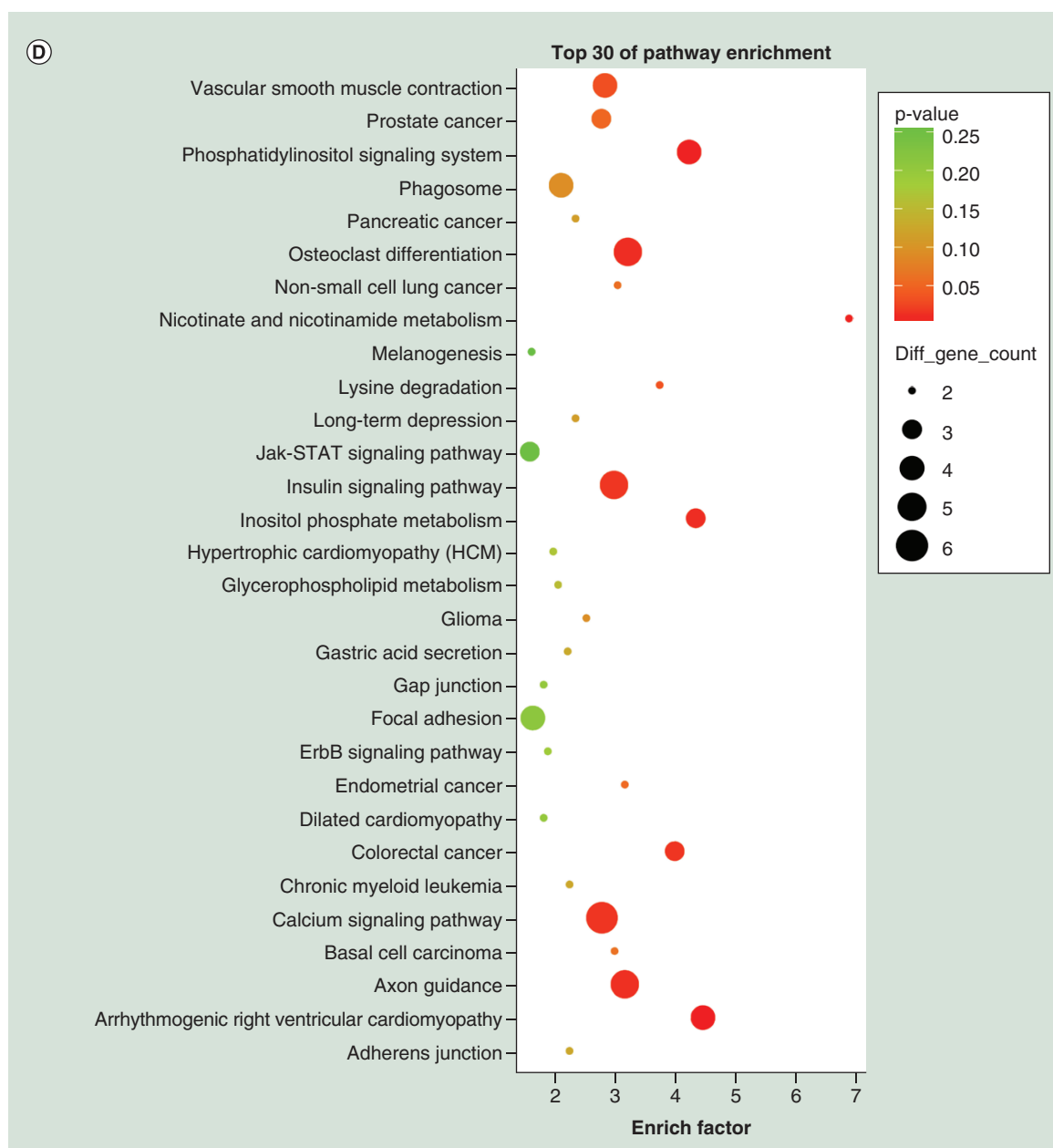
**Figure 5.    Top 30 enriched Gene Ontology terms and KEGG pathways for differentially expressed miRNAs between pulmonary tuberculosis and healthy individuals (cont.). (A)** Top 30 Gene Ontology terms for upregulated differentially expressed miRNAs (downregulation of target mRNAs) between PTB and healthy individuals. **(B)** Top 30 Gene Ontology terms for downregulated differentially expressed miRNAs (upregulation of target mRNAs) between PTB and healthy individuals. **(C)** Top 30 pathways for upregulated differentially expressed miRNAs (downregulation of target mRNAs) between PTB and healthy individuals. **(D)** Top 30 pathways for downregulated differentially expressed miRNAs (upregulation of target mRNAs) between PTB and healthy individuals. Gene number: number of target genes in a term or pathway. Rich factor: ratio of number of target genes divided by number of all the genes in a term or pathway.
PTB: Pulmonary tuberculosis.

miRNAs have been identified in PTB, miRNA–mRNA interactional regulatory mechanisms in the process of PTB are less reported. Screening comparing patients with PTB and healthy individuals (greater than or equal to twofold change; adjusted p < 0.05) identified 2565 mRNAs with significant levels of differential expression, with 1383 significantly upregulated and 1182 significantly downregulated. A series of immunity-related genes including
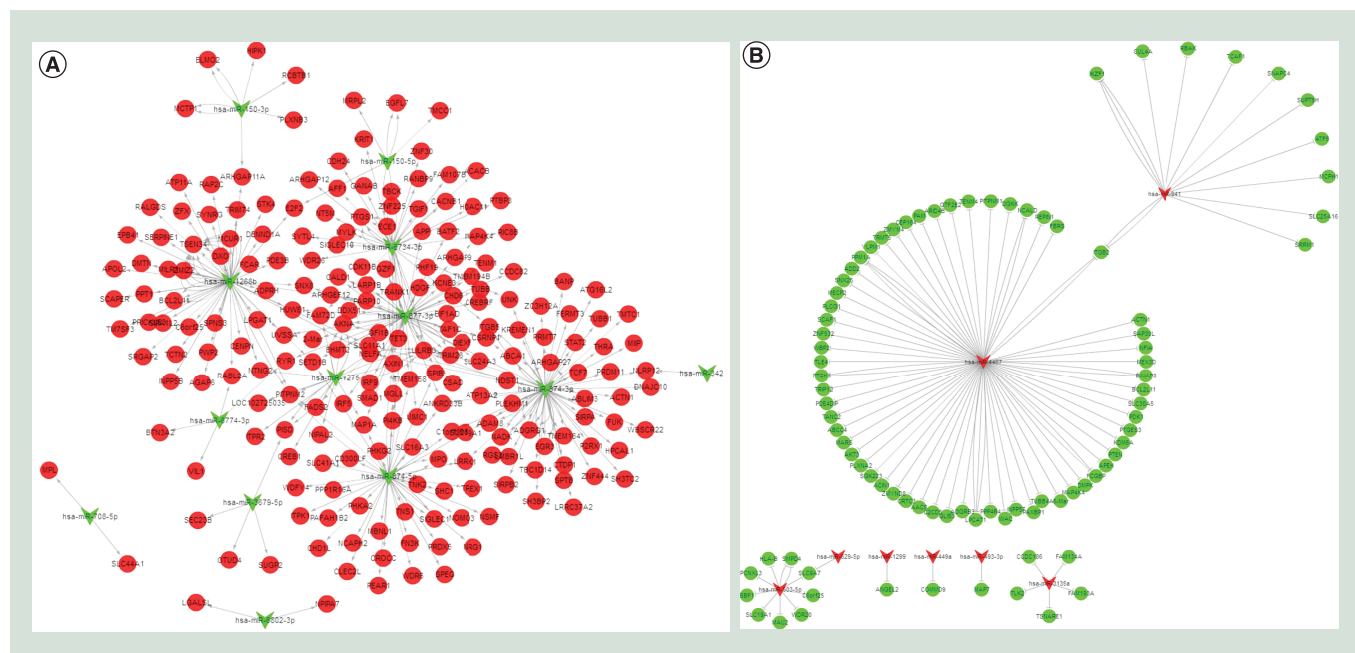
**Figure 6.    Proposed networks of putative interactions between miRNAs and mRNAs in pulmonary tuberculosis pathogenesis.**
Regulatory networks of miRNAs and mRNAs in PTB pathogenesis are illustrated by Cytoscape. Circle nodes, mRNAs; triangle nodes, miRNAs. **(A)** Positive correlation of miRNAs and mRNAs identified from PTB. **(B)** Negative correlation of miRNA and mRNA identified from PTB. All of the differentially expressed mRNA (fold change ≥2 and adjusted p < 0.05) was obtained from the whole transcriptome analysis, and the differentially expressed miRNA (fold change ≥2 and adjusted p < 0.05) was obtained from the small RNA sequencing analysis. Pearson correlation coefficients were computed using R (http://www.R-project.org) to determine the negative correlated between the expression levels of each miRNA and its mRNA targets.
PTB: Pulmonary tuberculosis.

leukocyte immunoglobulin-like receptor, Fc fragment of IgG binding protein and BCL2-like 11 were obtained. Differential expression analysis revealed 41 significantly differentially expressed miRNAs in patients with PTB versus healthy individuals (greater than or equal to twofold change; adjusted p < 0.05), including 18 upregulated miRNAs and 23 downregulated miRNAs. A series of important miRNAs included has-miR-150-3p, has-miR-150-5p, has-miR-874-5p and has-miR-941. Studies show that circulating miR-150, miR-146a and miR-125b identified from childhood TB may have diagnostic value with a combination of differentially expressed miRNAs in childhood TB [23]. Has-miR-3179 and has-miR-19b-2* were identified the most increased and decreased miRNAs in a comparison of sputum samples from patients with TB and controls [24]. In patients with PTB, miR-29 levels were increased in CD4+ T cells with an inverse correlation with IFN-γ mRNA expression [14]. MIR144* levels were upregulated in peripheral blood mononuclear cell (PBMCs) and active TB patients, targeting an autophagy and lysosomal protein (DRAM2), to contribute to the pathogenesis of TB by autophagic inhibition control [25]. Despite numerous studies on differentially expressed miRNAs in PTB, many challenges still exist in accurate normalization of miRNA levels for clinical use among different experiments. These results indicated the possibility that, as indicated in other reports, miRNAs are likely to participate in PTB pathogenesis.

Interaction networks between miRNAs, target genes and transcription factors are critical for an appropriate balance of gene expression in mammalian melanocytes [26]. Our target gene prediction of 41 differentially expressed miRNAs revealed 359 putative target genes. Has-miRNA-4467/ITGB2, has-miR-941/ITGB2, has-miR-1268b/BCL2L11 and has-miR150-3p/ARHGAP11A regulatory axes need to be further investigated in future experiments. A previous study reported miR-150 is underexpressed in active TB [27]. Its primary target is a negative regulator of NK-cell maturation and a reduction of miR-150 levels may indicate development of fewer mature NK cells, which are early innate effector cells controlling invading pathogens [27]. The positive and negative regulation of miRNA–mRNA pairs found in this study provides important clues for PTB pathogenesis.

We comprehensively investigated GO enrichment and KEGG pathways analysis for differentially expressed mRNA and targets of differentially expressed miRNA. With GO enrichment and pathway analysis together, we

found 41 differentially expressed miRNAs grouped into 298 KEGG pathways and 2565 differential expression mRNAs grouped into 283 KEGG pathways. Shared by the two methods were 282 KEGG pathways, including important pathways in immunity and metabolism pathway such as MAPK signaling, phagosome, calcium and HTLV-I infection signaling pathway. A previous study reported the activation of MAPK signaling in macrophages by nonpathogenic mycobacteria infection which leads to synthesis of microbicidal molecules, including TNF-$\alpha$, which mediates inflammatory immune responses and antipathogens [28]. Mannose-capped lipoarabinomannan from Mtb blocks phagosome maturation by inhibiting a signaling cascade of $Ca^{2+}$, calmodulin and PI3K. The arrest of phagosomal maturation by mannose-capped lipoarabinomannan is an effective mechanism used by mycobacteria for long-term survival in host cells [29]. Protective immunity to *M. tuberculosis* depends on correct function of T cells and their interaction with macrophages. HTLV-1 preferentially infects T cells and causes chronic T-cell dysfunction, which appears to impair the immune response to specific pathogens [30]. HTLV-1 infection may increase an individual's susceptibility to active TB [31]. Our results suggested that differentially expressed miRNAs and mRNAs associated with immune pathways might contribute to PTB pathogenesis.

## Conclusion

Taken together, our results provide novel insights into PTB pathogenesis. The network and pathway information presented here offer insights for the elucidation of detailed functions of mRNAs and miRNAs in PTB pathogenesis. However, the molecular roles that these dysregulated miRNAs and mRNAs play in PTB are not completely understood here. Future studies are needed to explore the potential mechanism of these dysregulated mRNAs and miRNAs in PTB pathogenesis.

### Summary points

- Host immunity is important in regulating progression of tuberculosis infection.
- The functions of mRNAs/miRNAs are closely correlated with pulmonary tuberculosis (PTB) pathogenesis.
- The whole transcriptome analysis and small RNA sequencing were applied to investigate the differentially expressed mRNAs and miRNAs from PTB patients and healthy individuals.
- Two thousand five hundred and sixty-five mRNAs were screened out with significant levels of differential expression comparing PTB patients with healthy individuals (greater than or equal to twofold change; adjusted $p < 0.05$), with 1383 significantly upregulated mRNAs and 1182 significantly downregulated mRNAs.
- Forty-one significantly differentially expressed miRNAs were identified in PTB versus healthy samples (greater than or equal to twofold change; adjusted $p < 0.05$), with 18 upregulated miRNAs and 23 downregulated miRNAs.
- A large number of the differentially expressed mRNAs and miRNAs were related to immune-related pathways, particularly the tuberculosis, phagosome and MAPK signaling pathway.
- Three hundred and fifty-nine potential target genes were identified for 41 differentially expressed miRNAs.
- Many of target genes were enriched to phagosome, calcium and insulin signaling pathway.

### Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/full/10.2217/epi-2018-0028

### Ethical conduct of research

This study was approved by the Human Ethics Committee of the First People's Hospital of Zhangjiagang, Jiangsu Province, China. All patients were informed and declared by written informed consent that their samples could be used for research.

## References

1. Sakamoto K. The pathology of *Mycobacterium tuberculosis* infection. *Vet. Pathol.* 49(3), 423–439 (2012).

2. WHO. Global tuberculosis report (2016). http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf?ua=1

3. Smith I. *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clin. Microbiol. Rev.* 16(3), 463–496 (2003).

4. Zhang X, Li Y, Li X *et al.* Association of the miR-146a, miR-149, miR-196a2 and miR-499 polymorphisms with susceptibility to pulmonary tuberculosis in the Chinese Uygur, Kazak and Southern Han populations. *BMC Infect. Dis.* 15, 41 (2015).

5. Khalilullah SA, Harapan H, Hasan NA, Winardi W, Ichsan I, Mulyadi M. Host genome polymorphisms and tuberculosis infection: what we have to say? *Egypt. J. Chest. Dis. Tuberc.* 63(1), 173–185 (2014).

6. Williams AE. Functional aspects of animal microRNAs. *Cell Mol. Life Sci.* 65(4), 545–562 (2008).

7. Shivdasani RA. MicroRNAs: regulators of gene expression and cell differentiation. *Blood* 108(12), 3646–3653 (2006).

8. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 19(1), 92–105 (2009).

9. Berry MP, Graham CM, McNab FW *et al.* An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 466(7309), 973–977 (2010).

10. Marin ND, Paris SC, Rojas M, Garcia LF. Functional profile of CD4$^+$ and CD8$^+$ T cells in latently infected individuals and patients with active TB. *Tuberculosis* 93(2), 155–166 (2013).

11. Du C, Liu C, Kang J *et al.* MicroRNA miR-326 regulates TH-17 differentiation and is associated with the pathogenesis of multiple sclerosis. *Nat. Immunol.* 10(12), 1252–1259 (2009).

12. O'Connell RM, Rao DS, Chaudhuri AA, Baltimore D. Physiological and pathological roles for microRNAs in the immune system. *Nat. Rev. Immunol.* 10(2), 111–122 (2010).

13. Taganov KD, Boldin MP, Chang KJ, Baltimore D. NF-κB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc. Natl Acad. Sci. USA* 103(33), 12481–12486 (2006).

14. Fu Y, Yi Z, Li J, Li R. Deregulated microRNAs in CD4$^+$ T cells from individuals with latent tuberculosis versus active tuberculosis. *J. Cell Mol. Med.* 18(3), 503–513 (2014).

15. Cui JY, Liang HW, Pan XL *et al.* Characterization of a novel panel of plasma microRNAs that discriminates between *Mycobacterium tuberculosis* infection and healthy individuals. *PLoS ONE* 12(9), e0184113 (2017).

16. Zhang X, Guo J, Fan S *et al.* Screening and identification of six serum microRNAs as novel potential combination biomarkers for pulmonary tuberculosis diagnosis. *PLoS ONE* 8(12), e81076 (2013).

17. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 11(10), R106 (2010).

18. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in *Drosophila*. *Genome Biol.* 5(1), R1 (2003).

19. Pasquinelli AE. Non-coding RNA: MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nat. Rev. Genet.* 13(4), 271–282 (2012).

20. Kim JK, Kim TS, Basu J, Jo EK. MicroRNA in innate immunity and autophagy during mycobacterial infection. *Cell Microbiol.* doi:10.1111/cmi.12687 (2017) (Epub ahead of print).

21. Peter ME. Targeting of mRNAs by multiple miRNAs: the next step. *Oncogene* 29(15), 2161–2164 (2010).

22. Ueberberg B, Kohns M, Mayatepek E, Jacobsen M. Are microRNAs suitable biomarkers of immunity to tuberculosis? *Mol. Cell Pediatr.* 1(1), 8 (2014).

23. Zhou M, Yu G, Yang X, Zhu C, Zhang Z, Zhan X. Circulating microRNAs as biomarkers for the early diagnosis of childhood tuberculosis infection. *Mol. Med. Rep.* 13(6), 4620–4626 (2016).

24. Yi Z, Fu Y, Ji R, Li R, Guan Z. Altered microRNA signatures in sputum of patients with active pulmonary tuberculosis. *PLoS ONE* 7(8), e43184 (2012).

25. Kim JK, Lee HM, Park KS *et al.* MIR144* inhibits antimicrobial responses against *Mycobacterium tuberculosis* in human monocytes and macrophages by targeting the autophagy protein DRAM2. *Autophagy* 13(2), 423–441 (2017).

26. Abdel-Malek Z, Scott MC, Suzuki I *et al.* The melanocortin-1 receptor is a key regulator of human cutaneous pigmentation. *Pigment Cell Res.* 13, 156–162 (2000).

27. Bezman NA, Chakraborty T, Bender T, Lanier LL. miR-150 regulates the development of NK and iNKT cells. *J. Exp. Med.* 208(13), 2717–2731 (2011).

28. Roach SK, Schorey JS. Differential regulation of the mitogen-activated protein kinases by pathogenic and nonpathogenic mycobacteria. *Infect. Immun.* 70(6), 3040–3052 (2002).

29.  Koul A, Herget T, Klebl B, Ullrich A. Interplay between mycobacteria and host signalling pathways. *Nat. Rev. Microbiol.* 2(3), 189–202 (2004).

30.  Suzuki M, Dezzutti CS, Okayama A *et al.* Modulation of T-cell responses to a recall antigen in human T-cell leukemia virus type 1-infected individuals. *Clin. Diagn. Lab. Immunol.* 6(5), 713–717 (1999).

31.  Verdonck K, Gonzalez E, Schrooten W, Vanham G, Gotuzzo E. HTLV-1 infection is associated with a history of active tuberculosis among family members of HTLV-1-infected patients in Peru. *Epidemiol. Infect.* 136(8), 1076–1083 (2008).

32.  Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9(4), 357–359 (2012).

33.  Friedlander MR, Mackowiak SD, Li N, Chen W, Rajewsky N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 40(1), 37–52 (2012).

34.  Mansson R, Tsapogas P, Akerlund M, Lagergren A, Gisler R, Sigvardsson M. Pearson correlation analysis of microarray data allows for the identification of genetic targets for early B-cell factor. *J. Biol. Chem.* 279(17), 17905–17913 (2004).