



Identification and Classification of Differentially Expressed Genes and Network Meta-Analysis Reveals Potential Molecular Signatures Associated With Tuberculosis

Aftab Alam¹, Nikhat Imam^{1,2}, Mohd Murshad Ahmed¹, Safia Tazyeen¹, Naaila Tamkeen³, Anam Farooqui¹, Md. Zubair Malik⁴ and Romana Ishrat^{1*}

OPEN ACCESS

Edited by:

Yusuf Akhter,
Babasaheb Bhimrao Ambedkar
University, India

Reviewed by:

Biju Issac,
Leidos Biomedical Research, Inc.,
United States
Syed Asfarul Haque,
ThermoFisher Scientific and Diamond
Light Source Ltd, United Kingdom

*Correspondence:

Romana Ishrat
romana05@gmail.com

Specialty section:

This article was submitted to
Bioinformatics and
Computational Biology,
a section of the journal
Frontiers in Genetics

Received: 09 April 2019

Accepted: 05 September 2019

Published: 04 November 2019

Citation:

Alam A, Imam N, Ahmed MM,
Tazyeen S, Tamkeen N, Farooqui A,
Malik MZ and Ishrat R (2019)
Identification and Classification of
Differentially Expressed Genes and
Network Meta-Analysis Reveals
Potential Molecular Signatures
Associated With Tuberculosis.
Front. Genet. 10:932.
doi: 10.3389/fgene.2019.00932

¹ Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, New Delhi, India, ² Department of Mathematics, Institute of Computer Science & Information Technology, Magadh University, Bodh Gaya, India, ³ Department of Biosciences, Jamia Millia Islamia, New Delhi, India, ⁴ School of Computational & Integrative Sciences, Jawaharlal Nehru University, New Delhi, India

Tuberculosis (TB) is one of deadly transmissible disease that causes death worldwide; however, only 10% of people infected with *Mycobacterium tuberculosis* develop disease, indicating that host genetic factors may play key role in determining susceptibility to TB disease. In this way, the analysis of gene expression profiling of TB infected individuals can give us a snapshot of actively expressed genes and transcripts under various conditions. In the present study, we have analyzed microarray data set and compared the gene expression profiles of patients with different datasets of healthy control, latent infection, and active TB. We observed the transition of genes from normal condition to different stages of the TB and identified and annotated those genes/pathways/processes that have important roles in TB disease during its cyclic interventions in the human body. We identified 488 genes that were differentially expressed at various stages of TB and allocated to pathways and gene set enrichment analysis. These pathways as well as GSEA's importance were evaluated according to the number of DEGs presents in both. In addition, we studied the gene regulatory networks that may help to further understand the molecular mechanism of immune response against the TB infection and provide us a new angle for future biomarker and therapeutic targets. In this study, we identified 26 leading hubs which are deeply rooted from top to bottom in the gene regulatory network and work as the backbone of the network. These leading hubs contains 31 key regulator genes, of which 14 genes were up-regulated and 17 genes were down-regulated. The proposed approach is based on gene-expression profiling, and network analysis approaches predict some unknown TB-associated genes, which can be considered (or can be tested) as reliable candidates for further (*in vivo/in vitro*) studies.

Keywords: meta-analysis, DEGs, gene ontology, KEGG, network analysis, gene knock-out, LCP

INTRODUCTION

Tuberculosis (TB) is a communicable disease generally caused by the bacterium *Mycobacterium tuberculosis* (MTB). The bacteria typically affects the lungs (pulmonary TB), but other body parts can be also affected (extra pulmonary TB) (Ahmad, 2011). The disease is communicable and spread through the air by expelling out the active MTB while coughing and sneezing (Schnappinger et al., 2003). In 2016, 10.4 million individuals were infected with TB, and 1.7 million died from the disease (including 0.4 million among people with HIV i.e., 40% of HIV deaths were due to TB). TB kills more adults in India than any other infectious disease (In 2016, an estimated 28 lakh cases occurred, and 4.5 lakh people died due to TB). India has the highest burden of both TB and advanced TB (like MDR TB) and second highest of HIV-associated TB. In India, the major challenge to curb the TB is poor primary healthcare system in rural areas due to deregulation of private health care leading to indiscriminate use of first- and second-line TB drugs, poverty, spreading HIV infection, and lack of administrative coordination among government functionary bodies. In our current study, we used meta-analysis of individual raw microarray data (GSE series) deposited in the GEO database which are obtained from various blood cell types (macrophages, monocytes, and CD4+) and cell lines (THP1) of individuals with different datasets (e.g., controls vs. TB disease, control vs. latent TB, latent TB vs. TB disease). We have performed gene-transition study of differentially expressed genes (DEGs) data and text mining between different stages of TB and classified the DEGs in stage-specific manner like normal to latent TB infections, normal to active TB, and latent infection to active TB. It is widely thought that, to know that function of a gene, it must be analyzed in the context of gene interaction network, because gene networks are commonly interpreted as encoding functional information in their connections. So, our study usually focused on the “guilt-by-association (GBA)” presumption which state that physically and functionally linked genes are possibly participating in the same biological pathways having comparable effects on the phenotypes (Lee et al., 2011). The concept of network theory is an imperative method to know the topological properties and the complex-system dynamics that correspond to their functional modules. The complex networks may be classified into four types of networks: (a) scale-free network, (b) small-world network, (c) random network, and (d) hierarchical network. For the biologist, hierarchical network has special interest because it incorporates the mien of modules, and distributed hubs (sparsely) regulate the network. So, we constructed the gene regulatory network and then analyzed its topological properties, because it helps to understand the structure of a network which facilitates in understanding the hidden mechanisms. Further, we identified important network

Abbreviations: MTB, *Mycobacterium tuberculosis*; DEGS, Differentially expressed genes; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; LCP, Local-community-paradigm; MDR, Multidrug-resistance TB; GBA, Guilt by association; MF, Molecular function; BP, Biological process; PC, Protein class; NC, Normal control; LI, Latent infection; ATB, Active TB; FKR, Fundamental key regulators; LAGO, Logically accelerated GO; GEO, Gene Expression Omnibus; GSE, GEO series; RMA, Robust multiarray average; LIMMA, Linear model for microarray analysis; CN, Common neighbors; HE, Hamiltonian energy; CPM, Constant Potts model; LCP-DP, LCP-decomposition-plot.

modules which contain fundamental key regulators that have a fundamental importance.

MATERIAL AND METHOD

Inclusion Criteria for Differentially Expressed Genes

A set of 11 microarray data sets were selected from the NCBI GEO repository database (<http://www.ncbi.nlm.nih.gov/geo/>). The datasets include GSE54992 (Cai et al., 2014), GSE52819 (Verway et al., 2013), GSE64335 (Bouttier et al., 2016), GSE11199 (Thuong et al., 2008), GSE98750 (Lavalett et al., 2017), GSE78233 (Cheng et al., 2017), GSE57736 (Guerra-Laso et al., 2015), GSE16250 (Reyes et al., 2015), GSE27882 (Liu et al., 2012), GSE57028 (Salamon et al., 2014), and GSE83456 (Blankley et al., 2016). The background data correction and data normalization were done by the robust multiarray average (RMA) in R affy and lumi packages to ensure unbiased and dysregulated gene expression data. The RMA method, which performs quantile normalization, was used to minimize inconsistencies due to normalization of the each Affymetrix GSE series. The RMA method was chosen over others due to its fine differential change detection and stable variance on a log scale and minimizes the false positive results. The specificity and sensitivity of RMA method are good during fold-change calculation to identify DEGs. Similarly, we have used lumi pipeline (Bioconductor package) which is exclusively developed to analyze Illumina data (BeadChip). It checks the data quality and data normalization and stabilizes the data variance. The gene expression data (GSE57736 and GSE27882) which is generated by Agilent platforms were already present as normalized data. In the present study, we used linear model for microarray analysis (LIMMA) package which is a highly recommended method to measure the differential expression of genes. It not only calculates simple t-test but also calculates moderate t-test and f-test by applying the *Empirical Bayes* approach and reducing the standard errors and gives us steady and reproducible outcomes even with a less quantity of arrays. So, limma package (Smyth, 2004) was used to identify the DEGs among the healthy control, latent TB, and active TB. We set the criteria to select significant differentially expressed genes i.e., the adjusted p-value was <0.05, and the fold change was >1.5. We have used the BRCW [Bioinformatics & Research Computing website (<http://jura.wi.mit.edu/bioc/tools/compare.php>)] to select the DEGs which is common in at least two datasets of gene expression profile. Doing this, we became more specific in the selection of DEGs and the chances of biased data compilation became negligible. According to the corresponding correlation between the probe and gene from the data, the probe numbers of the expression profile were changed into the corresponding gene symbols using the Synergizer database (Huang et al., 2009).

Gene Classification, Ontology, and Pathway Analysis of Degs

To know the significance of the identified DEGs, we have categorized them by GO-molecular function, GO-biological process, and

protein class using the Protein ANalysis THrough Evolutionary Relationships (PANTHER v.13.0) Classification System and analysis tools and DAVID database (<https://david.ncifcrf.gov/>) to enrich the given set of DEGs to possible GO terms (Mi et al., 2013) (Huang et al., 2009). The PANTHER overrepresentation study (Fisher's exact with FDR multiple test correction) was used to search the data against the PANTHER, and GO databases and p-values were set according to Bonferroni correction.

The GO analysis (gene ontology) is the useful method for annotation of genes and its products and characterization of biological attributes for high-throughput genome or transcriptome data (Ashburner et al., 2000). The differentially expressed genes among "active-TB" and "latent TB" on compared with "normal condition" were over-represented in various GO classes. The gene ontology provides and visualizes us the basic terms subdivided into three important categories, namely, BP (Biological process), MF (molecular functions), and biological pathways among these groups.

Gene Transition Among Different Stages of TB

In order to get the behavior of normal gene expression perturbation, we tried a normal way for finding the associated genes while moving from one stage to another. In all the transitions that we took into consideration have been provided by a list of up- and down-regulated genes. These gene(s) in both the cases (i.e., stage from which is transferred to the targeted stage) have its own meaning. This meaning to a gene(s) regulation gives us an opportunity to say something about the ongoing mechanobiology inside the cell. So, to observe these transitions, we framed our study in such a way (based on the data we've got) which considers every possible transition in view. These transitions are discussed in brief as follows. In this study, we made comparison of the gene-expression profiles among individuals with normal conditions, latent infection, and active TB. Thus, we observed the expression of genes from normal to different stages of the TB and try to arrest those genes which play a key role in susceptibility to TB.

- **Normal to Latent Infections:** In this section, we have taken those DEGs which are involved in between normal and latent TB conditions. In LTBI, the bacteria remain in inactive form for many years (years to decades) before transforming into TB disease. In this study, we have studied several BPs and important pathways that lead to the latent infection for identification of latently infected individuals.
- **Normal to Active TB:** In this section, we have taken those DEGs which are differentially expressed in TB disease condition as compared to healthy control and recognized those immune process and pathways which are prominent in TB disease.
- **Latent Infection to Active TB:** In this section, we have taken those DEGs which are involved in between latent TB and active TB diseases. The individual with LTBI eventually reactivates and becomes infectious, seriously influencing epidemiological situation. Mechanisms of MTB transition to dormancy and TB reactivation are inadequately understood, and biomarkers of latency remain largely mysterious (Kondratieva et al., 2014).

Topological Characterization of Networks

The classified genes of various stages of TB were used to construct their regulatory networks. The networks were built by STRING database (<https://string-db.org/>) and then visualized in Cytoscape v3.4 (Shannon et al., 2003). As the network was built, the first and foremost basic analysis are its topological properties. Topological analysis helps to understand the structure of a network which facilitates in understanding the hidden mechanisms. The following networks properties were analyzed to seek the important behaviors of the network:

- **Degree distribution:** The degree (k) of a node (gene) in a biological network is the number of links with other nodes, the probability distribution of these degrees called degree distribution ($P[k]$).

$$P(k) = \frac{n_k}{N} \quad (1)$$

where n_k = no. of nodes with degree k .

N = the total number of nodes in the network.

- **Neighborhood connectivity:** A node (gene) has number of neighbors and is considered as a connectivity of node. So, neighborhood connectivity ($C_N [K]$) is:

$$C_N(k) = \sum_q qP\left(\frac{q}{k}\right) \quad (2)$$

where $P\left(\frac{q}{k}\right)$ = the conditional probability.

- **Clustering co-efficient:** The ratio of number of edges (e_i) between the node's neighbor or maximum numbers of edges that could possibly exist between the nodes. So, the total network cluster coefficient is the average of cluster coefficient of all nodes (i^{th}) in the network.

$$C(k_i) = \frac{2e_i}{k_i(k_i - 1)} \quad (3)$$

- **Betweenness centrality:** In the complex network, a node's betweenness centrality (CB) represents the prominence of information flow through one node to another *via* shortest path (Brandes, 2001) (Mason and Verwoerd, 2007). From node (i) to node (j), the geodesic paths are shown by " $d_{ij}(v)$ " passing *via* node " v " and " d_{ij} ".

$$C_B(v) = \sum_{i,j,i \neq j \neq k} \frac{d_{ij}(v)}{d_{ij}} \quad (4)$$

- **Closeness centrality:** In the network, how quick info is circulated from one node (i) to another (j) is measured by closeness centrality (C_C). Closeness centrality is given by:

$$C_C(i) = \frac{n}{\sum_j d_{ij}} \quad (5)$$

- **Eigenvector centrality:** In the network, eigenvector centrality of a node “ i ($C_E[i]$)” is proportionate to the total of i 's neighbor centralities (Bonacich, 1987).

$$C_E(i) = \frac{1}{\lambda} \sum_{j \in nn(i)} v_j \quad (6)$$

where $nn(i)$ = closest neighbors of nodes (“ i ”).
 λ = eigenvalue of the eigenvector.
 $v_i = “Av_i = \lambda v_i”$ where “ A ” (adjacency matrix).

Community Analysis: Leading Eigenvector Method

In hierarchical network, to distinguish the nature of modular and its properties is important to explain the activities of network at various levels of hierarchy. In our study, the leading eigenvector method (LEV) (Newman, 2006a; Newman, 2006b) was used to detect the communities in R from package “igraph” (Gabor and Nepusz, 2006) (the community detection script is given **Supplementary Data 3**).

- **Modularity:** Modularity determines to measure the strength of division of a network into clusters or communities (Newman and Girvan, 2004).

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(C_i, C_j) \quad (7)$$

where m = Total no. of edges.
 A_{ij} = adjacency matrix of size “ $i \times j$ ”
 k = degrees.
 δ = function yields 1 if nodes “ i ” and “ j ” are in the same community.

Genes Tracing

To access the regulation of network, we first tried to find out the most influential nodes (genes) within the network. The gene tracing (up to motif level) was done purely on the appearance of the respective genes in various submodules obtained from the clustering. Then, these genes were used to get the picture of changes in the network organization in their absence.

Hub Gene Knockout

To know the changes of organization within the complex network in the absence of most influencing gene (node), we must remove constructed leading hubs (rich clubs) in the networks. We consecutively eliminated all the important hubs (one by one) from each network and measured the topological properties of this reorganized network to observe the regulating abilities of these hubs by calculating the degree of structural change due to their absence.

LCP-DP Approach to Estimate the Network Compactness

The LCP-decomposition-plot (LCP-DP) provides a way to characterize the topological properties of the network in 2D space of common neighbors (CN) index of connecting nodes and local community links (LCL) of each pair of interacting nodes in the network (Cannistraci et al., 2013). The LCP correlation defined by:

$$LCP - corr = \frac{cov(CN, LCL)}{\sigma_{CN} \sigma_{LCL}} \text{with } CN > 1 \quad (8)$$

where $cov(CN, LCL)$ = the covariance among LCL and CN .
 σ_{LCL} and σ_{CN} = standard deviations.
 (Note: MATLAB script is given in **Supplementary Data 2**).

Hamiltonian Energy Estimation: Energy Distribution in the Network

The HE is used to organize a network at a certain level by following the formalism of *constant Potts model* (Traag et al., 2011; Traag, 2013). The energy distribution at the global and modular levels of the network is given by HE. HE of a network can be measured by:

$$H^{[c]} = - \sum_c [e_c - \gamma n_c^2] \quad (9)$$

where e_c = no. of edges.
 n_c = no. nodes.
 “ c ” and “ γ ” = the resolution parameter

RESULTS

Differentially Expressed Genes (DEGs) Classification and Overrepresentation Analysis

A total of 5, 680 differentially expressed genes (DEGs) were identified after the extensive analysis of all the 11 GSE series, of which 2,660 were up-regulated, and 3,020 were down-regulated genes. These differentially expressed genes were clustered according to “GO-MF (molecular function),” “GO-BP (biological process),” and “PANTHER protein class” shown in **Figure 1**. All these predictive DEGs showed a broad range of protein classes which involved in various processes. The helicases and nucleases were found within the “nucleic acid binding” protein class. The “enzyme modulator” category features kinase, G-protein, phosphatase, and protease modulators. The structural motif and nuclear hormone receptors are part of the “transcription factor” protein class. The “hydrolases” is a sub-category of proteases and phosphatases. The “receptor” protein class includes cytokine receptors, protein-kinase receptors, ligand-gated ion channels, nuclear-hormone receptors, and G-protein-coupled receptors. Besides of these above protein classes, signaling molecules, transferase, oxidoreductase, and

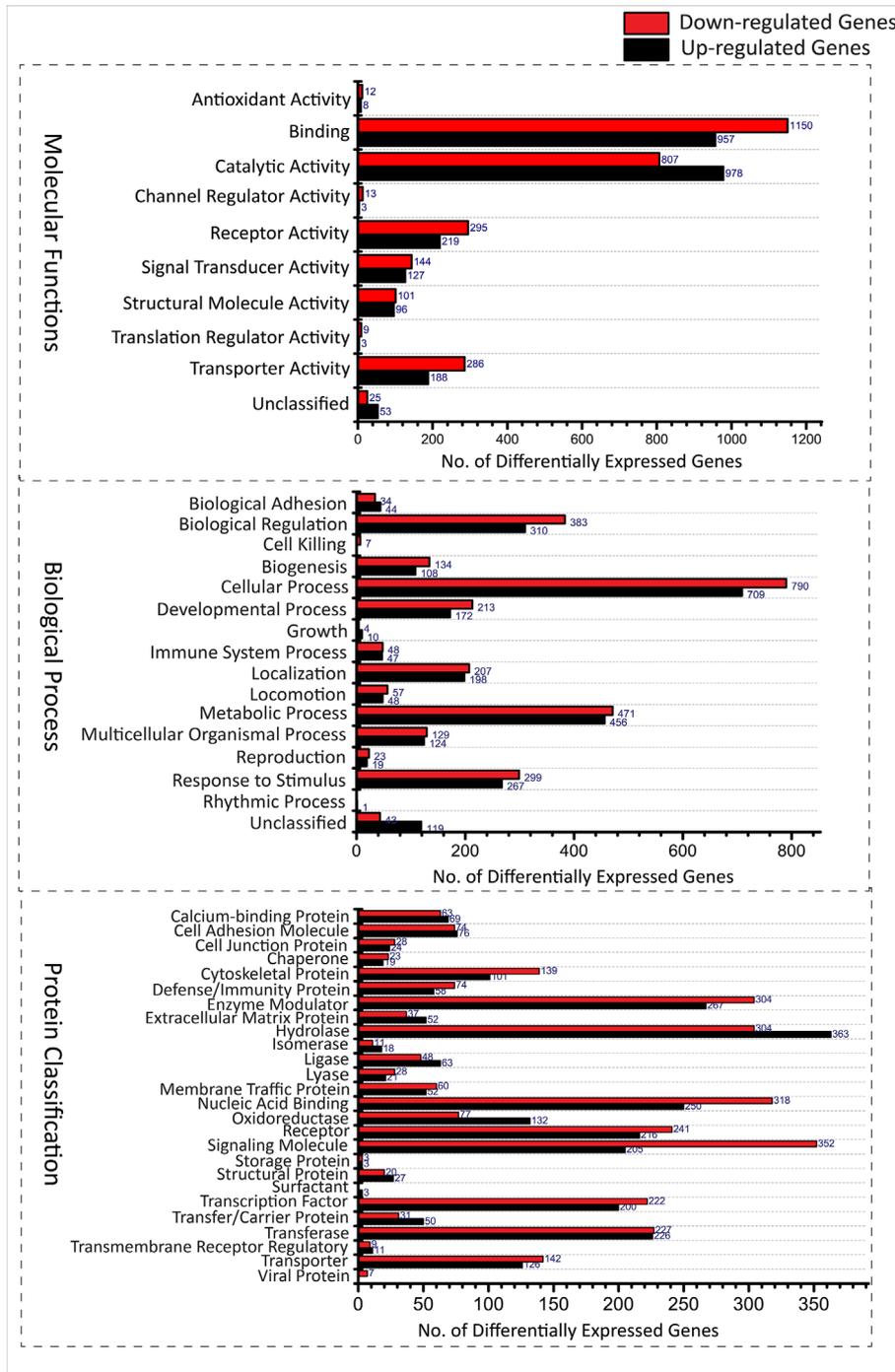


FIGURE 1 | Functional Classification of Differentially expressed genes from various GSE series associated with TB disease according to **(A)** Molecular functions [Transporter activity (U¹⁸⁸,D²⁸⁶), Translation regulator activity(U⁰³,D⁰⁹), Catalytic activity(U⁹⁷⁸,D³⁰⁷), Channel regulator activity(U⁰³,D¹³) Receptor activity(U²¹⁹,D²⁹⁵), Signal transducer activity(U¹²⁷,D¹⁴⁴), Antioxidant activity(U⁰⁸,D¹²), Structural molecule activity(U⁹⁶,D¹⁰¹), Binding(U⁹⁵⁷,D¹¹⁵⁰) and unclassified(U⁵³,D²⁵)], **(B)** Biological processes [(cellular component organization or biogenesis(U¹⁰⁸,D¹³⁴), cellular process(U⁷⁰⁹,D⁷⁹⁰), localization(U¹⁹⁸,D²⁰⁷), reproduction(U¹⁹,D²³), biological regulation(U³¹⁰,D³⁸³), response to stimulus(U²⁶⁷,D²⁹⁹), developmental process(U¹⁷²,D²¹³), rhythmic process(U⁰¹,D⁰⁰), multicellular organismal process (U¹²⁴,D¹²⁹), locomotion(U⁴⁸,D⁵⁷), biological adhesion(U⁴⁴,D³⁴), metabolic process(U⁴⁵⁶,D⁴⁷¹), growth(U¹⁰,D⁰⁴), immune system process(U⁴⁷,D⁴⁸), cell killing (U⁰⁰,D⁰⁷) and unclassified(U¹¹⁹,D⁴²)] and **(C)** Protein classes [(extracellular matrix protein(U⁵²,D³⁴), cytoskeletal protein(U¹⁰¹,D¹³⁹), transporter(U¹²⁶,D¹⁴²), transmembrane receptor regulatory/adaptor protein(U¹¹,D⁰⁹), transferase(U²²⁶,D²²⁷), oxidoreductase(U¹³²,D⁷⁷), lyase(U²¹,D²⁸), cell adhesion molecule(U⁷⁶,D⁷⁴), ligase(U⁸³,D⁸⁸), nucleic acid binding(U²⁵⁰,D³¹⁸), signaling molecule(U²⁰⁵,D³⁵²), enzyme modulator(U²⁰⁷,D³⁰⁴),calcium-binding protein(U⁸⁹,D⁸³), defense/immunity protein(U⁵⁸,D⁷⁴), hydrolase(U³⁶³,D³⁰⁴), transfer/carrier protein(U⁵⁰,D³¹), membrane traffic protein(U⁵²,D⁶⁰), transcription factor(U²⁰⁰,D²²²), chaperone(U¹⁹,D²³), cell junction protein(U²⁴,D²²), surfactant(U⁰³,D⁰⁰), structural protein(U²⁷,D²⁰), storage protein(U⁰³,D⁰³), isomerase(U¹⁸,D¹¹), receptor(U²¹⁶,D²⁴¹)]. (U, Up-regulated; D, Down-regulated).

transporter are also most abundant protein classes. The two most abundant GO-biological process groups—“metabolic process” and “cellular process” which are not astonishing as these processes carry those genes which involved in the most basic life processes. The “cellular process” includes cell cycle, cell–cell signaling, cell component movement, cytokinesis, and proliferation. The “metabolic process” section includes metabolism of lipid, protein, carbohydrate, cellular amino acid, and nucleobase-containing compound metabolism. The “biological regulation” includes metabolism, cell cycle, the regulation of apoptosis, catalytic activity, translation, and homeostasis. Further, these biological process responses to stimulus, localization, and developmental process also represent the significant number of proteins; the complete categories details are given in **Supplementary Data 1**. To know the probability of largely occupied protein classes and GO categories among the differentially expressed genes, we used PANTHER's overrepresentation analysis. When we compared with the reference genome, we found that any of the richest classes are overrepresented in the data (**Table 1**). The most abundant protein classes “chemokine,” “cytokine,” “hydrolase,” “ribosomal protein,” and “RNA-binding protein” were enriched along with the classes “signaling molecule” and “cell adhesion molecule.” The highly populated GO-biological processes were enriched in “cytokine-mediated signaling pathway,” “response to external stimulus,” “immune response,” “locomotion,” “signal transduction,” “cell communication,” “developmental process,” “cellular process,” and “MAPK cascade.” Similarly, the most abundant GO-molecular functions were enriched in “chemokine activity,” “cytokine activity,” “cytokine receptor binding,” “oxidoreductase activity,” and “receptor binding.”

Gene-Transition and GO Enrichment Analysis

All the 5,680 differentially expressed genes (DEGs) were divided into three different groups, normal vs. latent infection, normal vs. active TB, and latent infection vs. active TB. We filtered out a total number of 488 DEGs which are listed in (**Table 2**), and then, we also identified few genes which are commonly differentially expressed among various stages of TB (including NC vs. LI, NC vs. ATB, and LI vs. ATB) shown in **Figure 2**. Further, we isolated up- and down-regulated genes from each stage and performed GO-enrichment analysis using DAVID tool (v6.7). It is well preferred annotation database which used novel algorithms to extract biological information from large gene lists. The DEGs are enriched with a certain biological process or molecular function that are given in **Table 3**. Further, the most enriched pathways associated with these differentially expressed genes were identified, given in **Table 4**.

- **Normal to Latent infections:** In the early infection (latent infection), we have identified 28 genes which are differentially expressed in the host system, of which 12 genes (*IER5L*, *MS4A6A*, *DOK2*, *FZD2*, *NCK1-AS1*, *SNHG12*, *NLRC4*, *XPO7*, *SMA4*, *CD36*, *AFF1*, and *NDUFS8*) were up-regulated, and 16 genes (*IL1A*, *IL6*, *ACOD1*, *IL1B*, *ELOVL7*, *PTGS2*, *EREG*, *F3*, *IFIT1*, *TNF*, *KANK1*, *CCL4*, *CXCL11*, *PTX3*, *IRAK2*, and *AREG*) were down-regulated. The GO analysis of these DEGs

TABLE 1 | Overrepresented PANTHER protein class and GO ontology categories of all differentially expressed genes.

PANTHER Protein Classes	P-value	FDR
Chemokine	2.14E-08	2.29E-06
Cytokine	6.71E-08	4.79E-06
Hydrolase	5.82E-06	3.11E-04
Ribosomal protein	1.44E-05	6.15E-04
RNA-binding protein	2.36E-05	8.42E-04
Signaling molecule	1.46E-04	4.45E-03
Cell adhesion molecule	2.97E-04	7.95E-03
Protease	4.81E-04	1.14E-02
Receptor	1.52E-03	3.25E-02
GO-Molecular Functions		
Chemokine activity	1.19E-06	7.52E-05
Cytokine activity	1.18E-06	1.12E-04
Cytokine receptor binding	5.29E-05	1.67E-03
Oxidoreductase activity	4.85E-04	1.02E-02
Receptor binding	5.61E-04	1.07E-02
Hydrolase activity	3.87E-06	1.84E-04
Catalytic activity	9.50E-08	1.81E-05
Protein binding	1.20E-04	3.25E-03
Structural constituent of ribosome	1.66E-04	3.95E-03
GO-Biological Processes		
Cytokine-mediated signaling pathway	3.45E-08	4.21E-06
Response to external stimulus	3.75E-08	3.05E-06
Response to interferon gamma	1.65E-07	1.01E-05
Immune response	2.10E-07	1.02E-05
Sensory perception of chemical stimulus	2.15E-07	8.72E-06
Locomotion	3.38E-07	1.03E-05
Signal transduction	5.32E-07	1.44E-05
Immune system process	3.39E-06	8.27E-05
Cell communication	5.35E-06	1.19E-04
Cell proliferation	1.09E-05	2.23E-04
Response to biotic stimulus	1.59E-05	2.99E-04
Intracellular signal transduction	1.74E-05	3.03E-04
Death	4.17E-05	6.79E-04
Cell death	4.17E-05	6.36E-04
Cellular component movement	4.60E-05	6.61E-04
Developmental process	5.75E-05	7.80E-04
Apoptotic process	6.05E-05	7.77E-04
Cellular process	9.43E-05	1.15E-03
Sensory perception	9.74E-05	1.13E-03
Response to stress	1.04E-04	1.15E-03
RNA metabolic process	1.93E-04	2.04E-03
MAPK cascade	2.01E-04	2.05E-03
Response to stimulus	2.05E-04	2.00E-03
Cellular defense response	2.12E-04	1.99E-03
Endocytosis	3.17E-04	2.86E-03
Receptor-mediated endocytosis	7.23E-04	6.30E-03
Lipid metabolic process	9.30E-04	7.82E-03
Negative regulation of apoptotic process	1.01E-03	8.21E-03
Behavior	1.04E-03	8.15E-03
Localization	1.17E-03	8.90E-03
Regulation of catalytic activity	2.04E-03	1.51E-02
Sulfur compound metabolic process	2.06E-03	1.48E-02
Cell adhesion	2.78E-03	1.94E-02
Biological adhesion	2.78E-03	1.89E-02
Cell surface receptor signaling pathway	3.44E-03	2.27E-02
Cellular component biogenesis	3.88E-03	2.49E-02
Cell–cell adhesion	5.00E-03	3.13E-02
Cellular amino acid catabolic process	6.59E-03	4.02E-02
Regulation of molecular function	6.62E-03	3.94E-02
GO-Cellular Components		
Extracellular region	6.21E-05	1.96E-03
Extracellular space	5.48E-04	1.15E-02

(Continued)

TABLE 1 | Continued

GO-Cellular Components	P-value	FDR
Nucleus	2.01E-03	3.17E-02
Ribonucleoprotein complex	8.32E-06	5.24E-04
Nucleolus	3.28E-03	4.13E-02

Overrepresentation was determined by calculating the probability that the number of differentially expressed genes belonging to a category is larger or smaller than what would be expected based on a reference human genome. P-values are adjusted using a Bonferroni correction.

are enriched in “immune response,” “inflammatory response,” “chemokine activity,” “cellular process,” “biological regulation,” “regulation of metabolic process,” “protein binding,” “catalytic activity,” “bindings” etc. On the pathway analysis, most of the DEGs were found to be down-regulated (beneficial to pathogen) and enriched in very important pathways like toll-like receptors,

NF-kappaB signaling, cytokine–cytokine receptor interaction, MAPK signaling pathway, TB, and TNF signaling.

- **Normal to Active TB diseases:** A total of 266 differentially expressed genes (DEGs) were identified, of which 149 were up-regulated, and 117 were down-regulated among the various cases. In this condition, the DEGs were enriched in “inflammatory response,” “immune response,” “signal transduction,” “response to stimulus,” “apoptotic process,” “cellular process,” “metabolic process,” “binding,” “catalytic activity,” “receptor activity” etc. For the up-regulated genes, the most abundant pathways were “cytokine–cytokine receptor interaction,” “chemokine signaling,” “toll-like receptor signaling pathway,” “NF-kappa B signaling pathway,” “transcriptional misregulation in cancer,” and “pathways in cancer.” Similarly, for down-regulated genes, the most enriched pathways were “cell cycle,” “chemokine signaling pathway,” “NF-kappa B signaling pathway,” “TNF

TABLE 2 | The total number of DEGs from various GSE series associated with TB disease.

Normal to Latent Infection	
Up-regulated	Down-regulated
IER5L, MS4A6A, DOK2, FZD2, NCK1-AS1, SNHG12, NLRC4, XPO7, SMA4, CD36, AFF1, NDUFS8	IL1A, IL6, ACOD1, IL1B, ELOVL7, PTGS2, EREG, F3, IFIT1, TNF, KANK1, CCL4, CXCL11, PTX3, IRAK2, AREG
Normal to Active TB	
Up-regulated	Down-regulated
GBP5, ISG15, SAMD9L, SERPING1, ANKRD22, ETV7, EPSTI1, GBP4, RSAD2, AIM2, IFI44, IFIT3, FGL2, FYB, MND4, PAX5, OAS3, OAS1, IFI6, TNFSF10, UBE2L6, XAF1, STAT1, BST2, IFI35, STAT2, IFI44L, TRIM22, IFIH1, IFITM1, ATF3, BATF2, IFITM3, GBP1P1, RTP4, FCGR1B, C1QB, CEACAM1, FBXO6, SAMD4A, FRMD3, CMPK2, SELL, CFH, TLR8, LGALS3BP, SRGAP2, SECTM1, NCF1, SIGLEC1, APOL1, TRIM14, MB21D1, CARD16, FGD2, RNF213, CD163, PML, OAS2, OR52K3P, LY6E, RABGAP1L, P2RX7, NRG1, FBXO32, TYMP, PSMB9, NCF1C, PNPT1, CXCL13, GMPR, LAMP3, HESX1, C3AR1, STAT4, CXCL11, IFIT2, STAP1, ZC3H12A, RCAN1, HERC6, CCL20, CD83, CRIM1, GCH1, OASL, DLL4, MX2, EIF2AK2, EBI3, AXL, MGLL, CD80, IFIT5, IFIT1, CCL8, NFE2L3, PLSCR1, ICAM1, CXCL9, SLAMF7, CXCL10, HERC5, VCAM1, DDX58, NFKB1, SAMD9, IGFBP3, CCL3, CD274, BIRC3, IRF1, TAP1, PARP14, TMEM140, WARS, CASP1, GBP2, PSTPIP2, PARP9, RNASE6, FAM129C, FZD2, CD36, LRRK2, MS4A6A, CCR2, NAIIP, FCRL2, P2RY13, CLEC7A, PCDH9, CD300LF, CLEC4A, C10orf54, BAIAP2-AS1, C1orf162, SORT1, JAK2, VAMP5, SCO2, ODF3B, PSME2, LOC101930164, P2RY14, GBP1, GBP6, CARD17, FCGR1A	RNF141, SLC25A37, ID3, EMP2, SKP2, SLC2A3, SUN1, KIT, OLR1, FLNB, CCDC14, GAPT, DHRS9, IQGAP3, SESN3, GINS4, HIST1H4C, CD44, KANK1, KCNJ2, TNIP3, MIR146A, CXCL3, LOC644090, MSC, SOD2, NLRP3, SERPINB9, TNFRSF9, ARL5B, IL24, ADORA2A, PHLDA1, MYO10, CXCL8, DNAAF1, MIR3945HG, NR3C1, TNFAIP6, SERPINB2, TNF, MAP3K8, IL1A, AGO2, CSF3, SPAG9, KYNU, LOC440934, WNT5A, DENND4A, ACOD1, PTGS2, OSM, CCL4, PFKFB3, EREG, ITGB8, PTX3, IL36G, G0S2, SLC7A11, ZC3H12C, TNFAIP3, IL6, CCRL2, FERMT2, SLC04A1, SGPP2, FOSL2, CCL23, FLT1, SERPINB8, NUP98, SLC35F5, MN1, DDIT4, NAMPT, IRAK2, IL10, SLC7A5, AK4, CXCL2, UPB1, CEMIP, ADGRG2, FEZ1, THBS1, LACC1, CXCL1, TRAF1, PHLDA2, HEY1, LRP12, UBTD2, SLC39A8, PLPP3, SLC7A1, ATXN1, KMO, FNDC3B, IL1B, C11orf96, F3, PSEN1, BCAT1, GEM, TFP12, PLAUR, MAFF, TRIM36, ZNF697, INSIG1, DPYSL3, ATP2B1, NCR3LG1, MAMLD1, ZNF540
Latent infection to Infection to Active TB	
Up-regulated	Down-regulated
CLEC12B, CD36, CORO1B, SIGLEC16, LINC00484, AK5, MS4A6A, GPBAR1, RTN1, CREB5, DPYD, LDB2, FCN1, LRRK2, RASSF4, ANXA4, LPCAT2, SKAP2, CPPED1, RNASE2, PLSCR3, CLEC12A, BST1, FGD2, RAB3D, FGL2, PYCARD, CEBPA, MND4, CD33, PRAM1, LILRA1, SLC39A11, TNFSF13, SAMHD1, DIAPH2, FAR2, MSRB1, TBCK, FARS2, C14orf159, MSRB2, ATG16L2, DPYSL2, AIF1, HK3, SIGLEC7, FYB, RPS6KA4, TBC1D5, C10orf11, AGTRAP, PYGL, CARD9, NAGA, SLC9A9, C1RL, STX8, MTHFD1, KCTD12, CBR1, ASCL2, CPNE8, MBNL3, ANXA6, CALML4, HSDL2, SLC22A18, KDM1B, IDH1, DNAJC10, TBXAS1, SCLT1, HSD17B4, MGST2, NAIIP, JAML, ENTPD1, ASGR1, BLVRB, AOA, NIPAL2, NAAA, RAB24, TST, COMT, COMMD10, CYFIP1, TALDO1, ULK2, HDAC9, RBCK1, CEACAM4, OBF1, FUCA2, NREP, STX10, AKR7A2, PLOD1, TRIOBP, QDPR, FAM172A, CDK19, DPAGT1, PARVG, CLEC4A, SSBP4, PNKP, FBXL5, ASRGL1, CARS2, ATP11A, PLXNC1, TSPO, ARHGEF6, AGPAT3, HEXDC, PDSS2, PGM2, PRKCB, H2AFY, S100A9, SNX15, NINJ2, MAP4, PSTPIP1, GSTK1	IL1A, ITGB8, TFP12, COL1A1, MET, SERPINB2, GPRC5A, SLC7A11, CYP1A1, RAB7B, B3GALT2, DUSP4, CXCL1, MMP19, THBS1, HEY1, PPP1R10, ADAMDEC1, PDE4DIP, CCRL2, LMNA, ZC3H12C, RGS13, CXCL5, PMEPA1, SKIL, LACC1, DPYSL4, ABCG1, AHRH, MIR155, EPHA4, CD109, MIR3945HG, NCR3LG1, WHRN, FAM177A1, SLC39A8, HEY2, C11orf96, KMO, CYP1B1, DUSP16, NFKB1, LZTS3, CXCR3, PLIN2, BIRC3, DOCK4, EPB41L3, MYO5C, ZHX2, CD58, P2RY10, CD82, MAPK8, PARD6G, NAB1, ABCB4, AMPD3, STK38L, SPECC1L, IL10RB-AS1, BANP, ETV6, PDCD4, MPZL3, CASC7, GGA2

These genes are classified into various categories of TB like normal to latent infections that contain 28 genes, normal to active TB diseases that contain 265 genes, and latent infection to active TB that contains 195 genes in this category.

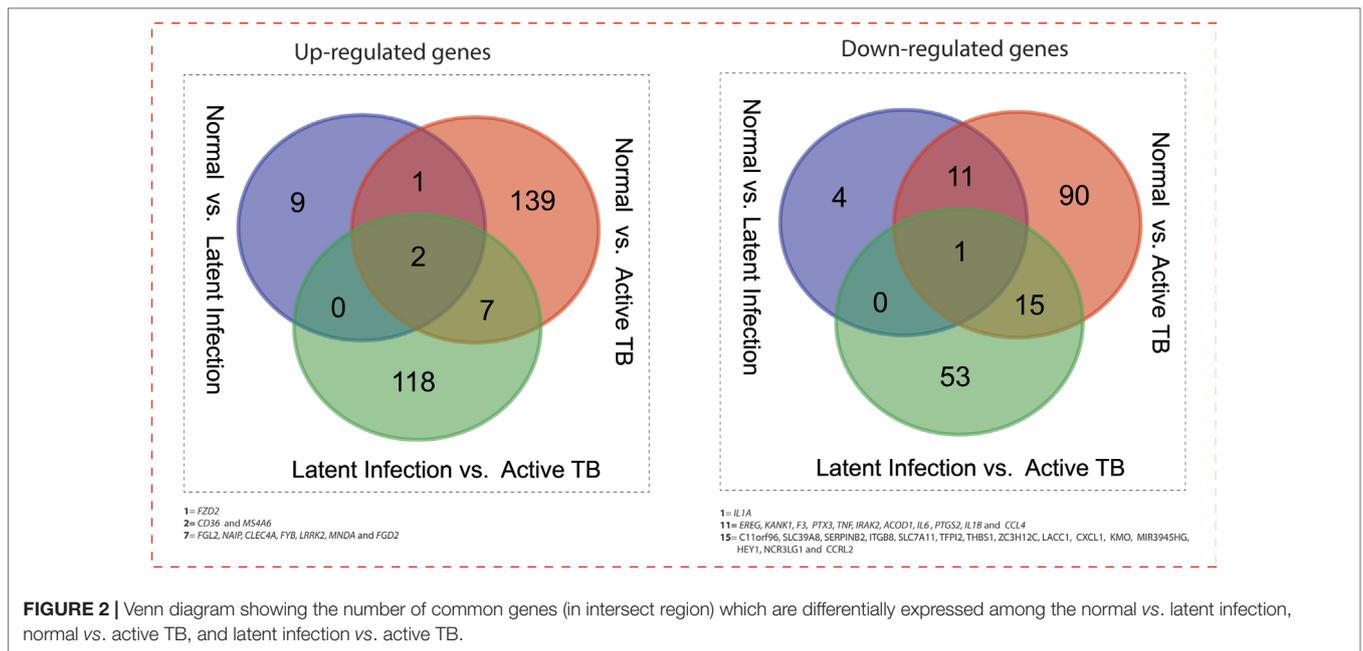


TABLE 3 | Gene set enrichment analysis of differentially expressed genes (DEGs) among active TB, LTBI, and normal condition.

Normal to Latent Infection	Normal to Active TB	Latent Infection to Active TB
Immune response	Inflammatory response	Response to stimulus
Inflammatory response	Immune response	Biological regulation
Transcription factor activity	Signal transduction	Localization
Transferase	Apoptotic process	Cell adhesion
Transporter	GTPase activity	Protein phosphorylation
MAPK cascade	Biological regulation	Immune response
Chemokine activity	Localization	Oxidoreductase activity
Cellular process	Angiogenesis	Hydrolase activity
Biological regulation	Cell adhesion	DNA-templated
Response to stimulus	Cellular process	Angiogenesis
Metabolic process	Response to stimulus	Binding
Protein binding	Metabolic process	Catalytic activity
Catalytic activity	Binding	Receptor activity
Bindings	Catalytic activity	Transporter activity
Signaling molecule	Receptor activity	Structural molecular activity
		Signal transduction
		Cellular process
		Metabolic process

signaling pathway,” “PI3K-Akt signaling pathway,” “metabolic pathways,” and “MAPK signaling pathway.”

- **Latent infection to Active TB Disease::** A total of 127 differentially expressed genes (DEGs) were found to be up-regulated, and 69 were down-regulated. These differentially expressed genes were enriched in “binding,” “catalytic activity,” “receptor activity,” “transporter activity,” “cellular and metabolic process,” “biological regulation,” “immune response,” and “oxidoreductase activity” including others. The most abundant enriched pathway for up-regulated genes belonged to “metabolic pathway,” “NOD-like receptor signaling pathway,” “regulation of actin cytoskeleton,” “biosynthesis of

antibiotics,” “metabolism of xenobiotics by cytochrome p450,” “thyroid hormone synthesis,” “regulation of autophagy,” and “peroxisome,” and for down-regulated genes, the enriched pathways were “chemokine signaling pathway,” “TNF signaling pathway,” “RAP1 signaling pathway,” “PI3K_AKT signaling,” “metabolic process,” “apoptotic,” “focal adhesion,” “MAPK signaling pathway,” “microRNAs in cancer,” and “amoebiasis.”

Gene Interaction: Hierarchical Scale-Free Network

The classified genes of various stages of TB (Table 2) were used to construct their regulatory network. We constructed six networks for Up and Down-regulated genes separately. The topological parameters of the networks obey power law distributions. The probability of clustering co-efficient $C(k)$, degree distributions $P(k)$, and neighborhood connectivity $C_N(k)$ exhibits fractal nature as shown in Figure 3. The results for all the networks are summarized as follows:

Normal to latent Inf.	Normal to active TB
$\begin{bmatrix} P \\ C \\ C_N \end{bmatrix} \sim \begin{bmatrix} k^{-A} \\ k^{-B} \\ k^{+C} \end{bmatrix} \rightarrow \begin{bmatrix} 0.418 \\ 0.218 \\ 2.099 \end{bmatrix}$	$\begin{bmatrix} P \\ C \\ C_N \end{bmatrix} \sim \begin{bmatrix} k^{-A} \\ k^{-B} \\ k^{+C} \end{bmatrix} \rightarrow \begin{bmatrix} 0.459 \\ 0.153 \\ 1.755 \end{bmatrix}$
$\begin{bmatrix} P \\ C \\ C_N \end{bmatrix} \sim \begin{bmatrix} k^{-A} \\ k^{-B} \\ k^{+C} \end{bmatrix} \rightarrow \begin{bmatrix} 0.315 \\ 0.413 \\ 1.33 \end{bmatrix}$	$\begin{bmatrix} P \\ C \\ C_N \end{bmatrix} \sim \begin{bmatrix} k^{-A} \\ k^{-B} \\ k^{+C} \end{bmatrix} \rightarrow \begin{bmatrix} 0.533 \\ 0.354 \\ 0.983 \end{bmatrix}$
Latent Inf. to active TB	
$\begin{bmatrix} P \\ C \\ C_N \end{bmatrix} \sim \begin{bmatrix} k^{-A} \\ k^{-B} \\ k^{+C} \end{bmatrix} \rightarrow \begin{bmatrix} 0.732 \\ 0.158 \\ 1.549 \end{bmatrix}$	→ For Up - regulated genes
$\begin{bmatrix} P \\ C \\ C_N \end{bmatrix} \sim \begin{bmatrix} k^{-A} \\ k^{-B} \\ k^{+C} \end{bmatrix} \rightarrow \begin{bmatrix} 0.719 \\ 0.563 \\ 1.093 \end{bmatrix}$	→ For Down - regulated genes

TABLE 4 | Pathways enriched by differentially expressed genes (DEGs) among TB, LTBI, and healthy control (HC).

Normal to Latent Infection			
Up-regulated		Down-regulated	
Transcriptional misregulation in cancer Phagosome, AMPK signaling pathway, ECM-receptor interaction, Hematopoietic cell lineage, PPAR signaling pathway, Adipocytokine signaling pathway	AFF1 CD36	Toll-like receptor NF-Kappa B signaling pathway Cytokine–cytokin receptor interaction MAPK signaling pathway Tuberculosis	CCL4, CXCL11, IL1B, IRAK2, IL6, TNF CCL4, IL1B, PTGS2, TNF CCL4, IL1A, IL1B, IL6, PTGS2 IL1A, IL1B, TNF IL1A, IL1B, IRAK2, IL6, TNF
Oxidative phosphorylation, non-alcoholic fatty liver disease (NAFLD), neurotrophin signaling pathway, Alzheimer's disease NOD-like receptor signaling pathway Wnt signaling pathway	NDUFS8 NLRC4 FZD2	TNF signaling	IL1B, IL6, PTGS2, TNF
Normal to Active TB			
Up-regulated		Down-regulated	
Cytokine–cytokine receptor interaction	CCL3, CCL8, CCR2, CXCL10, CXCL11, CXCL13, CXCL9, TNFSF10, TNF	Cell cycle	SKP2, IQGAP3, BCAT1, ID3, IL10, MAPK3K8, PHLDA1, PTGS2, THBS1, TRIM36
Chemokine signaling pathway	CCL3, CCL8, CCR2, CXCL10, CXCL11, CXCL13, CXCL9, JAK2, NCF1, NFKB1, STAT1, TLR8, TNF	Chemokine signaling pathway	CCL20, CCL23, CCL4, CXCL1, CXCL2, CXCL3, CXCL8
Toll-like receptor signaling pathway	CCL3, CXCL10, CXCL11, CXCL9, CD80, BIRC3, ITF1, IRAK2, NFKB1, STAT1, TLR8, TNF	NF-kappa B signaling pathway	CCL4, CXCL8, TNFAIP3, TRAF1, IL1B, PTGS2
NF-kappa B signaling pathway	DDX58, BIRC3, ICAM1, NFKB1, TNF, VCAM1	TNF signaling pathway	CCL20, CXCL1, CXCL2, CXCL3, TNFAIP3, IL1B, IL6, MAP3K8, PTGS2
Transcriptional misregulation in cancer	ETV7, FCGR1A, IGFBP3, NFKB1, PAX5, PML	PI3K-Akt signaling pathway	DDIT4, KIT, CSF3, FLT1, ITGB8, IL6, OSM, THBS1
Pathways in cancer	BIRC3, FZD2, NFKB1, PML, STAT1	Metabolic pathways	AK4, UPB1, BCAT1, DHRS9, KYNU, KMO, NAMP, PLPP3, PTGS2
		MAPK signaling pathway	FLNB, IL1A, IL1B, MAP3K8
Latent Infection to Active TB			
Up-regulated		Down-regulated	
Metabolic pathway	AGPAT3, AK5, BST1, CBR1, COMPT, DPYD, DPAGT1, HK3, HSD17B4, IDH1, LPCAT2, MTHFD1, PGM2, PYGL, QDPR, TST, TBXAS1, TALDO1	Chemokine signaling pathway TNF signaling pathway RAP1 signaling pathway PI3K_AKT signaling Metabolic process	CXCL1, CXCK5, CXCR3, NFKB1 CXCL1, BIRC3, MAPK8, NFKB1 MET, DOCK4, PARD6G, THBS1 MET, COL1A1, ITGB8, NFKB1, THBS1 AMPD3, B3GALT2, CYP1A1, KMO
NOD-like receptor signaling pathway Regulation of actin cytoskeleton Biosynthesis of antibiotics Metabolism of xenobiotics by cytochrome p450 Thyroid hormone synthesis	NAIP, PYCARD, CARD9, PSTPIP1 ARHGFE6, TRIOBP, CYFIP1, DIAPH2 AK5, HK3, IDH1, PGM2, TALDO1 AKR7A2, CBR1, GSTK1, MGST2 ASGR1, CREB5, PRKCB	Apoptotic	CXCR3, ETV6, SKIL, BIRC3, CYP1B1, EPB41L3, HEY2, IL1A, LMNA, MAPK8, NFKB1, PDCD4, SERPINB2, THBS1
Regulation of autophagy	PYCARD, TBC1D5, ATG16L2, LRRK2, ULK2	Focal adhesion	MET, BIRCC3, COL1A1, ITGB8, MAPK8, THBS1
Peroxisome	FAR2, GSTK1, HSD17B4, IDH1	MAPK signaling pathway MicroRNAs in cancer Amoebiasis	DUSP16, DUSP4, IL1A, MAPK8, NFKB1 MET, CYP1B1, MIT155, NFKB1, PDCD4, THBS1 RAB7B, COL1A1, NFKB1, SERDINB2

Our network behavior indicates hierarchical scale free network (Barabási and Oltvai, 2004) (Pastor-Satorras et al., 2001) (Ravasz and Barabási, 2003). The power law fits on the data points of the network's topological parameters were done and confirmed by following the standard statistical fitting method given by

Clauset *et al.* (Clauset et al., 2009). The p values for all data sets were calculated (against 2,500 random samplings) and found to be greater than 0.1, and data fitting goodness was less than 0.1. The values of **P(k)** and **C(k)** were negative, which implies that the network follows a hierarchical pattern, and **C_N(k)** was

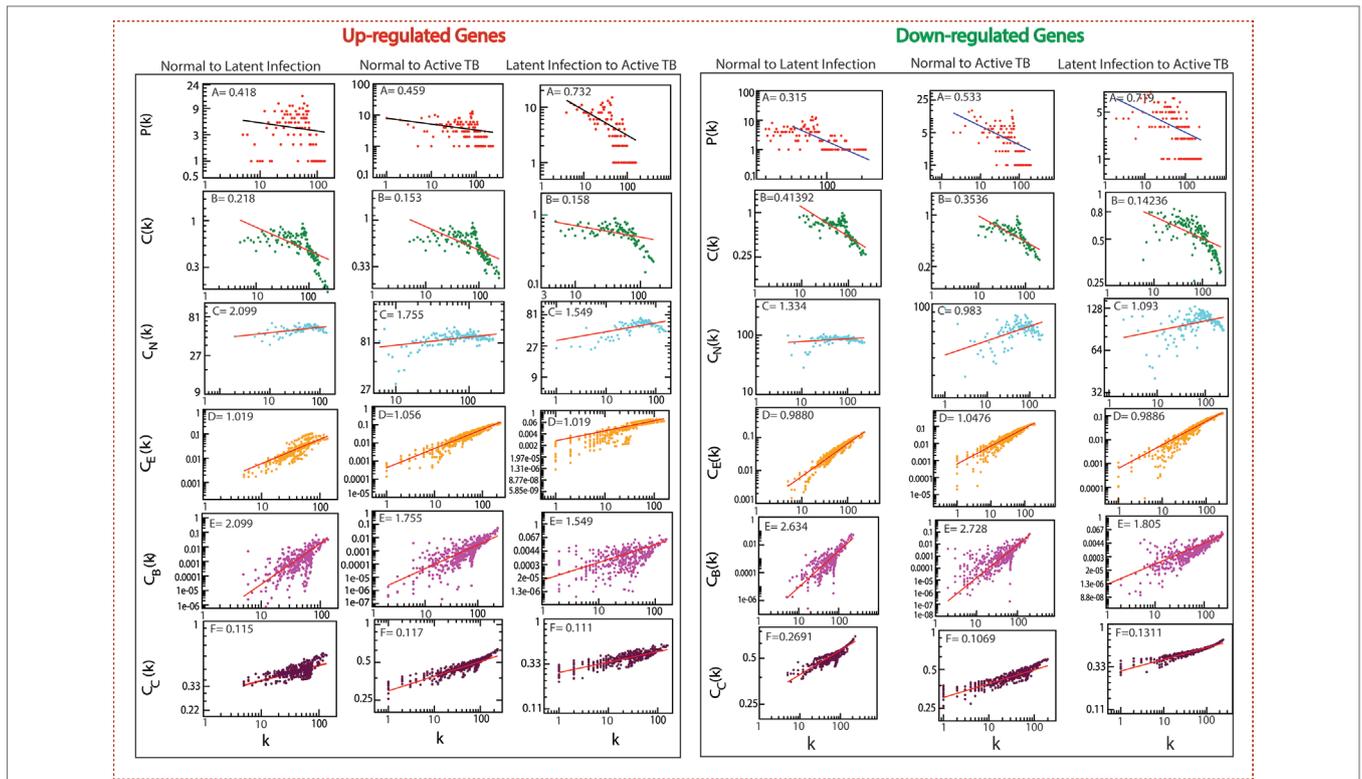


FIGURE 3 | Topological properties of all the six networks. The behaviors of degree distributions ($P(k)$), clustering co-efficient ($C(k)$), neighborhood connectivity ($C_N(k)$), betweenness ($C_B(k)$), closeness ($C_C(k)$), and eigenvector ($C_E(k)$) measurements as a function of degree k . The lines are fitted lines with power laws in the data sets.

positive, which means that the network follows the assortativity that identifies a huge cluster of degree-nodes (rich club), which regulates the network. The centrality parameters: betweenness (C_B), closeness (C_C), and eigenvector (C_E) centralities of the network also showed fractal behavior, and good connectivity of nodes in a network is distinguished by eigenvector or centrality $C_E(k)$. It calculates the effectiveness of the spreading (receiving) power of data of nodes from the network. These properties follow the power law behaviors as follows: -

<p>Normal to latent Inf.</p> $\begin{bmatrix} C_B \\ C_C \\ C_E \end{bmatrix} \sim \begin{bmatrix} k^E \\ k^F \\ k^D \end{bmatrix} \rightarrow \begin{bmatrix} 2.099 \\ 0.115 \\ 1.019 \end{bmatrix}$ $\begin{bmatrix} C_B \\ C_C \\ C_E \end{bmatrix} \sim \begin{bmatrix} k^E \\ k^F \\ k^D \end{bmatrix} \rightarrow \begin{bmatrix} 2.634 \\ 0.269 \\ 0.989 \end{bmatrix}$	<p>Normal to active TB</p> $\begin{bmatrix} C_B \\ C_C \\ C_E \end{bmatrix} \sim \begin{bmatrix} k^E \\ k^F \\ k^D \end{bmatrix} \rightarrow \begin{bmatrix} 1.755 \\ 0.117 \\ 1.056 \end{bmatrix}$ $\begin{bmatrix} C_B \\ C_C \\ C_E \end{bmatrix} \sim \begin{bmatrix} k^E \\ k^F \\ k^D \end{bmatrix} \rightarrow \begin{bmatrix} 2.728 \\ 0.107 \\ 1.048 \end{bmatrix}$
<p>Latent Inf. to active TB</p> $\begin{bmatrix} C_B \\ C_C \\ C_E \end{bmatrix} \sim \begin{bmatrix} k^E \\ k^F \\ k^D \end{bmatrix} \rightarrow \begin{bmatrix} 1.549 \\ 0.111 \\ 1.021 \end{bmatrix}$ $\begin{bmatrix} C_B \\ C_C \\ C_E \end{bmatrix} \sim \begin{bmatrix} k^E \\ k^F \\ k^D \end{bmatrix} \rightarrow \begin{bmatrix} 1.805 \\ 0.131 \\ 0.987 \end{bmatrix}$	<p>→ For Up - regulated genes</p> <p>→ For Down - regulated genes</p>

Identification of Key Regulators and Properties

Since the popularity of leading hubs gets changed according to the gene activities and its regulation, we can't say that all the predictive leading hubs are key regulators for disease progression, but some of these hubs can play a significant role, which we called them as fundamental key regulators (FKR). The structure of modular and its arrangement have been performed through Newman and Girvan's standard community finding algorithm at various levels of the organization. Using this community finding algorithm, we found that our six networks are hierarchically organized at various levels (**Supplementary Data 2**). The Hamiltonian energy (HE) and corresponding modularity (Q_N) are reduced as one goes from top to down organization of network (**Figure 4**).

We have calculated the probability to know the regulating ability of our 31 key regulators:

$$P_x(y^{[l]}) = \frac{y^{[l]}}{E^{[l]}}$$

where,

x = no. of edges ($y^{[l]}$)
 $(E^{[l]})$ = Total no. of edges of the network (modules and sub-modules).

The measured probability $P_x(y^l)$ of all the KR showed an increase in P_x as the level the increases from top to bottom

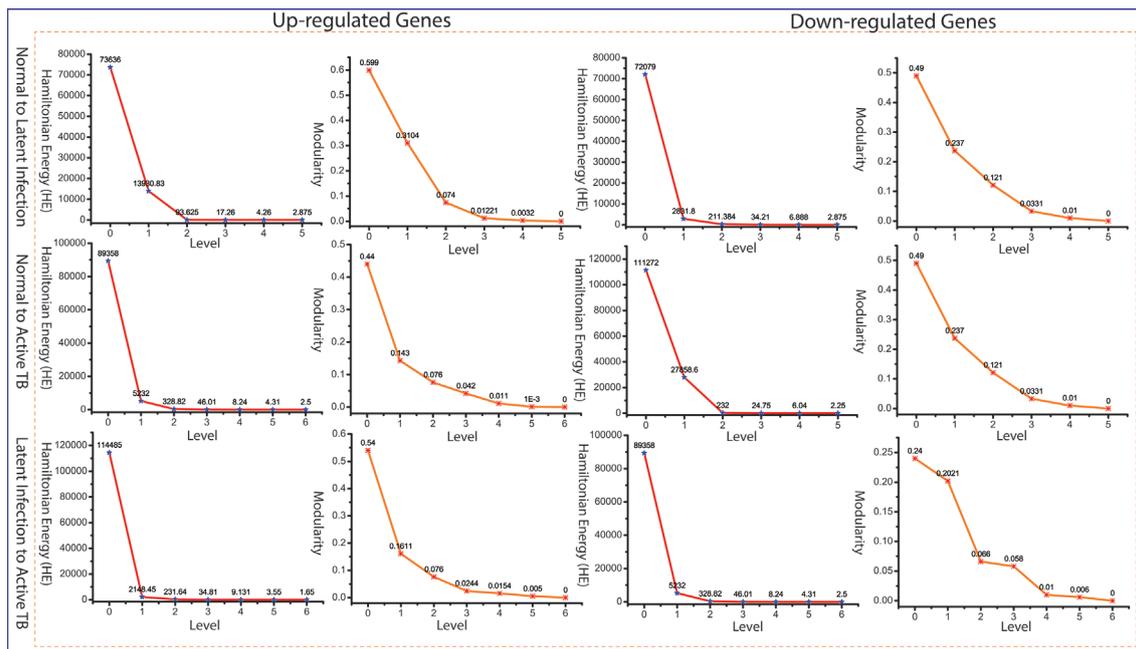


FIGURE 4 | Energy and modularity distribution in all networks quantified by Hamiltonian (HE) and modularity calculation as a function of network levels.

direction. At deeper level of the organization, regulation of FKR increases, and their activities become more prominent. Thus, these FKR becomes backbone of the network organization, stabilization, and active workers at grassroots level. The FKR are shown in **Figures 5 and 6**.

As the results, we identified a total of 31 FKR genes from various stages of TB. In normal to latent infections, we identified eight genes, of which three genes (FZD2, NDUFS8, NLRC4) were up-regulated, and another five genes (CCL4, IL1B, IL1A, TNF, AREG) were down-regulated. While in normal to active TB, we found 15 genes out of which eight genes (EIF2AK2, SAMD9L, IFI44L, DDX58, IFI44, HERC5, NFE2L3, LRRK2) were up-regulated, and seven genes (CSF3, MAP3K8, IRAK2, TNFAIP3, TRAF1, PLAUR, and CD44) were down-regulated, and similarly, for latent to active TB diseases, three genes (TST, MTHFD1, CARS2) were up-regulated, and five genes (ETV6, NFKB1, MET, DOCK4, PARD6G) were down-regulated. The functional interpretations of these 31 key regulators were annotated by GO-TermFinder (LAGO) (Boyle et al., 2004) using biological process terms at P-value 0.01; the results are interpreted in **Figure 7**. The gene specific pathways were identified by KEGG pathway database (Kanehisa et al., 2016), and the results are shown in **Figure 8**.

Local Perturbations of Network

The gene knockout experiment of all the hubs/motif from the parent networks may able to highlight the local perturbations driven by these individual hub or motif, and their effect on global network properties. It has been revealed that the network is tolerant to hub's deletion which implies that the important network elements are still remains after elimination of hubs

at level 0 (parent network). One of main causes is that the P-P interactions networks are too dense to be broken into fragments by only removing hubs. However, the elimination of these hubs/motif from the complete network does cause significant variations in the network properties, where $P(k)$ and $C(k)$ change significantly in complete network level, whereas $C_N(k)$ changes slightly. Likewise, the variations in the exponents of centrality measurements have also showed significant changes, as shown in **Figure 9**. Since, it is clear from the differences in the exponents of topological parameters that network perturbation increases when it goes to deeper level (top to down direction). In our case, most of the perturbation increases after the 3rd level; at this level elimination of key regulator from network almost breaks down the submodules existing in the deeper levels; such type of behavior shows that local perturbation of network is highest at deeper levels.

The Local-Community-Paradigm: Evidence of Self-Organization

The LCP architecture supports the quick transfer of data across several network modules and through local processing too. We have analyzed all the six networks to check self-organization behavior at different levels using LCP method. For different level, the calculated LCP-corr of all the modules or sub-modules are shown in **Figures 10 and 11**. The average values of LCP-corr (we ignored modules which having zero LCP-corr) were greater than 0.853 at each level. This shows that the network maintains compactness and self-organization and has efficient data processing. It characterizes robust LCP networks that are dynamic in nature and heterogeneous, which help in network re-organization and evolution.

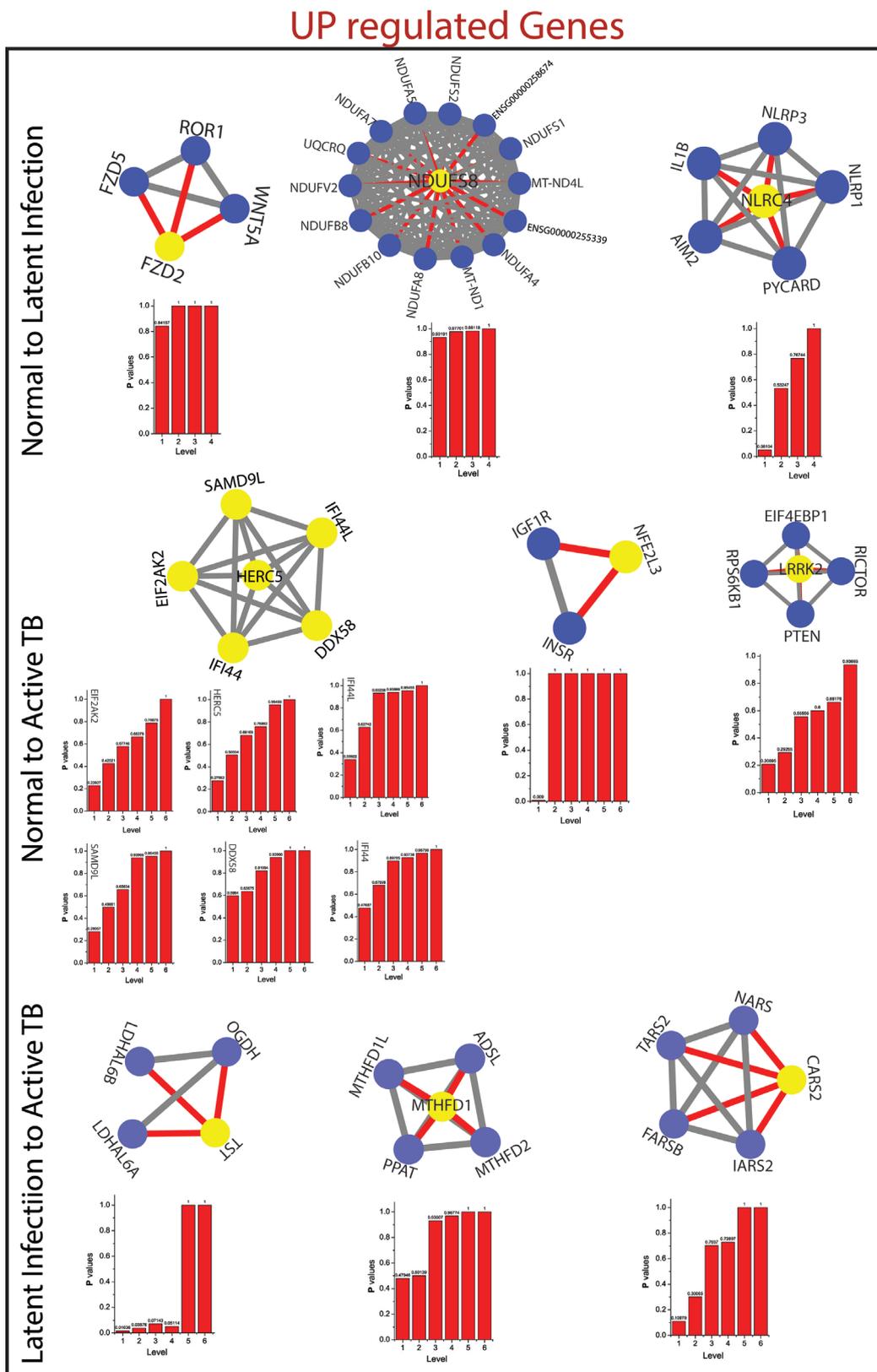


FIGURE 5 | This figure shown the fundamental key regulator from various stage of TB obtained from main networks to motif/hub level through various modules/sub-modules at various level of organization. The probability distribution of the 14 up-regulated key genes as a function of level of organization.

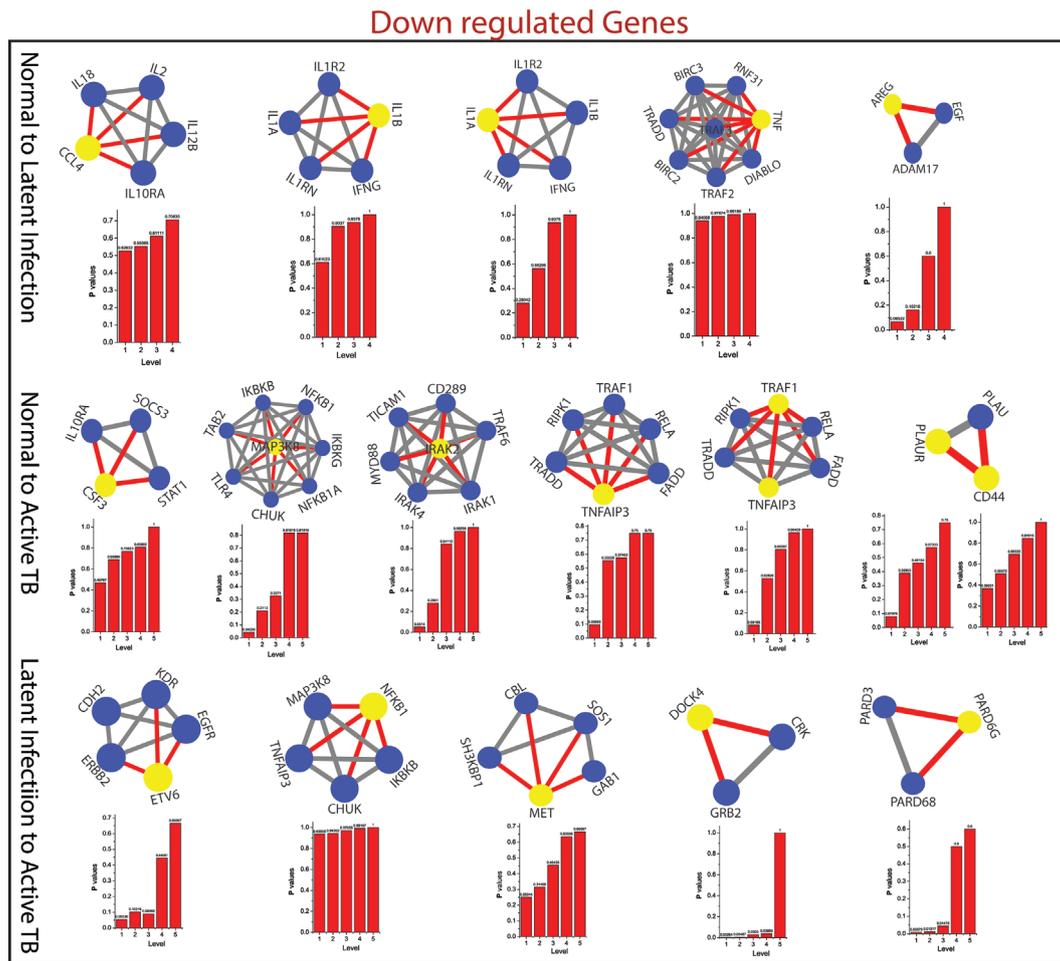


FIGURE 6 | This figure shown the fundamental key regulator from various stage of TB obtained from main networks to motif/hub through various modules/sub-modules at various level of organization. The probability distribution of the 15 down-regulated key genes as a function of level of organization.

DISCUSSION

TB disease remains to be one of the most significant infectious diseases and a leading cause of death worldwide. Currently, the main challenge is to develop a delicate and effective method to identify the latent TB infection (LTBI) because 90% of cases of latent infection with MTB do not show any symptoms/signs but have a 10% lifetime possibility of transforming into active TB. As we know, many events have happened in response to TB infection cycle. According to the cycle of infection by Young et al., 2008, the exposure of microorganism up to its development to latent stage, there are stringent amount of processes that have appeared in the host cell. The long fight of MTB to establish its generation by stabilizing its environment for metabolism has evolved with an ability to overcome immunity. In this scenario of transferring from initial infection stage to latent stage has a profound increase in the overall immune response outside the host cell and an increased intensity of basic cellular machine establishment process for the pathogen cell. In case, the elimination eliminates the eliminator even after the induction of T-cell but able to suppress the effect;

pathogen cells (Pcells) have two choices, either develop army or attack the adaption of immunity. In order to decide to enter the resting phase (i.e., latent), MTB's survival instincts allow it to develop its metabolic environment by downregulating antigen presentation and the release of cytokines too (like IFN γ). On the other hand, if adaptive immunity has taken its chance, then MTB takes the defensive mode by subverting various normal cell cycle functioning inside the macrophages thus making a halt on the maturation of phagosomes.

To stop the disease epidemic, early diagnostic method or techniques are required. In this way, the gene expression profiling has uncovered the differences in the transcriptome among normal condition, latent infection, and active TB. These results not only revealed significant genetic biomarkers indicative of LTBI, TB-disease conditions but also recognized transcriptionally regulated genes that vary in biological functions. In the current study, we have identified differentially expressed genes (DEGs) from various stages of disease and found that most of the DEGs relative to normal vs. TB disease condition (including latent infection) are enriched in important



FIGURE 7 | This heatmap shows the 31 hub genes with their involvement in various biological processes.

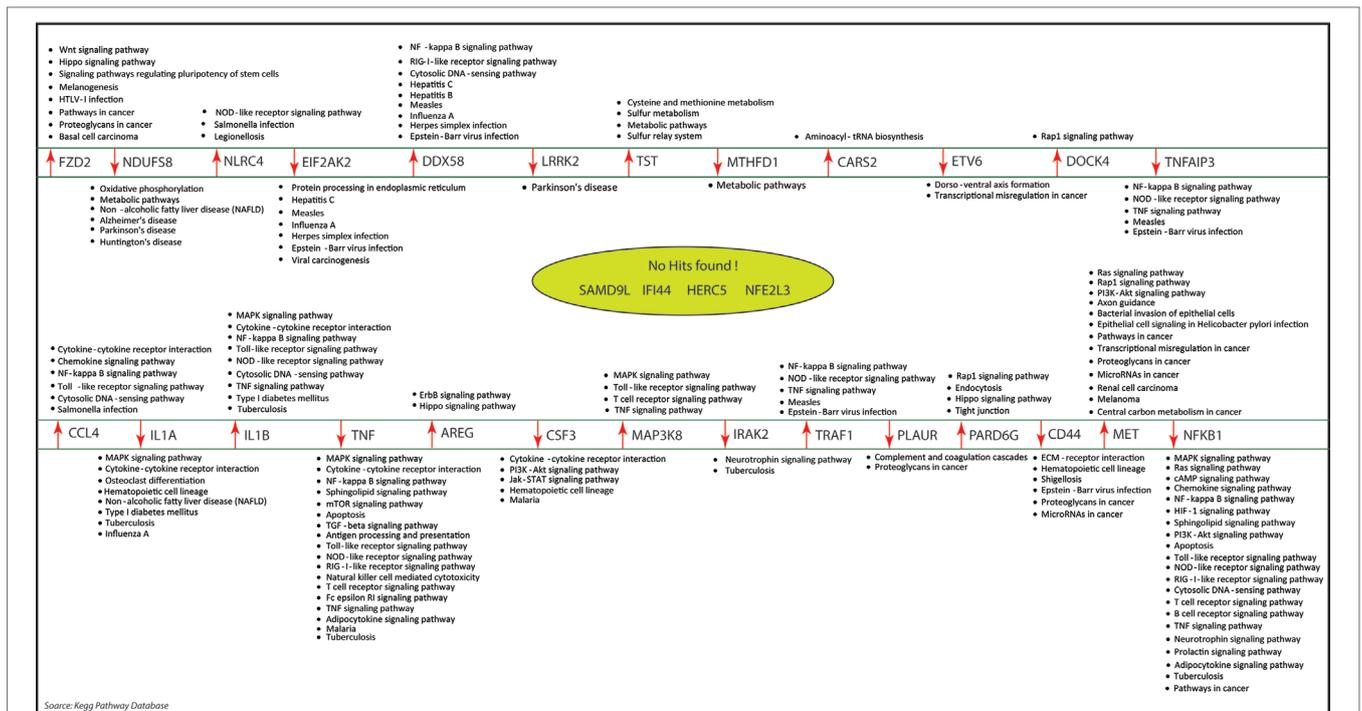


FIGURE 8 | The molecular pathways associated with our key regulators were identified by KEGG pathway database. Out of the 31 key regulators, we did not find any pathways information of five genes namely: SAMD9L, IFI44L, IFI44, HERC5, and NF2L3.

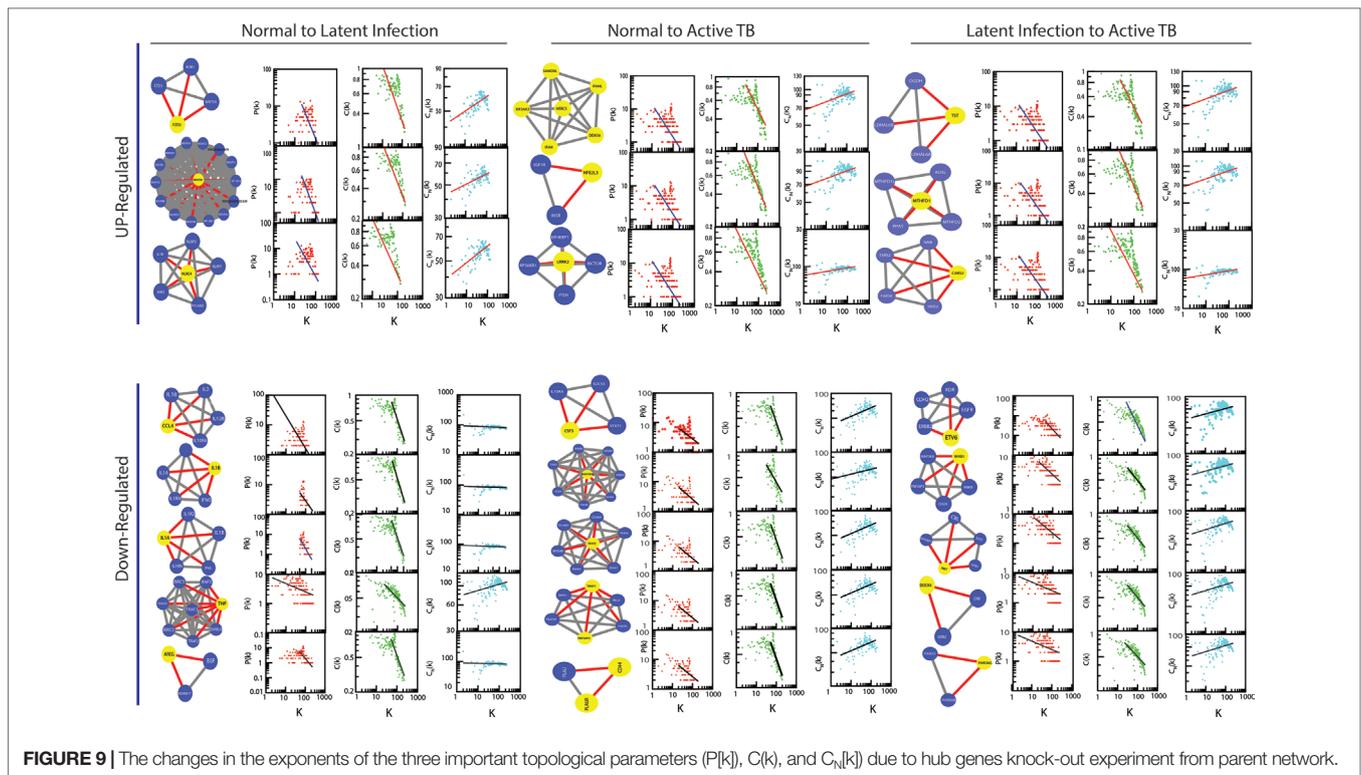
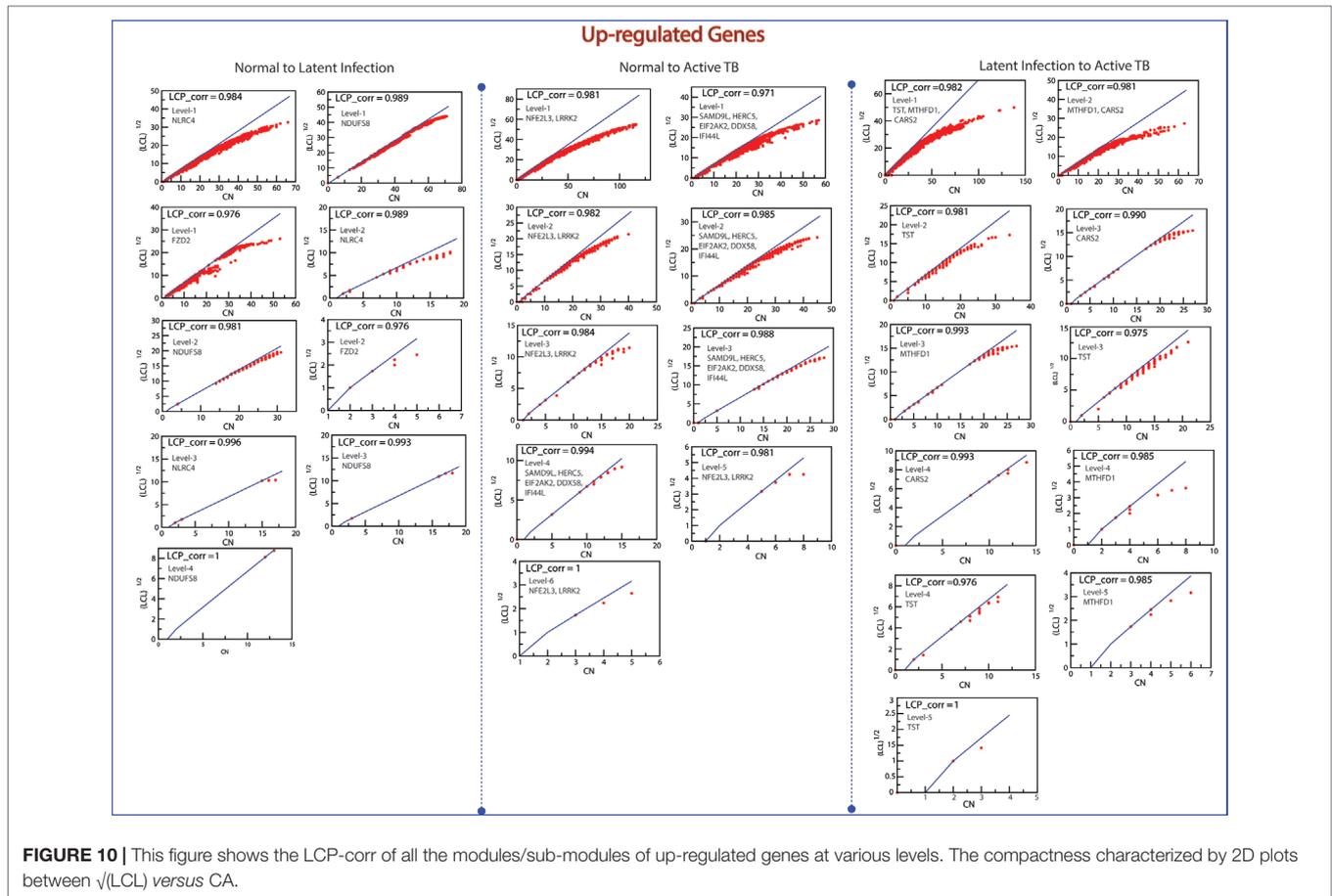


FIGURE 9 | The changes in the exponents of the three important topological parameters ($P[k]$), $C(k)$, and $C_N[k]$) due to hub genes knock-out experiment from parent network.

pathway like-toll like receptor, NOD-like receptor signaling pathway, MAPK signaling pathway, TNF signaling, chemokine signaling pathway, PI3K-Akt signaling pathway, apoptosis etc. In fact, pathogen is identified by the receptors on the surface of immune cells, and toll-like receptors (TLRs) are one of them. Different TLRs including TLR2, 4, 9, and 8 play an important roles in TB infection (Faridgozar and Nikouejad, 2017). These receptors (toll-like, NOD-like etc.) are expressed irrespective of whether they participate in immune signaling or immunity. The interaction of MTB with these receptors (like TLRs) initiates an intercellular signaling cascade that culminates in a pro-inflammatory response (cytokines and chemokines that serve as a signal for infection). As we know that MAPK signaling pathway is important for transcription and non-transcription responses of the immune system, most of the pathogens during infection hijack the immune system by targeting the MAPK signaling pathway (Soares-Silva et al., 2016). The TNF signaling pathway is involved in the regulation of immune cells, and it can regulate immune responses (Kim, 2018). However, pathogen can target the TNF signaling pathway for its survival in the host cell. The PI3K-Akt signaling pathway plays an important role in apoptosis, autophagy, metabolism, cell growth, and differentiation. The expression of FoxP3 by inhibiting the activation of transcription factor Forkhead-O3a (Foxo1-3a) is negatively regulated by this pathway (Zhang et al., 2017). The FoxP3+Treg cell activation which will assist to set up a new target for the involvement of TB immunotherapy molecules as part of the immune-escape mechanism to provide a theoretical basis is inhibited by *M. tuberculosis* (Scott-Browne et al., 2007). On comparing between latent infection and TB disease, few additional pathways

are involved, like regulation of actin cytoskeleton during MTB infection; to maintain the stability of the cytoskeleton, macrophages cells themselves are also trying to regulate cytoskeletal-associated proteins (Wang et al., 2015b). Thyroid hormones (hormones, T4 and T3) are produced by thyroid gland, which is essential for the regulation of metabolic processes throughout the body. The regulation of autophagy is important for host in response to invading mycobacteria; the host defense mechanism identifies pathogen motifs through innate receptors but also releases appropriate cytokines. The autophagic pathways are regulated by these innate signals by regulation of genes of this pathway during infection (Jo, 2013). Recently, a study reports that the enzyme Msm_ACTase (from bacteria) helps in scavenging increased amount of H_2O_2 due to upregulation of genes involved in peroxisome pathway which gives an insight into a new idea as to how the MTB bacteria surpasses the host defense in MTB infection (Ganguli, 2015). Moreover, macrophages are the main effector cells responsible for killing pathogen (MTB) *via* different mechanisms, including apoptosis. But important genes in this pathway were down-regulated by MTB that slow down or stop the apoptotic process for its survival in host system (Behler et al., 2015). Most of the microRNAs that were found deregulated in cancer were affected by MTB in a similar way as its infestations, like deletion, mutation, and epigenetic silencing (Garzon et al., 2009). The apoptotic pathways are important to kill the infected macrophage. During MTB infection, the pathogen try to control the timing and mode of host cell death for its persistence and replication (Schaaf et al., 2017)

In addition, we have built the networks which emphasize on genes that were regulated by network. The constructed network



of classified genes from various stages of TB showed hierarchical nature, which indicated that the networks have system level organization including modules/sub-modules which are interconnected. Since the network's nature is hierarchical, its synchronization confirms several important functional regulations of the network, but individual gene activities are not so important. In our networks (including up- and down-regulated genes), a total of 31 key regulators were identified by affecting motifs and module regulation, showing their biological importance and serve as the foundation of network activities and their regulations, and could be the most probable target gene for disease control. We performed extensive analysis of all the key regulators (31 genes) to find evidence of their association with TB or which are directly or indirectly involved in host immune response and confirmed by manually searching for evidence in the literatures as presented:

- **Down-regulated genes:** We found that *TNF* and *IL1* (*A* and *B*) are key mediators present in severe inflammatory diseases. However, both *IL-1* and *TNF* receptor pathways are important for the control of *MTB* infection, and it is critical to assess the respective role of *IL1A*, *IL1B*, and *TNF* (Bourigault et al., 2013) (Cavalcanti et al., 2012). The expression of *CCL4* is high in the late phase of the active TB disease, but there are low levels of expression during early

infection (Rangel-Santiago et al., 2016). It has been currently reported that *AREG* plays an important role in orchestrating both host resistance and tolerance mechanisms. Although *AREG* is known as epithelial cell-derived factor, the recent studies showed that *AREG* can be expressed by multiple populations of activated immune cells in inflammatory conditions (Zaiss et al., 2015). The *NFKB1* activates the immune cells by up-regulating the expression of various cytokines in the host immune system (Ghosh et al., 1998). The gene *CD44* plays an important role in innate and adaptive immune responses, in the acute inflammatory response to both infectious and sterile stimuli. Also during infection, *CD44* may influence host defense by affecting phagocytosis (van der Windt et al., 2010). In our present study, we found that *NFKB1* and *CD44* are down-regulated which means that *MTB* is burglarized and seizing the host immune system by slowing down the expression of *NFKB1*. It is well known that *MET* (hepatocyte growth factor receptor) regulates several functions of immune cells, including cytokine production, differentiation and maturation, cellular migration and adhesion, and T-cell effector function, and HGF exerts anti-inflammatory activities through *MET* signaling (Sagi and Hieronymus, 2018). A group of researchers has elucidated that *PLAUR* domain containing *Lypd8* that inhibits bacterial invasion of colonic epithelia (Okumura et al., 2016). The gene

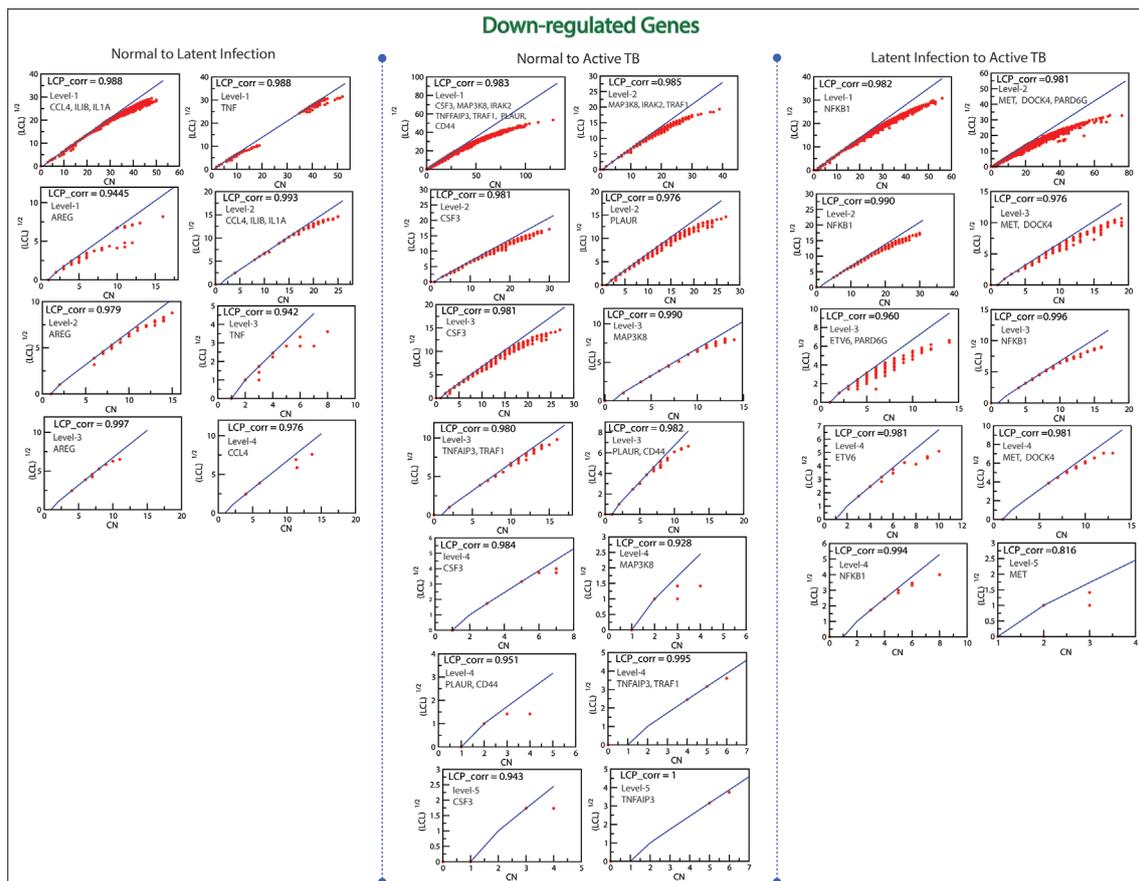


FIGURE 11 | This figure shows the LCP-corr of all the modules/sub-modules of down-regulated genes at various levels. The compactness characterized by 2D plots between $\sqrt{\text{LCL}}$ versus CA.

TRAF1 has an active or passive interaction with many tumor necrosis factor receptor (TNFR), and it has been also shown that *TRAF1* plays as an antiapoptotic role in lymphoma cells by activation of NFKB (Wan et al., 2016). Further, the gene *TNFAIP3* (*A20*) is a cytoplasmic zinc-finger protein which works as a negative-feedback regulator of NFKB activation (Vereecke et al., 2011). It has been shown that *IRAK2* is an important factor and potential novel biomarkers for human antiviral innate immunity (Wang et al., 2015a). Also, *MAP3K8* is a serine-threonine kinase and has a critical function in integrating host immune responses to complex pathogens (Mielke et al., 2009). Similarly, S. Srijata, et al. has suggested that AgNP attenuate the NF- κ B pathway as indicated by the downregulation of NF- κ B target genes *CSF3* and subsequently inhibit MTB-induced proinflammatory responses (Sarkar et al., 2015). The downregulation of these genes has allowed the process of host defense to cease and give time to bacteria for progression.

- **Up-regulated genes:** We found that *NLRC4* gene is associated with inflammasome signaling. Its upregulation means inflammasome activation (play an important role in host defense against MTB) not only leads to cytokine secretion but may also cause pyroptosis (Ting et al., 2008) (Bergsbaken

et al., 2009). The *NDUFS8* genes also play an important role in host immunity by increasing the expression level (Lohman et al., 2017). The gene *FZD2* (frizzled class receptor 2) is Wnt pathway component (Wnt signaling), and Wnt signaling pathway plays a key role at different stages of TB development (Villaseñor et al., 2017). It has been reported that *LRRK2* is involved in the IFN- γ response and host response to pathogens (Gardet et al., 2010), and *LRRK2* also inhibits the immune response transcription factor NFAT1 (Liu et al., 2011). The genes *DDX58* interact with IRGM and promote its K63-linked polyubiquitination, indicating that IRGM is positioned at a nexus of various innate immunity (Chauhan et al., 2016) while *SAMD9L* and *EIF2AK2* play key roles in the innate immune responses to multiple stimuli. The gene *EIF2AK2* is also involved in the regulation of signal transduction, apoptosis, cell proliferation, and differentiation (Lemos de Matos et al., 2013) (Onomoto et al., 2012) (Reineke and Lloyd, 2015). It is well known that type-I IFN induces potent cellular defense against viral infection mediated by up-regulation of ISGs like *IFI44*, *IFI44L*, and *HERC5* (Jeon et al., 2010; Power et al., 2015). In our study, the expression of *IFI44* and *HERC5* were found to be up-regulated that means host system tried to resist bacterial infection.

We didn't find any literature evidence that supports the role of four genes (*NFE2L3*, *TST*, *MTHFD1*, and *CARS2*) and three genes (*ETV6*, *DOCK4*, and *PARD6G*) which are down-regulated and up-regulated respectively in host immune response against the pathogen, but most of these genes are involved in various types of cancers like *ETV6* is involved in prostate cancer as well as colorectal cancer susceptibility (Fararjeh and Liu, 2019; Wang et al., 2016); mutation in *DOCK4* can cause ovarian, prostate, glioma, and colorectal cancers (Sundaravel et al., 2015), and *PARD6G* suppresses cell proliferation and is targeted by loss-of-function mutations in several cancers (Marques et al., 2016). Besides of these *NFE2L3* (transcription factor) which involved in several cellular processes like inflammation, stress response, differentiation, and carcinogenesis (Kannan et al., 2015), *MTHFD1* are key players in folate metabolism, which is essential for *de novo* purine synthesis, and several defects in this pathway have been associated with immunodeficiency (Keller et al., 2013), and *CARS2* is associated with a severe progressive myoclonic epilepsy most resembling MERRF syndrome (Hallmann et al., 2014). The *TST* gene is a mitochondrial matrix enzyme (Pallini et al., 1991). It may play roles in cyanide detoxification, the formation of iron-sulfur proteins, and the modification of sulfur-containing enzymes.

CONCLUSION

In this study, we have compared gene-expression profiling of active TB patient and latent TB patients, as well as healthy controls. We have combined different data sets from multiple microarray experiment, and this combined effect size method for the gene expression analysis gave us more biologically consistent and conservative results. We have identified differentially expressed genes and their associated particular biological processes and pathways among healthy controls, active TB and LTBI. The gene interaction networks employed to obtain functional modules and also uncovered novel key regulators for active TB and LTBI which might play an important role in regulating the expression of multiple TB-related host factors. In conclusion, our study provides new insights into host genes and pathways important for TB infection. Experimental validation of these hypotheses would confirm the credibility of the key regulators and make easier the development of a cost-effective and sensitive molecular diagnostic platform for TB disease.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <https://www.ncbi.nlm.nih.gov/gds>

REFERENCES

- Ahmad, S. (2011). Pathogenesis, immunology, and diagnosis of latent *Mycobacterium tuberculosis* Infection. *Clin. Dev. Immunol.* 2011, 1–17. doi: 10.1155/2011/814943
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

AUTHOR CONTRIBUTIONS

RI conceived the study design instructed on data analysis. AA curated data and performed statistical and network analyses. NI, MM and AF implemented key statistical tools. AA, NI, MA, ST, AF, and NT drafted the manuscript. All the authors read, edited and approved of the manuscript.

FUNDING

This work was supported by the Department of Health Research (DHR), Ministry of Health and Family Welfare, Government of India. Under the scheme of “Young Scientist” fellowship (R.12014/06/2019-HR).

ACKNOWLEDGMENTS

The authors are grateful to the Centre for Interdisciplinary Research in Basic Sciences (CIRBSc), Jamia Millia Islamia-110025 for providing the research infrastructure. AA acknowledges the Department of Health Research (DHR) New Delhi, India for the award of “Young Scientist” fellowship (R.12014/06/2019-HR). NT is thankful to Council of Scientific and Industrial Research (CSIR), New Delhi, India, NI, MA, ST and AF are thankful to Indian Council of Medical Research (ICMR), New Delhi, India. RI duly acknowledges the Indian Council of Medical Research (ICMR), Council of Scientific and Industrial Research (CSIR) and Department of Health and Research (DHR).

Note: “This manuscript (Old version) has been released as a Pre-Print at [<https://www.biorxiv.org/content/10.1101/414110v2>]”.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.00932/full#supplementary-material>

SUPPLEMENTARY DATA 1 | Functional Classification of Differentially expressed according to (A) Molecular functions, (B) Biological processes and (C) Protein classes.

SUPPLEMENTARY DATA 2 | The modular path of important hub genes from complete network to motif level with the structures of modules/sub-modules at various levels.

SUPPLEMENTARY DATA 3 | (A) Community Finding (Network cluster Analysis) script in R. (B) LCP-correlation and LCP-decomposition-plot (LCP-DP approach to estimate the network compactness) method and MATLAB commands.

- Barabási, A.-L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113. doi: 10.1038/nrg1272
- Behler, F., Maus, R., Bohling, J., Knippenberg, S., Kirchhof, G., Nagata, M., et al. (2015). Macrophage-inducible C-type lectin mincle-expressing dendritic cells contribute to control of splenic *Mycobacterium bovis* BCG infection in mice. *Infect. Immun.* 83, 184–196. doi: 10.1128/IAI.02500-14
- Bergsbaken, T., Fink, S. L., and Cookson, B. T. (2009). Pyroptosis: host cell death and inflammation. *Nat. Rev. Microbiol.* 7, 99–109. doi: 10.1038/nrmicro2070

- Blankley, S., Graham, C. M., Turner, J., Berry, M. P. R., Bloom, C. I., Xu, Z., et al. (2016). The transcriptional signature of active tuberculosis reflects symptom status in extra-pulmonary and pulmonary tuberculosis. *PLoS One* 11, e0162220. doi: 10.1371/journal.pone.0162220
- Bonacich, P. (1987). Power and centrality: a family of measures. *Am. J. Sociol.* 92, 1170–1182. doi: 10.1086/228631
- Bourigault, M.-L., Segueni, N., Rose, S., Court, N., Vacher, R., Vasseur, V., et al. (2013). Relative contribution of IL-1 α , IL-1 β and TNF to the host response to *Mycobacterium tuberculosis* and attenuated *M. bovis* BCG: IL-1 α , IL-1 β versus TNF in host response to *M. tuberculosis*. *Immun. Inflamm. Dis.* 1, 47–62. doi: 10.1002/iid3.9
- Bouttier, M., Laperriere, D., Memari, B., Mangiapane, J., Fiore, A., Mitchell, E., et al. (2016). Alu repeats as transcriptional regulatory platforms in macrophage responses to *M. tuberculosis* infection. *Nucleic Acids Res.* 44, 10571–10587. doi: 10.1093/nar/gkw782
- Boyle, E. I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J. M., et al. (2004). GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinforma. Oxf. Engl.* 20, 3710–3715. doi: 10.1093/bioinformatics/bth456
- Brandes, U. (2001). A faster algorithm for betweenness centrality*. *J. Math. Sociol.* 25, 163–177. doi: 10.1080/0022250X.2001.9990249
- Cai, Y., Yang, Q., Tang, Y., Zhang, M., Liu, H., Zhang, G., et al. (2014). Increased complement C1q level marks active disease in human tuberculosis. *PLoS One* 9, e92340. doi: 10.1371/journal.pone.0092340
- Cannistraci, C. V., Alanis-Lobato, G., and Ravasi, T. (2013). From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks. *Sci. Rep.* 3, 1613. doi: 10.1038/srep01613
- Cavalcanti, Y. V. N., Brelaz, M. C. A., Lemoine Neves, J. K., de, A., Ferraz, J. C., and Pereira, V. R. A. (2012). Role of TNF- α , IFN- γ , and IL-10 in the development of pulmonary tuberculosis. *Pulm. Med.* 2012, 1–10. doi: 10.1155/2012/745483
- Chauhan, S., Mandell, M. A., and Deretic, V. (2016). Mechanism of action of the tuberculosis and Crohn disease risk factor IRGM in autophagy. *Autophagy* 12, 429–431. doi: 10.1080/15548627.2015.1084457
- Cheng, C. Y., Gutierrez, N. M., Marzuki, M. B., Lu, X., Foreman, T. W., Paleja, B., et al. (2017). Host sirtuin 1 regulates mycobacterial immunopathogenesis and represents a therapeutic target against tuberculosis. *Sci. Immunol.* 2. doi: 10.1126/sciimmunol.aaj1789
- Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Rev.* 51, 661–703. doi: 10.1137/070710111
- Fararjeh, A. S., and Liu, Y.-N. (2019). ZBTB46, SPDEF, and ETV6: Novel potential biomarkers and therapeutic targets in castration-resistant prostate cancer. *Int. J. Mol. Sci.* 20, 2802. doi: 10.3390/ijms20112802
- Faridoghar, M., and Nikouinejad, H. (2017). New findings of toll-like receptors involved in *Mycobacterium tuberculosis* infection. *Pathog. Glob. Health* 111, 256–264. doi: 10.1080/20477724.2017.1351080
- Gabor, C., and Nepusz, T. (2006). The igraph software package for complex network research. *Inter J Comp Syst* 1695, 1–9.
- Ganguli, G. (2015). Mycobacterium tuberculosis acetyltransferase reduces the oxidative stress response through expression of peroxisomal membrane transporter protein. *Open Forum Infect. Dis.* 2. doi: 10.1093/ofid/ofv131.70
- Gardet, A., Benita, Y., Li, C., Sands, B. E., Ballester, I., Stevens, C., et al. (2010). LRRK2 is involved in the IFN-response and host response to pathogens. *J. Immunol.* 185, 5577–5585. doi: 10.4049/jimmunol.1000548
- Garzon, R., Calin, G. A., and Croce, C. M. (2009). MicroRNAs in cancer. *Annu. Rev. Med.* 60, 167–179. doi: 10.1146/annurev.med.59.053006.104707
- Ghosh, S., May, M. J., and Kopp, E. B. (1998). NF- κ B and Rel proteins: evolutionarily conserved mediators of immune responses. *Annu. Rev. Immunol.* 16, 225–260. doi: 10.1146/annurev.immunol.16.1.225
- Guerra-Laso, J. M., Raposo-García, S., García-García, S., Diez-Tascón, C., and Rivero-Lezcano, O. M. (2015). Microarray analysis of *Mycobacterium tuberculosis*-infected monocytes reveals IL26 as a new candidate gene for tuberculosis susceptibility. *Immunology* 144, 291–301. doi: 10.1111/imm.12371
- Hallmann, K., Zsurka, G., Moskau-Hartmann, S., Kirschner, J., Korinthenberg, R., Ruppert, A.-K., et al. (2014). A homozygous splice-site mutation in CARS2 is associated with progressive myoclonic epilepsy. *Neurology* 83, 2183–2187. doi: 10.1212/WNL.0000000000001055
- Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211
- Jeon, Y. J., Yoo, H. M., and Chung, C. H. (2010). ISG15 and immune diseases. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* 1802, 485–496. doi: 10.1016/j.bbadis.2010.02.006
- Jo, E.-K. (2013). Autophagy as an innate defense against mycobacteria. *Pathog. Dis.* 67, 108–118. doi: 10.1111/2049-632X.12023
- Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44, D457–D462. doi: 10.1093/nar/gkv1070
- Kannan, M. B., Dodard-Friedman, I., and Blank, V. (2015). Stringent control of NFE2L3 (nuclear factor, erythroid 2-like 3; NRF3) protein degradation by FBW7 (F-box/WD repeat-containing protein 7) and glycogen synthase kinase 3 (GSK3). *J. Biol. Chem.* 290, 26292–26302. doi: 10.1074/jbc.M115.666446
- Keller, M. D., Ganesh, J., Heltzer, M., Paessler, M., Bergqvist, A. G. C., Baluarte, H. J., et al. (2013). Severe combined immunodeficiency resulting from mutations in MTHFD1. *Pediatrics* 131, e629–e634. doi: 10.1542/peds.2012-0899
- Kim, J. (2018). Regulation of immune cell functions by metabolic reprogramming. *J. Immunol. Res.* 2018, 1–12. doi: 10.1155/2018/8605471
- Kondratieva, T., Azhikina, T., Nikonenko, B., Kaprelyants, A., and Apt, A. (2014). Latent tuberculosis infection: what we know about its genetic control? *Tuberc. Edinb. Scotl.* 94, 462–468. doi: 10.1016/j.tube.2014.06.009
- Lavalett, L., Rodriguez, H., Ortega, H., Sadee, W., Schlesinger, L. S., and Barrera, L. F. (2017). Alveolar macrophages from tuberculosis patients display an altered inflammatory gene expression profile. *Tuberc. Edinb. Scotl.* 107, 156–167. doi: 10.1016/j.tube.2017.08.012
- Lee, I., Blom, U. M., Wang, P. I., Shim, J. E., and Marcotte, E. M. (2011). Prioritizing candidate disease genes by network-based boosting of genome-wide association data. *Genome Res.* 21, 1109–1121. doi: 10.1101/gr.118992.110
- Lemos de Matos, A., Liu, J., McFadden, G., and Esteves, P. J. (2013). Evolution and divergence of the mammalian SAMD9/SAMD9L gene family. *BMC Evol. Biol.* 13, 121. doi: 10.1186/1471-2148-13-121
- Liu, K., Zhang, Y., Hu, S., Yu, Y., Yang, Q., Jin, D., et al. (2012). Increased levels of BAFF and APRIL related to human active pulmonary tuberculosis. *PLoS One* 7, e38429. doi: 10.1371/journal.pone.0038429
- Liu, Z., Lee, J., Krummey, S., Lu, W., Cai, H., and Lenardo, M. J. (2011). The kinase LRRK2 is a regulator of the transcription factor NFAT that modulates the severity of inflammatory bowel disease. *Nat. Immunol.* 12, 1063–1070. doi: 10.1038/ni.2113
- Lohman, B. K., Steinel, N. C., Weber, J. N., and Bolnick, D. I. (2017). Gene expression contributes to the recent evolution of host resistance in a model host parasite system. *Front. Immunol.* 8, 107. doi: 10.3389/fimmu.2017.01071
- Marques, E., Englund, J. I., Tervonen, T. A., Virkunen, E., Laakso, M., Myllynen, M., et al. (2016). Par6G suppresses cell proliferation and is targeted by loss-of-function mutations in multiple cancers. *Oncogene* 35, 1386–1398. doi: 10.1038/onc.2015.196
- Mason, O., and Verwoerd, M. (2007). Graph theory and networks in biology. *IET Syst. Biol.* 1, 89–119. doi: 10.1049/iet-syb:20060038
- Mi, H., Muruganujan, A., and Thomas, P. D. (2013). PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.* 41, D377–D386. doi: 10.1093/nar/gks1118
- Mielke, L. A., Elkins, K. L., Wei, L., Starr, R., Tschlis, P. N., O'Shea, J. J., et al. (2009). Tumor progression locus 2 (Map3k8) is critical for host defense against *Listeria monocytogenes* and IL-1 production. *J. Immunol.* 183, 7984–7993. doi: 10.4049/jimmunol.0901336
- Newman, M. E. J. (2006a). Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E* 74, 036104. doi: 10.1103/PhysRevE.74.036104
- Newman, M. E. J. (2006b). Modularity and community structure in networks. *Proc. Natl. Acad. Sci.* 103, 8577–8582. doi: 10.1073/pnas.0601602103
- Newman, M. E. J., and Girvan, M. (2004). Finding and evaluating community structure in networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 69, 026113. doi: 10.1103/PhysRevE.69.026113
- Okumura, R., Kurakawa, T., Nakano, T., Kayama, H., Kinoshita, M., Motooka, D., et al. (2016). Lypd8 promotes the segregation of flagellated microbiota and colonic epithelia. *Nature* 532, 117–121. doi: 10.1038/nature17406
- Onomoto, K., Jogi, M., Yoo, J.-S., Narita, R., Morimoto, S., Takemura, A., et al. (2012). Critical role of an antiviral stress granule containing RIG-I and PKR in viral detection and innate immunity. *PLoS One* 7, e43031. doi: 10.1371/journal.pone.0043031

- Pallini, R., Guazzi, G. C., Cannella, C., and Cacace, M. G. (1991). Cloning and sequence analysis of the human liver rhodanese: comparison with the bovine and chicken enzymes. *Biochem. Biophys. Res. Commun.* 180, 887–893. doi: 10.1016/S0006-291X(05)81148-9
- Pastor-Satorras, R., Vázquez, A., and Vespignani, A. (2001). Dynamical and correlation properties of the internet. *Phys. Rev. Lett.* 87, 258701. doi: 10.1103/PhysRevLett.87.258701
- Power, D., Santos, N., Dieringer, M., Yu, J., Huang, H., Simpson, S., et al. (2015). IFI44 suppresses HIV-1 LTR promoter activity and facilitates its latency. *Virology* 481, 142–150. doi: 10.1016/j.virol.2015.02.046
- Rangel-Santiago, J. F., Baay-Guzman, G. J., Duran-Padilla, M. A., Lopez-Bochm, K. A., Garcia-Romero, B. L., Hernandez-Cueto, D. D., et al. (2016). A novel role of yin-yang-1 in pulmonary tuberculosis through the regulation of the chemokine CCL4. *Tuberculosis* 96, 87–95. doi: 10.1016/j.tube.2015.10.013
- Ravasz, E., and Barabási, A.-L. (2003). Hierarchical organization in complex networks. *Phys. Rev. E* 67, 026112. doi: 10.1103/PhysRevE.67.026112
- Reineke, L. C., and Lloyd, R. E. (2015). The stress granule protein G3BP1 recruits protein kinase R to promote multiple innate immune antiviral responses. *J. Virol.* 89, 2575–2589. doi: 10.1128/JVI.02791-14
- Reyes, N., Bettin, A., Reyes, I., and Geliebter, J. (2015). Microarray analysis of the in vitro granulomatous response to Mycobacterium tuberculosis H37Ra. *Colomb. Medica Cali Colomb.* 46, 26–32. doi: 10.25100/cm.v46i1.1570
- Sagi, Z., and Hieronymus, T. (2018). The impact of the epithelial–mesenchymal transition regulator hepatocyte growth factor receptor/met on skin immunity by modulating Langerhans cell migration. *Front. Immunol.* 9, 517. doi: 10.3389/fimmu.2018.00517
- Salamon, H., Bruiners, N., Lakehal, K., Shi, L., Ravi, J., Yamaguchi, K. D., et al. (2014). Cutting edge: vitamin D regulates lipid metabolism in Mycobacterium tuberculosis infection. *J. Immunol. Baltim. Md 1950* 193, 30–34. doi: 10.4049/jimmunol.1400736
- Sarkar, S., Leo, B. F., Carranza, C., Chen, S., Rivas-Santiago, C., Porter, A. E., et al. (2015). Modulation of human macrophage responses to Mycobacterium tuberculosis by silver nanoparticles of different size and surface modification. *PLOS One* 10, e0143077. doi: 10.1371/journal.pone.0143077
- Schaaf, K., Smith, S. R., Duverger, A., Wagner, F., Wolschendorf, F., Westfall, A. O., et al. (2017). Mycobacterium tuberculosis exploits the PPM1A signaling pathway to block host macrophage apoptosis. *Sci. Rep.* 7, 42101. doi: 10.1038/srep42101
- Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M., et al. (2003). Transcriptional adaptation of Mycobacterium tuberculosis within macrophages: insights into the phagosomal environment. *J. Exp. Med.* 198, 693–704. doi: 10.1084/jem.20030846
- Scott-Browne, J. P., Shafiani, S., Tucker-Heard, G., Ishida-Tsubota, K., Fontenot, J. D., Rudensky, A. Y., et al. (2007). Expansion and function of Foxp3-expressing T regulatory cells during tuberculosis. *J. Exp. Med.* 204, 2159–2169. doi: 10.1084/jem.20062105
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 3, Article3. doi: 10.2202/1544-6115.1027
- Soares-Silva, M., Diniz, F. F., Gomes, G. N., and Bahia, D. (2016). The mitogen-activated protein kinase (MAPK) pathway: role in immune evasion by trypanosomatids. *Front. Microbiol.* 7, 183. doi: 10.3389/fmicb.2016.00183
- Sundaravel, S., Duggan, R., Bhagat, T., Ebenezer, D. L., Liu, H., Yu, Y., et al. (2015). Reduced *DOCK4* expression leads to erythroid dysplasia in myelodysplastic syndromes. *Proc. Natl. Acad. Sci.* 112, E6359–E6368. doi: 10.1073/pnas.1516394112
- Thuong, N. T. T., Dunstan, S. J., Chau, T. T. H., Thorsson, V., Simmons, C. P., Quyen, N. T. H., et al. (2008). Identification of tuberculosis susceptibility genes with human macrophage gene expression profiles. *PLoS Pathog.* 4, e1000229. doi: 10.1371/journal.ppat.1000229
- Ting, J. P.-Y., Willingham, S. B., and Bergstralh, D. T. (2008). NLRs at the intersection of cell death and immunity. *Nat. Rev. Immunol.* 8, 372–379. doi: 10.1038/nri2296
- Traag, V. A., Krings, G., and Van Dooren, P. (2013). Significant scales in community structure. *Sci. Rep.* 3, 2930. doi: 10.1038/srep02930
- Traag, V. A., Van Dooren, P., and Nesterov, Y. (2011). Narrow scope for resolution-limit-free community detection. *Phys. Rev. E* 84, 016114. doi: 10.1103/PhysRevE.84.016114
- van der Windt, G. J. W., Florquin, S., de Vos, A. F., van't Veer, C., Queiroz, K. C. S., Liang, J., et al. (2010). CD44 deficiency is associated with increased bacterial clearance but enhanced lung inflammation during gram-negative pneumonia. *Am. J. Pathol.* 177, 2483–2494. doi: 10.2353/ajpath.2010.100562
- Vereecke, L., Beyaert, R., and van Loo, G. (2011). Genetic relationships between *A20/TNFAIP3*, chronic inflammation and autoimmune disease. *Biochem. Soc. Trans.* 39, 1086–1091. doi: 10.1042/BST0391086
- Verway, M., Bouttier, M., Wang, T.-T., Carrier, M., Calderon, M., An, B.-S., et al. (2013). Vitamin D induces interleukin-1 β expression: paracrine macrophage epithelial signaling controls M. tuberculosis infection. *PLoS Pathog.* 9, e1003407. doi: 10.1371/journal.ppat.1003407
- Villaseñor, T., Madrid-Paulino, E., Maldonado-Bravo, R., Urbán-Aragón, A., Pérez-Martínez, L., and Pedraza-Alva, G. (2017). Activation of the Wnt pathway by Mycobacterium tuberculosis: a Wnt–Wnt Situation. *Front. Immunol.* 8, 50. doi: 10.3389/fimmu.2017.00050
- Wan, X.-K., Yuan, S.-L., Wang, Y.-C., Tao, H.-X., Jiang, W., Guan, Z.-Y., et al. (2016). *Helicobacter pylori* inhibits the cleavage of TRAF1 via a CagA-dependent mechanism. *World J. Gastroenterol.* 22, 10566. doi: 10.3748/wjg.v22.i48.10566
- Wang, H., El Maadidi, S., Fischer, J., Grabski, E., Dickhöfer, S., Klimosch, S., et al. (2015a). A frequent hypofunctional *IRAK2* variant is associated with reduced spontaneous hepatitis C virus clearance. *Hepatology* 62, 1375–1387. doi: 10.1002/hep.28105
- Wang, J., Wang, Z., Yao, Y., Wu, J., Tang, X., Gu, T., et al. (2015b). The fibroblast growth factor-2 arrests Mycobacterium avium sp. paratuberculosis growth and immunomodulates host response in macrophages. *Tuberc. Edinb. Scotl.* 95, 505–514. doi: 10.1016/j.tube.2015.04.006
- Wang, M., Gu, D., Du, M., Xu, Z., Zhang, S., Zhu, L., et al. (2016). Common genetic variation in ETV6 is associated with colorectal cancer susceptibility. *Nat. Commun.* 7, 11478. doi: 10.1038/ncomms11478
- Zaiss, D. M. W., Gause, W. C., Osborne, L. C., and Artis, D. (2015). Emerging functions of amphiregulin in orchestrating immunity, inflammation, and tissue repair. *Immunity* 42, 216–226. doi: 10.1016/j.immuni.2015.01.020
- Zhang, X., Huang, T., Wu, Y., Peng, W., Xie, H., Pan, M., et al. (2017). Inhibition of the PI3K-Akt-mTOR signaling pathway in T lymphocytes in patients with active tuberculosis. *Int. J. Infect. Dis.* 59, 110–117. doi: 10.1016/j.ijid.2017.04.004

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Alam, Imam, Ahmed, Tazyeen, Tamkeen, Farooqui, Malik and Ishrat. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.