Beni-Suef University Journal of Basic and Applied Sciences

RESEARCH

Open Access

Exploring key molecular signatures of immune responses and pathways associated with tuberculosis in comorbid diabetes mellitus: a systems biology approach

G. Tamizh Selvan[1], Pavan Gollapalli[1,2*], Praveenkumar Shetty[1,3] and N. Suchetha Kumari[1,3]

Abstract

Background: Comorbid type 2 diabetes mellitus (T2DM) increases the risk for tuberculosis (TB) and its associated complications, although the pathological connections between T2DM and TB are unknown. The current research aims to identify shared molecular gene signatures and pathways that affirm the epidemiological association of T2DM and TB and afford clues on mechanistic basis of their association through integrative systems biology and bioinformatics approaches. Earlier research has found specific molecular markers linked to T2DM and TB, but, despite their importance, only offered a limited understanding of the genesis of this comorbidity. Our investigation used a network medicine method to find possible T2DM-TB molecular mediators.

Results: Functional annotation clustering, interaction networks, network cluster analysis, and network topology were part of our systematic investigation of T2DM-TB linked with 1603 differentially expressed genes (DEGs). The functional enrichment and gene interaction network analysis emphasized the importance of cytokine/chemokine signalling, T cell receptor signalling route, NF-kappa B signalling pathway and Jak-STAT signalling system. Furthermore, network analysis revealed significant DEGs such as *ITGAM* and *STAT1*, which may be necessary for T2DM-TB immune responses. Furthermore, these two genes are modulators in clusters C4 and C5, abundant in cytokine/chemokine signalling and Jak-STAT signalling pathways.

Conclusions: Our analyses highlight the role of ITGAM and STAT1 in T2DM-TB-associated pathways and advances our knowledge of the genetic processes driving this comorbidity.

Keywords: Gene interaction network, Functional enrichment analysis, Tuberculosis, type 2 diabetes mellitus, Clustering analysis, *ITGAM*, *STAT1*

1 Background

The vital challenge of integrating public health and care delivery for comorbidity has been raised due to non-communicable and communicable diseases [1]. In the world's tuberculosis (TB) epidemic regions, the prevalence of TB comorbidity with non-communicable (malnutrition, smoking, excessive alcohol use, diabetes) and other communicable (HIV/AIDS, malaria, Influenza, Helminths) diseases is very high. Amongst those, one of the major public health concerns in present decade is the adverse effect of acquired host-derived factors like type 2 diabetes mellitus (T2DM) that increases the risk and severity of TB. In India, a recent report suggests that nearly 14.8% of TB cases are in pre-diabetes patients [2]. It is considered that the risk of TB in T2DM individuals is nearly threefold greater when

*Correspondence: gollapallipavan@nitte.edu.in

[1] Central Research Laboratory, KS Hegde Medical Academy, Nitte (Deemed to Be University), Mangalore, Karnataka 575018, India
Full list of author information is available at the end of the article

compared to non-diabetic patients [3]. The prevalence of T2DM is projected to increase by 67% by 2035; in India, the co-burden of T2DM-TB may lead to major public health crises [4].

The underlying biochemical and epidemiological interactions between T2DM and TB remain too limited. Perhaps, the association between TB and T2DM was the oldest to be identified, but the molecular mechanisms from the immunological point of view behind this comorbidity have only been explored recently [5, 6]. An altered humoral innate immune response is mediated through cytokine levels. Thus, the macrophages and polymorphonuclear cells lose functioning and exhibit altered chemotaxis, phagocytosis and killing efficiency in T2DM patients [7]. Few studies have demonstrated that in addition to increased absolute neutrophil count in blood, the elevated plasma levels of HO-1, TIMP-4 and angiogenic factors (VEGF-A, C, D, Ra, R2, R3, angiopoietin 1, angiopoietin 2 and Tie 2 receptors) act as potential biomarkers for monitoring therapeutic responses in T2DM-TB comorbidity [8]. It has been identified that an altered metabolic environment compromises the host immune responses and produces changes that allow infection and facilitate disease progression of TB at the onset of the pre-diabetes stage. Impaired insulin signalling, glucose and lipid homeostasis alter the gene signalling profile in both organs and serum during prediabetic and diabetic conditions [9].

A bidirectional cost-effective screening strategy is needed in countries with high T2DM-TB comorbidity. Routine techniques like fasting blood glucose testing, oral glucose tolerance test combined with HbA1c for diabetes and sputum smearing examination, and chest radiography for TB are used for screening. However, these methods are considered as low-sensitive methods. Hence, there is an urgent need to develop feasible, more sensitive and cost-effective screening technologies to diagnose both the diseases. However, in people with T2DM, it is difficult to change the recommended standard TB treatment regimens or specify clinical case management of TB due to inadequate data availability. Very poor control of T2DM on the pathogenesis of TB provides a rational basis for testing with combined anti-microbial and anti-inflammatory therapies in T2DM patients with TB. The dual burden of this comorbidity can be alleviated by a better understanding of the molecular basis of TB susceptibility in T2DM and by developing a new diagnostic and therapeutic strategy. To understand the molecular complexity that hides the factors that enhance the T2DM patients for TB susceptibility, we present an integrative approach of network biology to identify and validate a distinct genes signature as a potential biomarker for early diagnosis and formulate future therapeutic use.

2 Methods

2.1 Data extraction of differentially expressed genes (DEGs) based on the literature

The primary source of seed genes for the network creation and topological analysis was retrieved through a literature survey specific to whole blood/serum gene expression data [9].

2.2 Gene ontology (GO) term and pathway enrichment analysis

The functional enrichment analysis allows us to understand all genes' biological information and functions. In a hierarchically structured method, annotations are used to categorize genes into biological/cellular/molecular keywords. GO is a database that contains a structured, unique, and regulated vocabulary of gene annotation and combines large-scale biological data [10]. The Kyoto Encyclopedia of Genes and Genomes (KEGG) database collects online genomes, enzymatic pathways, and biological substances databases [11]. DAVID (Database for Annotation, Visualization, and Integrated Discovery) Bioinformatics Resources 6.8 was used to perform functional annotation of the nodes in the cluster following the clustering of the protein–protein interaction (PPI) network [12]. The GO and KEGG pathway analyses are included in the functional annotation. The GO categories analysis provides a standard descriptive framework for functionally annotating and categorizing gene sets. GO categories include biological process, cellular component, and molecular function. KEGG pathways employ an artificial pathway diagram to link molecular interactions and reaction networks. The Benjamini approach was used to control the false discovery rate to fix the P value. The Benjamini approach is helpful in large-scale multiple testing challenges based on discrete test statistics. Its asymptotic (as the number of hypotheses approaches infinity) properties are obtained, superseding prior findings.

2.3 Integrated gene interaction network construction and topology analysis

Based on the literature study, the DEGs data were used to construct the gene interaction network. To generate the gene interaction network, selected genes were queried against the Search Tool for the Retrieval of Interacting Genes/Proteins database (STRING 11.0) [13]. STRING uses a list of proteins/genes as input to find the closest neighbours with direct connections. The STRING network was constructed using text mining, experiments, databases, co-expression, neighbourhood, gene fusion, and co-occurrence. Each interaction was assigned a confidence level or a combined score [14]. The combined scores typically range from 0 to 1, with the lowest number

indicating a low probability of occurrence and the highest number indicating a high probability of occurrence [15, 16]. The Cytoscape 3.3.0 program displays the gene interaction network, and default parameters were used to determine network node properties [17].

A topological study of the interaction network was performed using the Cytoscape plug-in NetworkAnalyzer. The topological parameters calculated include the number of nodes, connecting edges, network diameter, density, radius, centralization, heterogeneity, clustering coefficient, characteristic path length, distribution of node degrees, neighbourhood connectivity, average clustering coefficients and shortest path lengths [18]. The two major topological characteristics in network theory, degree (*k*) and betweenness centrality (BC), were used to evaluate the nodes in a network. The average distance between nodes is also measured, known as closeness centrality (CC). The shortest path is defined as the length of all geodesics from or to a network vertex. The node with the highest BC significantly impacts network traffic, making it an important global attribute in a network. Network diameter, shortest path length and diameter are other topological features of the biological network produced [19].

2.4 Network cluster analysis

A complex biological network is intricately linked to a complex biological process, and several subnetworks or functional modules (clusters) of proteins are involved. These modules impact each participating node in the network with a specific role, regardless of how they affect the core network [20, 21]. The densely connected regions in the core network were predicted using module analysis using Molecular Cluster Detection (MCODE) 4.1, a Cytoscape plug-in. The local neighbourhood density of weighted nodes was used to identify the network's heavily linked areas [22]. All of the parameters were kept constant, including the degree threshold (2), node score threshold (0.2), k-core threshold (2) and network max depth (100). Each vertex in a subgraph is specified by k-core and has a degree of at least '*k*' [14]. MCODE will be unaffected by the expected high false-positive rate in large-scale network interaction data. After the PPI network was clustered, DAVID Bioinformatics Resources 6.8 was used to perform functional annotation of the nodes in the cluster.

3 Results

3.1 Extraction of DEGs data

The literature-based screening related to T2DM-TB from the PubMed database was performed to integrated gene interaction network analysis. We selected the gene expression data specific to whole blood/serum from the published literature (PMID: 28,515,464). Based on the preliminary results, we used the following inclusion criteria: $p < 0.05$ and $|\log FC| \geq 1.0$, yielding 1603 differentially expressed genes (Additional file 1: Table S1).

3.2 Functional enrichment of the GO terms and KEGG pathways

The DAVID tool and GO functional enrichment analysis were used to decode the functionalities of a large number of genes. There were 567 GO keywords altogether, with 348, 84 and 85 GO terms relevant to biological processes (Fig. 1A), cellular components (Fig. 1B) and molecular function (Fig. 1C), respectively. Similarly, eleven new KEGG pathway ontology elements have been added (Fig. 1D). The top 10 biological processes enriched include cell surface receptor signalling pathway, SRP-dependent co-translational protein targeting to membrane, negative regulation of viral genome replication, interferon-γ (IFN-γ)-mediated signalling pathway, adaptive immune response, type I IFN signalling pathway, defence response to virus inflammatory response, innate immune response and immune response. The cellular components include membrane raft, plasma membrane, focal adhesion, nucleoplasm, membrane, T cell receptor complex, cytoplasm, external side of the plasma membrane, extracellular exosome and cytosol. The molecular functions include protein homodimerization activity, ATP binding, GTPase activity enzyme binding, transmembrane signalling receptor activity, virus receptor activity, double-stranded RNA binding, receptor activity, poly(A) RNA binding and protein binding. The top KEGG pathways enriched include haematopoietic cell lineage, primary immunodeficiency, T cell receptor signalling pathway, phagosome, NF-kappa B signalling pathway and TB.

3.3 Integrated gene interaction network

The STRING server's principal interaction data created an extended core network that included unweighted and undirected binary interactions. The extended core network consists of 1115 nodes connected by 6703 edges (Fig. 2). The extended network includes one giant network composed of 1058 nodes connected with 6672 edges (Fig. 3) and 4 separated small components that are derived from the seed proteins, extended synaptotagmin-1 (ESYT1), SCY1-like protein 2 (SCYL2), outer dynein arm-docking complex subunit 4 (TTC25) and ankyrin repeat domain-containing protein 50 (ANKRD50). The centrality metrics were utilized in a biological network to connect the communicative nodes that reflect the importance of functioning genes. The value of a gene is precisely proportionate to its relevance in linking regulatory molecules. Three centralities were

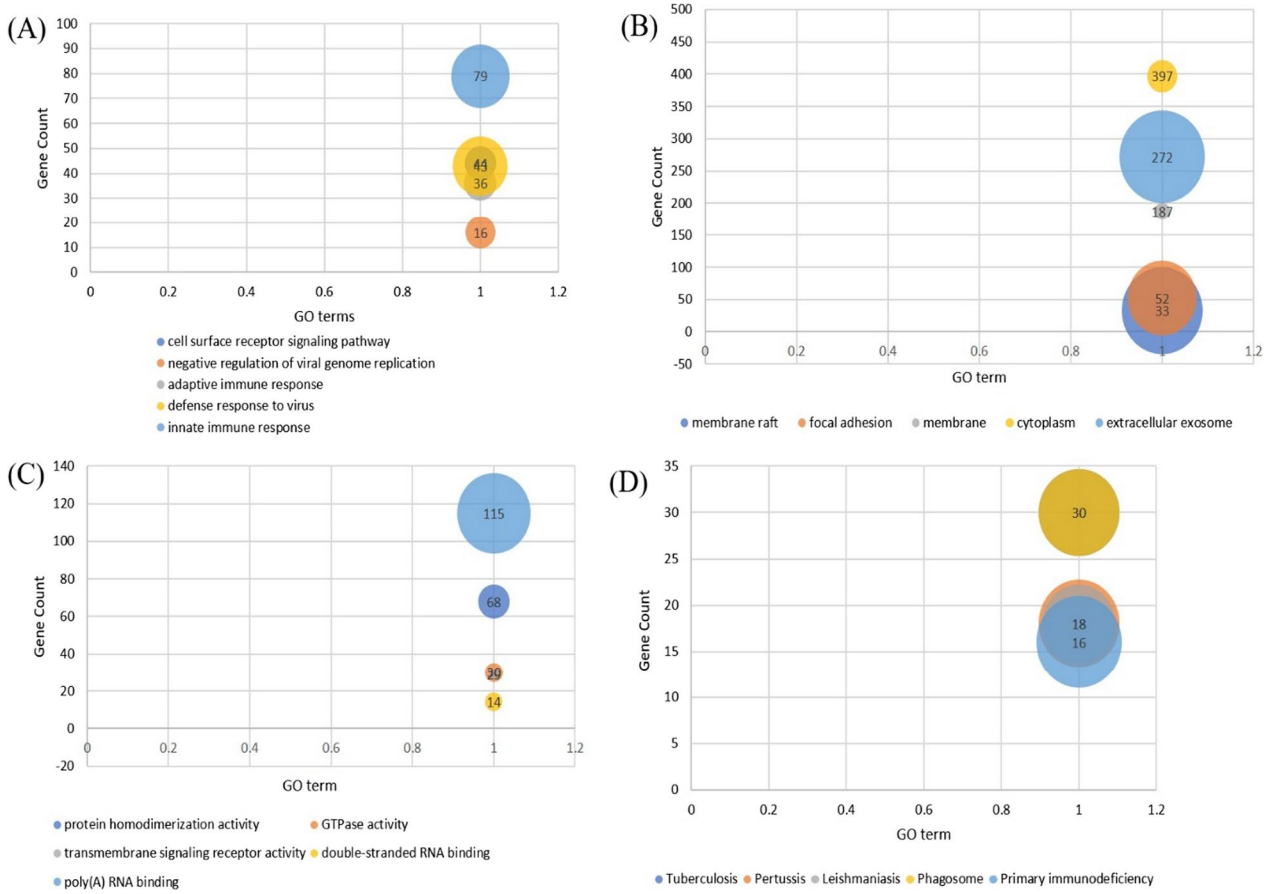

Fig. 1 The functional enrichment analysis of DEGs was performed by DAVID (https://david.ncifcrf.gov/tools.jsp). GO functional enrichment analysis results from DEGs relating to **A** biological process, **B** cellular components, **C** molecular function and **D** KEGG pathways. The size of the bubble represents the number of genes associated with the terms

estimated for each gene in the PPI network: degree, BC and CC. The number of interactions between a specific node and all other nodes was denoted by the degree of its nodes.

With a moderate degree of 12.61 and a clustering value of 0.441, the giant network has a significant degree of 73. This recovers the fundamental PPI network property of a few nodes with fewer connections and a small number of highly connected nodes [23]. The shortest pathways in the network will connect the two randomly chosen nodes, and a histogram was used to display the network's BC, CC, average clustering coefficient and distribution (Fig. 4A–D). Table 1 illustrates the results of each node in the core network's topological parameter analysis, including degree (k), BC and CC. The BC in the network was 0.002, and the average shortest path length was 4.025. The average number of neighbours for 1058 nodes was 12.61. The complete DEGs network centrality distribution list may be seen in Additional file 1: Table S2. Topological variables such as degree, average shortest path length and clustering coefficient explain these relationships [14].

Lower average shortest path length and larger degree and clustering coefficient showed strong linkages between DEGs. ITGAM (integrin alpha M) is a hub protein with the most significant degree ($k=73$) and seventh highest BC value of 0.0366, whereas MYC (Myc proto-oncogene protein) is a bottleneck protein with the highest BC value of 0.0567, according to our comprehensive network topology study. STAT1 (signal transducer and activator of transcription 1-alpha/beta) has the highest CC value of 0.3544, indicating that it is at the centre of the network. Based on the degree values, Table 2 displays the top ten genes with the most functional interactors and Table 3 shows the top 20 genes with the shortest average path length and highest nearest centrality values. The genes with the most direct connections are the network's hub molecules, and these genes are the molecular interactions' regulatory points. Hub genes can be utilized as therapeutic targets in creating novel medicines

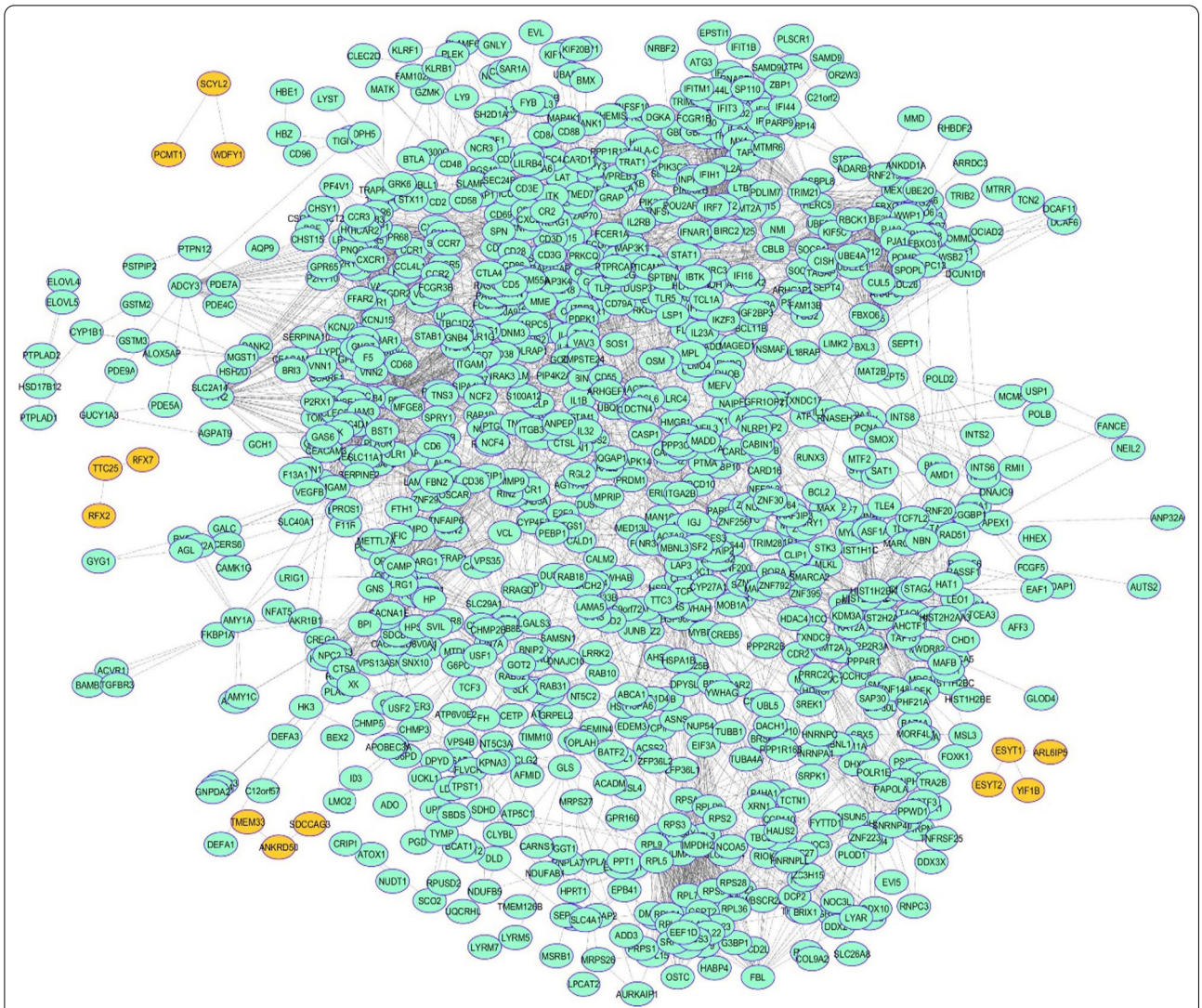

Fig. 2 Overview of the extended network includes one giant network (green) and 4 separate small components (orange) which are derived, respectively

and can assist in a better understanding of molecular processes.

3.4 Significance of key nodes in the network

A small number of highly interconnected nodes in a scale-free distribution network are more significant than any other less-connected nodes [22, 24]. Even though nodes are involved in a small number of processes with little connectivity, they have a high BC. They are more conserved than other proteins because they serve as a connection between different modules, making them bottleneck proteins. Furthermore, nodes with BC and/or degree values greater than the mean plus standard deviation were chosen. There were 192 nodes with high degree values (Additional file 1: Table S4 and Fig. 5A) and 102 nodes with high BC values (Additional file 1: Table S5 and Fig. 5B). Besides, 70 nodes with substantial degrees and BC were discovered (Table 4). MYC, mitogen-activated protein kinase 14 (MAPK14), epidermal growth factor (EGF), signal transducer and activator of transcription (STAT) were among the top 20 critical genes with both a large degree and a high BC. Using graph theory, it is possible to uncover the hidden properties of biological communication systems. Protein–protein interaction networks can efficiently evaluate and estimate the possibility of existing but undiscovered interactions between proteins/genes [25]. The biological relevance of proteins was associated with topological properties in several PPI

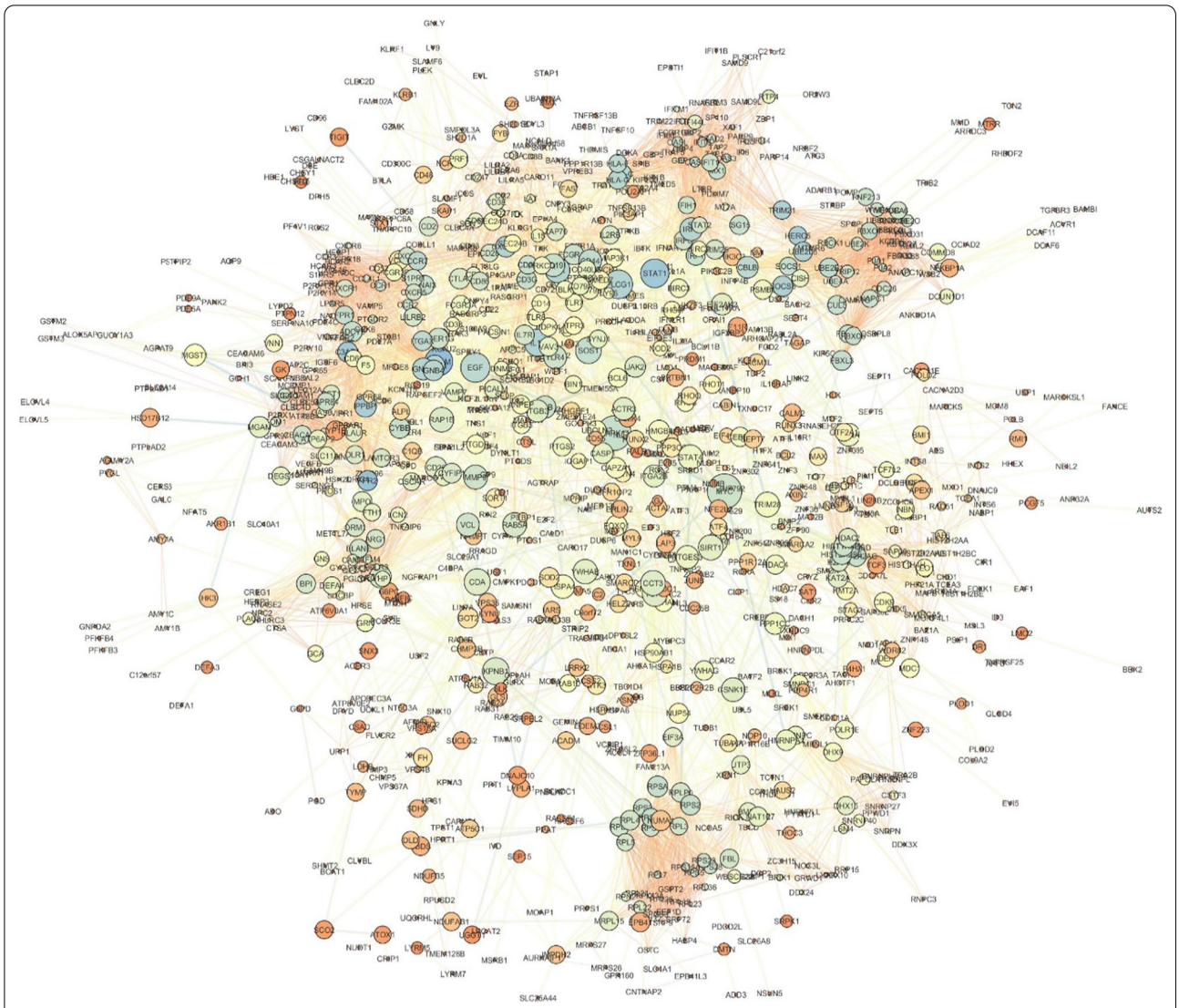

Fig. 3 Topology of the giant network extracted from the extended network is the biggest component in the extended network. It consisted of 1058 nodes and 6672 edges. Key nodes in the giant network are highlighted in different colours (low value to bright colour). The size of the nodes corresponds to their BC values (low value to small size; low values are represented in orange colour and large value with blue colour)

networks. Its connection shows the gene/protein and their topological responsibilities, known as hubs, may be categorized depending on their location [26].

3.5 Backbone network analysis

In the backbone network, HSPA4 has the greatest degree and BC value. On the other hand, EGF was discovered in the centre with the highest CC value, suggesting that EGF may be involved in the formation of the T2DM-TB immunological response (Fig. 6 and Additional file 1: Table S6). In granulomatous tissues and macrophages, EGF is a growth factor for pathogenic mycobacteria, and it may assist both intracellular and extracellular mycobacteria to grow faster at the infection site [27]. The host response to mycobacteria is connected to necrosis with or without granuloma formation. EGF, fibroblast growth factor and transforming growth factor are cytokines and growth factors. In areas of caseation necrosis and granulomatous inflammation, identification of strongly linked regions in the PPI network is common [28]. As central neighbours, the *ITGAM* and *STAT1* genes have been demonstrated to interact directly with the EGF.

3.6 Cluster analysis of gene interaction network

The MCODE technique was utilized from a gene interaction network to discover strongly linked proteins.

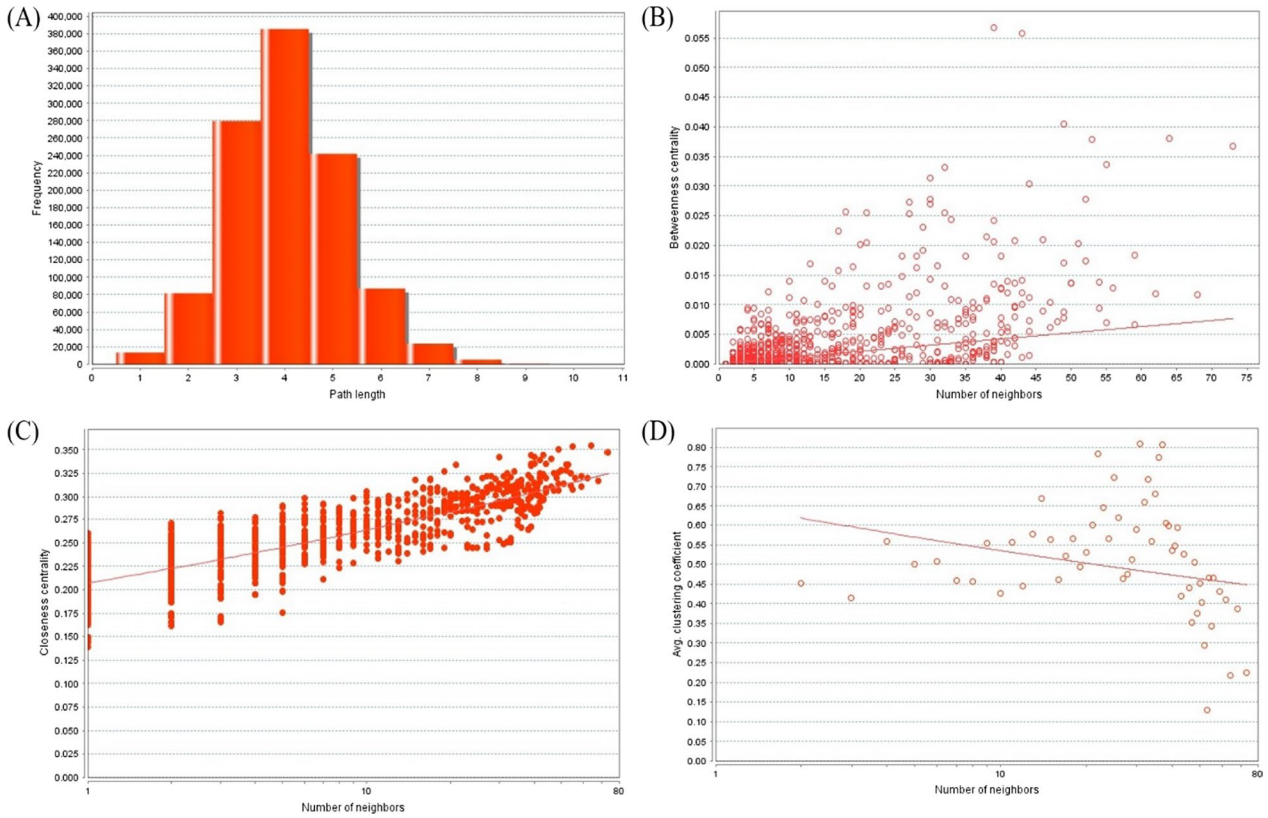

Fig. 4 Topological analysis of a gene interaction network for DEGs with 1058 nodes and 6672 edges. **A** The histogram represents the shortest path length distribution, demonstrating the network's small-world feature; **B** BC distribution fitted by a power law; **C** represents the clustering coefficient of the network's nodes; **D** average clustering coefficient of the network's nodes

Table 1 General network measurements for giant and backbone networks for the DEGs in T2DM-TB

Network parameters	Giant network	Backbone network
Number of nodes	1058	52
Number of edges	6672	296
Clustering coefficient	0.441	0.492
Network diameter	10	4
Network centralization	0.057	0.257
Shortest path	1,118,306	2652
Characteristic path length	4.026	1.990
Average number of neighbours	12.612	11.385
Correlation values by fitting power law for the network		
Betweenness centrality (BC)	0.521	0.733
Closeness centrality (CC)	0.828	0.943
Average clustering coefficient distribution	0.141	0..568
Neighbourhood connectivity distribution	0.868	0.488
Topological coefficient	0.770	0.444

Table 2 List of top 10 genes with a greater number of functional interactors in the gene interaction network of DEGs in T2DM-TB

Gene name	Description	Degree	Functional partners
ITGAM	Integrin alpha M	73	GPR97, CYBB, MMP9, ATP8B4, CERCAMB, SLC2A3, CD68, CLAME, CLEC5A, CLEA12A, BST1, MRO, ARG1, MCEMF1, SELP, RAP1B, CLEC4D, OLR1, ATP6AP2, RAP20, MGAM, CD69, CR1, GPR84, TOM1, P2RX1, SLC11A1, DEGS1, LILRB2, CEACAM1, JAM13, TLR4, CCR1, CD36, LAMTOP3, CEACAM13, ITGB5, NBEA1, PLAUR, CTLA4, IL-7R, CCR7, C3AR1, CCR2, ITHAX, ITGB3, IL-1β, TLR8, CPN, FPR2, CD5, CD19, ITGB7, IL-15, TLR7, TLR1, CD14, FGER1G, CXCL-10, IL-10, CD28, CD2, CD3E, STAT1, FCGR1A, ANPEP, CD48, HMGB1, CD40LG, CD44, IRF8, SLAMF1
FPR2	N-formyl peptide receptor 2	68	CXCL10, CCR7, C3AR1, CCR2, ITGAX, FCER1G, TLR8, ITGAM, PLAUR, NBEAL2, CEACAM3, LAMTOR3, CD36, CCR1, GPR84, TOM1, P2RX, SLC11A1, DEGS1, LILRB2, CEACAM1, MCEMP1, MGAM, RAP20, ATP6AP2, OLR1, CLEC4D, CLEC5A, CLEC12, BST1, MCEMP1, SLC2A3, CEACAM8, ATP8B4, CYBB, GPR97, GNB4, GPR18, P2RY14, CAMP, ADCY3, HEBP1, PPBP, S1PR1, CCL4L1, P2RY13, LPAR5, S1PR5, HCAR3, FPR1, GPR65, GNG7, F2ML, CCR3, HCAR2, GNAI3, CXCR1, CXCL16, PTGDR2, CXCR6, GNG11, GPR68, P2RY10, FFAR2, PNOC, GPR183, CXCR5
STAT1	Signal transducer and activator of transcription 1-alpha/beta	64	CXCR3, PTGS2, IL-1β, ITGAM, ISG15, FGFR10P2, IL-10, LAP3, EGF, IRF8, PLCG1, TLR8, EIF2AK2, IFI44L, SOCS3, CXCL10, OASL, IRF7, JAK2, SOCS1, LCK, BCL8, MAPK14, CUX1, RUNX2, STAT4, IL-2Rβ, UBE2E1, TRIM25, UBE2L6, IFI35, HERC5, OSM, OAM, OAS3, MX1, GBP2, XAF1, IFIH1, GBP5, OAS1, IFI6, IFGAX, STAT2, GBP1, IL-10Rβ, IFNAR1, IL-23A, LMC4, IFNLR1, CISH, TRIM22, RSAD2, IL-23A, LMC4, IFNLR1, CISH, TRIM22, RSAD2, IGIT3, IGIT1, PSMB9, SP110, IFIT2, IFIT2, IFI44, SAMD9L, PARP9, PARP14
C3AR1	C3a anaphylatoxin chemotactic receptor	62	TLR4, CD86, ITGAM, CXCL10, CXCR3, RAR1B, CD68, CYBB, CEACAMB, CCR3, PTGDR2, CXCR5, CCR1, CCR7, C3AR1, CLEC5A, CLEC4D, GPR84, CD36, OLR1, CEACAM1, GNB4, ADCY3, GNA13, GPR183, TOM1, CCR2, LAMTOR3, P2RX1, CEACAN3, DEGS1, PLAUR, CCLAL1, PPBP, P2RY14, HEBP1, FPR1, HCAR3, HCAR2, GNG11, P2RY13, FCER1G, GNG7, GPR18, CXCR6, S1PR1, PNOC, LPAR5, S1PR5, CXB1
PPBP	Platelet basic protein	59	CAMP, MMP9, SELP, CXCL10, CXCR3, FPR2, CCR3, PTGDR2, CXCR5, CXCL16, SERP1NG1, OLFM4, MMRN1, GAS6, VEGFB, FAM49B, PROSI, FTH1, OSCAR, HP, GNB4, CCR1, CCR7, C3AR1, CDA, LRG1, FOLR3, QPCT, CTSD, METL7A, TCN1, PGLYRF1, ADCY3, GNAI3, GPR183, CCR2, TNFAIF6, VWF, CYFIP1, F13A, F5, ORM1, CCL41, P2RY14, HEBP1, FPR1, HCAR3, HCAR2, GNG11, P2RY13, EGF, GNG7, GPR18, CXCR6, S1PR1, PNOC, LPAR5, S1PR5, CXCR1
HERC5	E3 ISG15–protein ligase HERC5	59	KCTD6, CUL5, WWP1, PLCG1, STAT1, ANAPC1, EIF2AK2, IFI44L, FDXO7, FBXL5, SOCS1, ISG14, IRF7, OAS1, RNF138, SOCS3, TRIP12, U8E2E1, TRIM25, UBE2L6, IFI35, IFIT1, OAS3, MX1, PJA1, UBE2G2, FBXL3, TRIM21, RSAD2, IFIT3, EF4G2, IFI6, OAS1, RBCK1, IFCH1, XAF1, FBXO9, UBA6, CBLB, UBE2Q2, UBE2K, UBE20, FBXO31, EF4E3, IFI44, SAMD9L, CDC26, ASB8, KLHL2, FLNB, IFIT2, KBTBD9L, UBE4A, RNF213, MEX3C, PJA2, ANAPC1B, KLHL5
SOCS3	Suppressor of cytokine signalling 3	56	TRIP12, RNF13B, SOCS1, FBXL5, FBXO6, KCTD6, CUL5, WWP1, STAT1, ANAPC1, FBXO7, UBE2E, UBE2L6, HERC5, PJA1, UBE2G2, FBXL3, TRIM21, RBCK1, FBXO9, UBA6, CBLB, UBE2Q2, UBE2K, UBE20, FBXO31, CDC26, ASBB, KLHL3, ANAPC1, PJA2, MEX3C, RNF213, UBE4A, KBTBD7, KLHL2, IL-10, PTGS3, IL-1β, IRF1, JAK2, OSM, STAT2, LCK, IL2RB, LMO4, IL-23A, IFNAR1, CISH
IL-10	Interleukin-10	55	MAPK14, CCR7, CTLA4, PTGS2, IL-1β, STAT1, SOCS3, IL-2Rβ, IFNAR1, ARG1, STAT2, CXCL10, CXCR3, MMP9, CCR2, MPO, IL-1RN, CCR1, CD69, ITGAX, IL-23A, CD19, TLR8, ITGAM, HSPA4, DJSP1, FCGR1A, HLA-G, STAT4, IL-7R, FCFR2, IL-10RB, CCLAL1, EGF, IFNLR1, TLR5, CASP, FPR1, CD86, CD44, TLR1, TLR7, CD28, IL-15, TLR4, CD40LG, NFL3, PROM1, LILRB2, PRF1, CD27, CD83, CD68, TNFSF13B, NOD2
TRIM21	E3 ubiquitin-protein ligase TRIM21	55	HLA-G, KCTD6, CUL5, WWP1, FCGR1A, ANAPC1, CD44, FBXO7, TRAR1, SOCS3, IRF1, FBXO6, FBXL5, SOCS1, HLA-C, HLA-F, GBP4, RNF138, MT2A, TRIP12, IFI30, FCGR1B, UBE2E1, UBE2L6, IEF7, HERC5, OASL, OAS3, OAS1, PJA1, UBE2G2, GBP5, FBXL3, IFI35, GBP1, IRF8, RBCK1, GBP2, FBXO9, UBA6, CBLB, UBE2Q2, UBE2K, UBE20, FBXO31, CDC26, ASB8, KLML2, KBTBD7, UBE4A, RNF213, MEX3C, PJA2, ANAPC1, KLHL3
CXCL-10	C-X-C motif chemokine 10	54	CXCR3, MMP9, CCR2, CCR1, TLR4, TLR8, IL-15, IL-1β, IL-10, CD86, FPR1, IFI44L, GBP5, CCR7, STAT1, FBR2, CCR3, PTGDR2, CXCR5, CXCL16, OASL, IRF7, ISG15, RSAD2, CXCR5, CXCL16, OASL, IRF7, ISG15, RSAD2, IFIT3, IFIT1, MX1, CCL4L1, ADCY3, GNAI3, GPR183, OAS1, GBR1, C3AR1, IRF1, IFIT2, IFIH1, ITGAM, PPBP, P2RY14, HEBP1, HCAR3, HCAR2, GNG11, P2RY13, GNG7, GPP18, CXCR6, S1PR1, PNOC, LPAR5, S1PR5, CXCR1

Module or clustering analysis was carried out using the MCODE technique. The clusters were filtered using the features specified in the approach to ensure the efficiency of functional partners towards the core network of T2DM-TB DEGs. Based on the number of interactions between each node, the genes with most closely

Table 3 Topological parameter analysis using NetworkAnalyzer: All the genes in the network are analysed for the topological parameter (top 20 genes with high CC and least average shortest path length)

Genes	Avg shortest path length	Clustering coefficient	CC	Stree	Degree	BC	Neighbourhood connectivity	Radiality	Topological coefficient
STAT1	2.821192	0.216766	0.35446	544,986	64	0.037979	30.35938	0.817881	0.07397104
IL-10	2.826869	0.171044	0.353748	493,032	55	0.033681	28.34545	0.817313	0.06629764
EGF	2.852412	0.207483	0.35058	449,756	49	0.040504	26.14286	0.814759	0.06377551
ITGAM	2.882687	0.224886	0.346899	560,624	73	0.036693	29.83562	0.811731	0.07492944
MAPK14	2.900662	0.07309	0.344749	862,926	43	0.055699	23.09302	0.809934	0.06424581
TLR4	2.904447	0.17139	0.3443	276,442	39	0.020693	28.66667	0.809555	0.07713042
IL-1β	2.92053	0.15641	0.342404	243,614	40	0.018198	24.55	0.807947	0.0637013
JAK2	2.924314	0.172414	0.341961	179,722	30	0.014197	26.7	0.807569	0.07544854
CD44	2.935667	0.263768	0.340638	291,368	46	0.02095	32.3913	0.806433	0.08540373
CD19	2.950804	0.209059	0.338891	265,706	42	0.020744	25.2381	0.80492	0.06997354
FCGR1A	2.982025	0.320513	0.335343	239,952	40	0.012874	32.775	0.801798	0.1021028
SOS1	2.987701	0.124736	0.334706	354,612	44	0.030432	19.04545	0.80123	0.05614578
MYC	2.987701	0.097166	0.334706	611,210	39	0.056718	17.38462	0.80123	0.05597453
LCK	2.989593	0.168175	0.334494	451,424	52	0.027723	25.28846	0.801041	0.0680593
PLCG1	2.991485	0.129173	0.334282	531,098	53	0.037789	24.5283	0.800851	0.06662897
STAT4	2.997162	0.133333	0.333649	271,400	21	0.02555	22.71429	0.800284	0.07681018
MMP9	3.005676	0.287449	0.332704	299,480	39	0.024128	27.51282	0.799432	0.08581731
GNG7	3.047304	0.417798	0.328159	417,720	52	0.017441	29.46154	0.79527	0.11145105
GNG11	3.049196	0.452245	0.327955	380,978	50	0.013659	30.64	0.79508	0.11518797
GNB4	3.049196	0.452245	0.327955	380,978	50	0.013659	30.64	0.79508	0.11518797

The top 20 genes with the least average shortest path length and the highest closeness centrality values are listed. The average shortest path length gives the expected distance between the two connected nodes, and genes with the shortest path length and high closeness centrality are considered as the controlling points of molecular communication. The cluster coefficient value lies in between 0 and 1. The value 0 indicates the node with less than 2 neighbours

related interactors are grouped [29]. Seven closely connected groups emerged due to the clustering analysis of the genes in interaction network. Table 5 lists the cluster scores, and Fig. 7 depicts the proteins implicated in each cluster (C1–C6). Furthermore, the DAVID GO and pathway enrichment tool was grouped using the P value for each cluster (C1–C6) because a P value of 0.05 shows significant findings. The P values and functional annotations for functional partners such as biological process, molecular function and cellular component were examined.

4 Discussion

We found 1603 DEGs substantially linked to T2DM-TB in our current study from the literature search. We conducted comprehensive studies that included functional annotation clustering using GO and KEGG enrichment analysis, the development of interaction networks, and network cluster/module to gain the molecular understanding that might aid to discover novel underlying processes implicated in T2DM-TB. The functional enrichment analysis revealed that genes were primarily involved in the GO terms immune response, innate immune response, inflammatory response, type I IFN signalling pathway, adaptive immune response, as well as the KEGG pathways T cell receptor signalling pathway, primary immunodeficiency, haematopoietic cell lineage, NF-kappa B signalling pathway and TB, among others. A range of network topological analyses was used to identify proteins with high degree and/or BC and CC values. The top 10 genes with the largest number of enriched functional interactors were discovered to better understand protein functioning in cellular processes. The network's hub molecules are the genes with the most direct connections, and these genes are the regulatory sites of functional molecular interactions. These hub proteins may be exploited as therapeutic targets in developing novel drugs to understand chemical processes [30].

MYC is a bottleneck protein with the greatest BC value, while ITGAM is a hub protein with the largest and seventh highest BC values. STAT1 has the greatest CC value, indicating that it is closest to the network's core. ITGAM is involved in various adhesive interactions between monocytes, macrophages and granulocytes and the absorption of complement-coated particles and pathogens [31]. It is identical to CR-3, which binds the R-G-D peptide in C3b and is the receptor for the iC3b fragment of the third complement component. Fibrinogen, factor X and ICAM1 receptors for the integrin ITGAM/ITGB2.

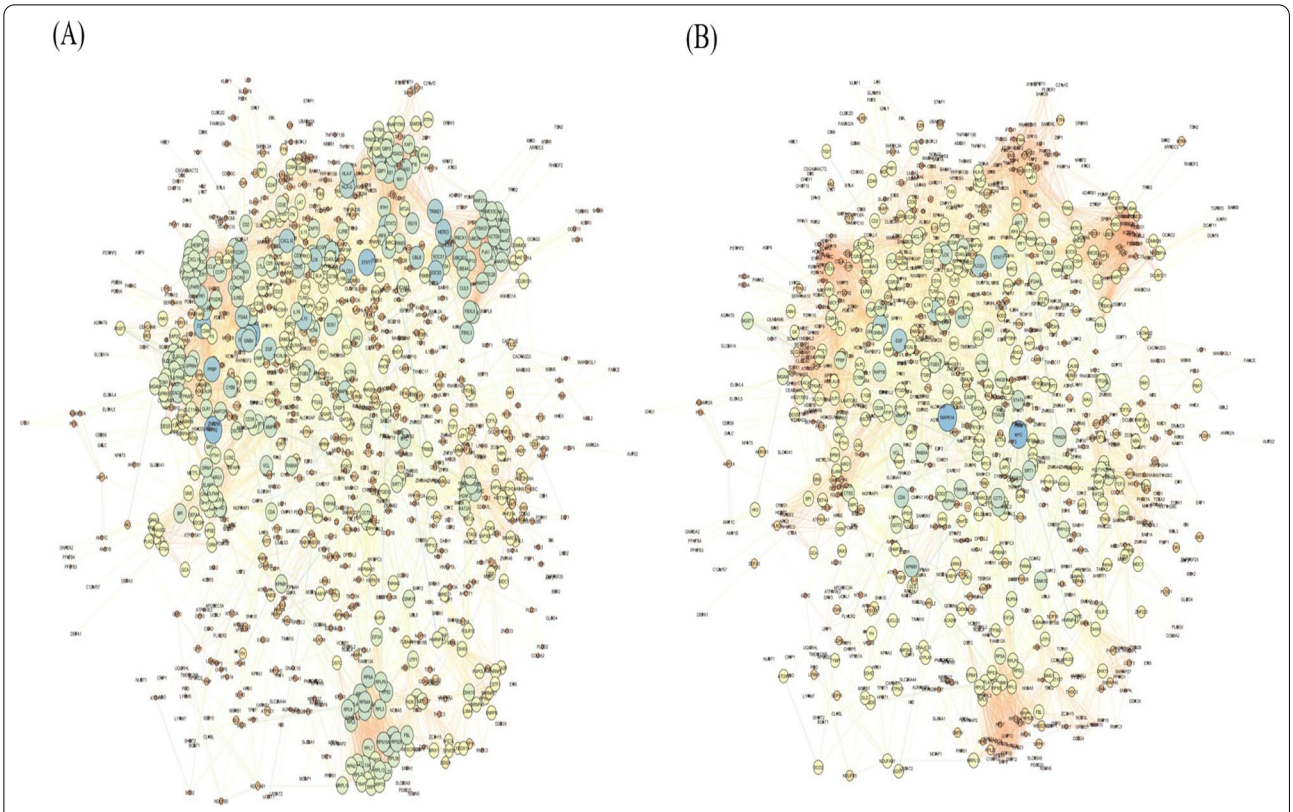

Fig. 5 Topology of the core network consisting of 1058 nodes connected via 6672 edges. **A** The key nodes were mapped with respect to degree and against the CC. The size of the nodes corresponds to their degree values (low value to smallest size) and colour to their CC values (bright colour); **B** the key nodes were mapped with respect to BC and against the CC. The size of the nodes corresponds to their BC values (low value to smallest size) and colour to their CC values (bright colour)

It detects fibrinogen gamma chain P1 and P2 peptides and controls neutrophil movement. CD177-PRTN3-mediated activation of neutrophils requires the beta subunit ITGB2/TNF-primed CD18. Apoptosis in extravasated neutrophils may be regulated by phagocytosis. This factor may influence mast cell development. Microglia works with TYROBP/DAP12 to regulate the generation of superoxide ions in the microglia, which induce neuronal death throughout brain development [32]. Integrins have been discovered to be needed for host control of infection in the case of tuberculosis (Fig. 8). In mice lacking CD11a (ITGAM) and CD18, survival significantly decreases. Furthermore, compared to wild-type mice, CD11a knock-out animals have a reduced capacity to control *M. tuberculosis* and have fewer effector T cells in the lungs [33].

Myc protein regulates proinflammatory cytokine responses and restricts mycobacteria intracellular growth by activating the IRAK1-dependent pathway. It would be fascinating to see if Myc has a similar role in diseases other than bacteria. Nonetheless, because Myc is required for various cellular responses in various cell types [34], differentiated blood macrophages from mice with inducible conditional knock-out of Myc should be investigated to determine the specific in vivo role of Myc in regulating innate immune responses. Understanding how induced these reactions are might lead to Myc-enhancing medicines to combat TB epidemics [35]. Cellular myelocytomatosis (cMyc) is a transcription factor that regulates cell proliferation and belongs to the proto-oncogene family [36]. Myc's N-terminal transactivation domain has merged with the basic helix–loop–helix leucine zipper domain, which binds to the CACGTG E-box DNA sequence [37]. This connection facilitates the recruitment of histone acetyltransferase and elongation factors, which can change the transcriptional response of numerous genes. By binding to open chromatin of glycolysis and glutaminolysis target genes, Myc controls metabolic reprogramming and allows for efficient transcription. Myc forms a dimer with Max, a DNA-binding helix–loop–helix leucine zipper protein, to alter gene expression [36]. *M. pneumoniae* infection of human peripheral blood mononuclear cells (PBMCs) with varied pathogenicity *M. pneumoniae* infection of

Table 4 List of DEGs (nodes) with large degree and high BC and their CC values

Genes	Functional description
MYC	Myc proto-oncogene protein
MAPK14	Mitogen-activated protein kinase 14
EGF	Epidermal growth factor
STAT1	Signal transducer and activator of transcription 1-alpha/beta
PLCG1	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-1
ITGAM	Integrin alpha M
IL-10	Interleukin-10
KPNB1	Importin subunit beta-1
CDA	Cytidine deaminase
SOS1	Son of sevenless homolog 1
TCP1	T-complex protein 1 subunit alpha
LCK	Tyrosine-protein kinase Lck
SIRT1	NAD-dependent protein deacetylase sirtuin-1
CCT3	T-complex protein 1 subunit gamma
RAP1B	Ras-related protein Rap-1b
YWHAB	14–3-3 protein beta/alpha
RAB5A	Ras-related protein Rab-5A
MMP9	Matrix metalloproteinase-9
CSNK1E	Casein kinase I isoform epsilon
VCL	Vinculin
CD44	CD44 antigen
CD19	B-lymphocyte antigen CD19
TLR4	Toll-like receptor 4
IRF7	Interferon regulatory factor 7
ACTR3	Actin-related protein 3
PPBP	Platelet basic protein
ITGB3	Integrin beta-3
CTSD	Cathepsin D
IL-1β	Interleukin-1 beta
GNG7	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-7
CBLB	E3 ubiquitin-protein ligase CBL-B
CD36	Platelet glycoprotein 4
EIF2AK2	Interferon-induced, double-stranded RNA-activated protein kinase
JAK2	Tyrosine-protein kinase JAK2
RPSA	40S ribosomal protein SA
HIST2H2AC	Histone H2A type 2-C
CXCL-10	C-X-C motif chemokine 10
GNG11	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-11
GNB4	Guanine nucleotide-binding protein subunit beta-4
CYBB	Cytochrome b-245 heavy chain
FCER1G	High affinity immunoglobulin epsilon receptor subunit gamma
CTLA4	Cytotoxic T-lymphocyte protein 4
FCGR1A	High affinity immunoglobulin gamma Fc receptor I
SOCS3	Suppressor of cytokine signalling 3
HIST2H2BE	Histone H2B type 2-E
KAT2A	Histone acetyltransferase KAT2A
RPS2	40S ribosomal protein S2
ADCY3	Adenylate cyclase type 3
C3AR1	C3a anaphylatoxin chemotactic receptor
FPR2	N-formyl peptide receptor 2
MPO	Myeloperoxidase
FBXL3	F-box/LRR-repeat protein 3
HIST1H2BD	Histone H2B type 1-D
CYFIP1	Cytoplasmic FMR1-interacting protein 1
HIST1H2BK	Histone H2B type 1-K
IL-7R	Interleukin-7 receptor subunit alpha
RPL4	60S ribosomal protein L4
ITGAX	Integrin alpha-X
CD3G	T cell surface glycoprotein CD3 gamma chain
GNAI3	Guanine nucleotide-binding protein G(i) subunit alpha-3
IRF1	Interferon regulatory factor 1
ISG15	Ubiquitin-like protein ISG15
UBE2L6	Ubiquitin/ISG15-conjugating enzyme E2 L6
RPLP0	60S acidic ribosomal protein P0
RPL9	60S ribosomal protein L9
BPI	Bactericidal permeability-increasing protein
UBE2E1	Ubiquitin-conjugating enzyme E2 E1
MRPL15	39S ribosomal protein L15, mitochondrial
TRIM25	E3 ubiquitin/ISG15 ligase TRIM25
HLA-G	HLA class I histocompatibility antigen, alpha chain G

human PBMCs with different mycobacterial species. The Wnt/beta-catenin signalling system was discovered to activate cMyc via the MAPK/ERK pathway, leading to the overexpression of essential cytokines such as TNF-α and IL-6, which limit mycobacterial development [35]. Myc was implicated in the anti-mycobacterial response in this case without impacting cell proliferation or changing the G0/G1 cell cycle phase of macrophages.

STAT1 modulates cellular responses to IFNs, the cytokine KITLG/SCF, and other cytokines and growth factors as a signal transducer and transcription activator [38]. Signalling via protein kinases occurs after type I IFN (IFN-α and IFN-β) binds to cell surface receptors, resulting in the activation of Jak kinases (TYK2 and JAK1) and tyrosine phosphorylation of STAT1 and STAT2 [38]. STAT1 is phosphorylated on tyrosine and serine in response to type II IFN (IFN-γ) [39]. It then forms an IFN-γ activated factor homodimer, migrates to the nucleus and binds to the IFN-γ activated sequence, turning the cell antiviral. It responds to KITLG/SCF and KIT signals by being active. Activated FGFR1, FGFR2, FGFR3 and FGFR4 may mediate

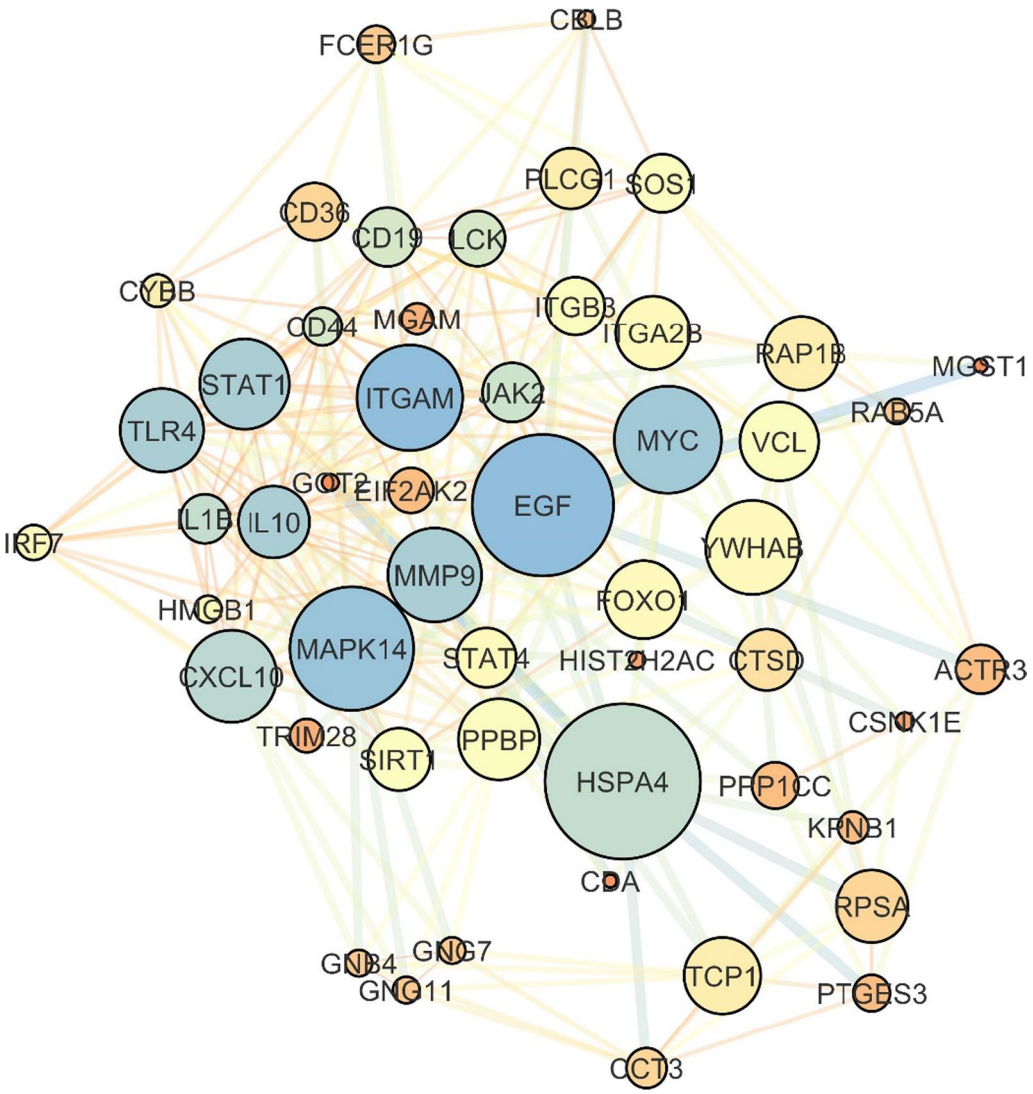

Fig. 6 Topology of the backbone network that consists of 52 nodes with a high BC value and 296 edges. The size of the nodes corresponds to their BC values

cellular responses [40]. STAT1 is one of the principal genes linked with the IFN signalling pathway, according to Yi et al. [41], and it plays a vital role in the immunological defence against TB infection. Other key genes in the network include N-formyl peptide receptor 2 (FPR2), a low-affinity receptor for N-formyl-methionyl peptides, and potent neutrophil chemotactic agents that activate neutrophils when bound [42]. C3AR1 is a receptor that promotes chemotaxis, granule enzyme release and superoxide anion generation in response to the chemotactic and inflammatory peptide anaphylatoxin C3a [43]. Kumar et al. [44] show that chemokines are disease severity markers in pulmonary TB, suggesting greater bacterial burden and delayed culture conversion. Ubiquitination serves as a pathogen defence mechanism for the host. Ubiquitin ligase 3 was thought to have a role in ubiquitinating M. tuberculosis in order to direct autophagic mycobacteria to elimination [45]. SOCS3 (suppressor of cytokine signalling 3) is a protein that plays a role in the negative regulation of cytokines that signal via the JAK/STAT pathway. It inhibits cytokine signal transmission by binding to tyrosine kinase receptors such as IL-6ST/gp130, LIF, erythropoietin, insulin, IL-12, GCSF and leptin receptors. JAK2 kinase activity is inhibited, and IL-6 signalling is regulated when JAK2 is bound. Erythropoiesis in the foetal liver is inhibited and regulates T-helper type 2 cell-mediated allergic reactions [46]. Interleukin-10

Table 5 Genes belonging to each cluster with respective MCODE scores and clustering coefficients: clusters were ranked based on the MCODE scores which implied that C1 had the highest total density around each node in the cluster

Clusters	MCODE score	Clustering coefficient	Nodes	Edges	Gene
C1	36	1	36	630	CUL5, SOCS1, KLHL2, RNF138, TRIM21, CBLB, UBE2Q2, PJA2, UBE2L6, PJA1, HERC5, ASB8, RBCK1, UBE2G2, WWP1, CDC26, RNF213, TRIP12, FBXL3, FBXO9, MEX3C, ANAPC1, UBE4A, FBXO31, KLHL3, ANAPC13, FBXL5, UBE2E1, KCTD6, FBXO7, SOCS3, UBE2K, FBXO6, UBE2O, KBTBD7, UBA6
C2	32	1	32	496	CCR2, PNOC, FPR1, GNG7, C3AR1, FPR2, S1PR5, CXCL10, CXCR6, PTGDR2, GNAI3, GPR18, CCR3, LPAR5, ADCY3, GPR183, HCAR3, HCAR2, CCL4L1, P2RY14, CCR1, P2RY13, PPBP, CCR7, CXCL16, CXCR5, GNB4, CXCR1, HEBP1, GNG11, CXCR3, S1PR1
C3	28.414	0.959	30	412	C18orf32, RPL7, RPS3, RPS23, RPS15A, RPS8, RPL13A, RPL23, RPS28, RPS5, SPCS3, RPN2, EIF3A, RPL13, RPS2, RPL4, RPL22, MRPL15, RPLP0, RPL36, RPL3, RPL21, RPS4X, RPL26, RPL9, RPSA, RPL5, GSPT2, SRP9, SRP72
C4	22.366	0.944	72	794	CEACAM1, LILRB2, P2RX1, PGLYRP1, BST1, CYFIP1, TCN1, CLEC4D, BPI, OLR1, FCER1G, QPCT, ITGAM, MCEMP1, HLA-F, GPR97, CAMP, PLAUR, SLC2A3, CLEC12A, OSCAR, ORM1, CEACAM3, CYBB, GBP2, ATP8B4, CLEC5A, ITGAX, LAMTOR3, ARG1, HLA-G, VCL, HLA-C, KPNB1, CTSD, DEGS1, OLFM4, RSAD2, MX1, SLC11A1, IRF7, ATP6AP2, RAP2C, CD36, OASL, LRG1, NBEAL2, XAF1, MGAM, IFI6, HP, HPSE, IFIT1, IFIT3, IFIT2, STAT2, ELANE, FOLR3, IFITM1, IFI35, OAS3, IRF1, IRF8, LCN2, RNASEL, CEACAM8, ISG15, GPR84, OAS1, DEFA4, IFITM3, TOM1
C5	11.556	0.861	55	312	METTL7A, WSB2, MMP9, CISH, PICALM, PACSIN1, FCGR1A, CDA, CREG1, STAT1, IL-7R, RTP4, DCUN1D1, TRIM25, GBP5, CD44, CTSA, SCARB2, VAMP7, GCA, LDLRAP1, FTH1, REPS2, SYNJ1, GBP1, SDCBP, BIN1, MT2A, FCHO1, GRN, CD3D, EIF2AK2, ARPC5, TRIM22, IFI30, DNM3, GNS, IFIH1, RNASE2, FCGR1B, CD3G, IFI44, NHLRC3, IFI44L, ACTR3, NPC2, RAB5A, HEBP2, AGFG1, MPO, TNFAIP6, GBP4, PLAC8, EGF, COMMD8
C6	10.889	0.905	19	98	DHX15, HNRNPC, CSTF3, CSNK1E, LSM4, SMNDC1, TRA2B, HNRNPL, WBSCR22, PAPOLA, SNRNP27, DHX9, SNRNP40, SNRPN, PPWD1, BMS1, UTP3, HNRNPA1, RIOK3

C1–C6 are that ensure the effectiveness of functional partners towards the core network of T2DM-TB DEGs

(IL-10) is a key immune regulatory cytokine that acts on numerous immune system cells and has potent anti-inflammatory properties, reducing inflammation-induced tissue damage. IL-10 binds to its heterotetrameric receptor, including IL-10RA and IL-10RB, causing STAT3 to be phosphorylated by JAK1 and STAT2 [47]. E3 ubiquitin-protein ligase TRIM21 forms a complex in cooperation with the E2 UBE2D2 that is used not only for USP4 and IKBKB but also for its self-ubiquitination. In response to IFN-γ, TRIM21 regulates innate immunity and the inflammatory response [48]. CXCL-10 is a proinflammatory cytokine involved in several activities, including chemotaxis, differentiation, activation of peripheral immune cells, cell growth control, apoptosis and angiostatic effects modulation [49, 50]. The CXCL-10/CXCR3 axis is also crucial in neurons responding to brain damage for activating microglia, the central nervous system's resident macrophage population, and directing them to the lesion site. This process of recruitment is critical for neuronal remodelling.

The 1058 genes in the network were clustered using Cytoscape MCODE. MCODE discovered 19 highly related gene clusters based on the number of direct connections and gene connectivity in the network. Based on a minimal MCODE score, we selected six clusters for functional enrichment and pathway analysis. Cluster C1 had the highest MCODE score of 36, with 36 genes and 630 functional connections, whereas cluster C4 had the densest interactions, with 72 genes and 794 functional interactions (Fig. 7). The clusters C2, C3, C5 and C6 showed 32, 30, 55 and 19 genes, respectively, with 496, 412, 312 and 98 edges. The purpose of the functional enrichment study was to extract as much information as possible on the functional association of these DEGs at the molecular level. The DAVID database was used to predict GO keywords such as biological processes (BP), molecular functions (MF) and cellular components (CC), as well as KEGG pathways linked to different immune response mechanisms were elevated in T2DM people during TB comorbidity. BPs are GO words that define the capability of numerous cellular interactions and are crucial for the cell's survival. At the molecular level, MFs are the essential actions of the gene product, such as binding or catalysis [10]. We discovered a highly linked area using cluster analysis that comprised the seed proteins identified as essential genes in the network, such as ITGAM and STAT1 in clusters 4 and 5. Although the immune responses to TB and T2DM have been widely studied, the immune responses to TB in T2DM patients are yet unknown. During TB infection, the host's innate immunity stimulates macrophages, resulting in phagocytosis and the activation of cytokines. The ITGAM and

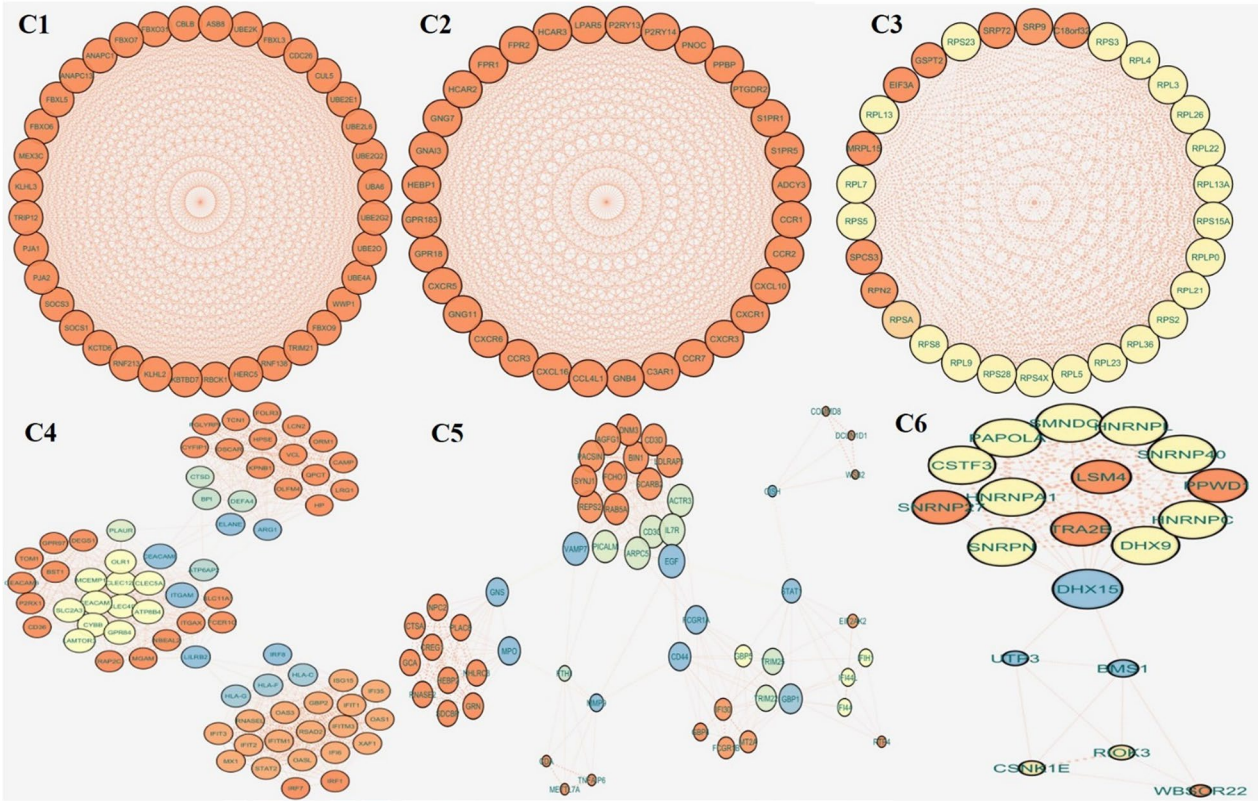

Fig. 7 Gene clusters as viewed in Cytoscape. For each distribution of genes belonging to each cluster accordingly as C1–C6. The different colours in each network represent the BC value (low values to bright colour and high values to dark colour)

CD11B are considered as macrophage marker [51]. In non-diabetic conditions, cytokines get activated, resulting in the production of reactive oxygen species (ROS), which kills TB. The immunosuppression results in a loss of self-defence, and anti-TB drug interactions with anti-diabetic treatments result in a lower therapeutic index [52]. Because of the immune-compromised state of T2DM patients, changes in cytokine release contribute to TB survival due to reduced ROS formation (Fig. 9).

The enriched BP entries are related to cluster C4, where ITGAM includes BP like cell adhesion, integrin-mediated signalling pathway, ectodermal cell differentiation, extracellular matrix organization, toll-like receptor 4 signalling pathway and leukocyte migration. The enriched cellular components include extracellular space, plasma membrane, integrin complex, cell surface, an integral component of membrane and extracellular exosome. Glycoprotein binding, protein binding, metal ion binding and protein heterodimerization are molecular activities. Rap1 signalling pathway, phagosome, cell adhesion molecules, haematopoietic cell lineage, leukocyte transendothelial migration, control of actin cytoskeleton and TB are among the KEGG pathways enriched. The nonsynonymous ITGAM mutations rs1143679 and rs1143678/rs113683 lead to altered Mac-1 function on neutrophils. In M. tuberculosis-infected dendritic cells, the distribution of integrin beta-2 is also significantly changed. Critical pathways include toll-like receptor, RAP1 signalling route, NOD-like receptor signalling pathway, MAPK signalling network, TNF signalling, chemokine signalling pathway, PI3K-Akt signalling pathway and apoptosis, and others are abundant in TB disease settings (including latent infection). Receptors recognize the pathogen on the surface of immune cells, which is toll-like receptors [53].

Similarly, negative regulation of transcription from the RNA polymerase II promoter, negative regulation of endothelial cell proliferation, positive regulation of mesenchymal cell proliferation, negative regulation of mesenchymal-to-epithelial transition involved in metanephros morphogenesis, transcription, DNA-templated and apoptotic cell death are among the enriched biological processes of cluster C5 related to STAT1. Nuclear chromatin, nucleus, nucleoplasm, nucleolus, cytoplasm, cytosol, cell–cell adherens junction, axon, dendrite and

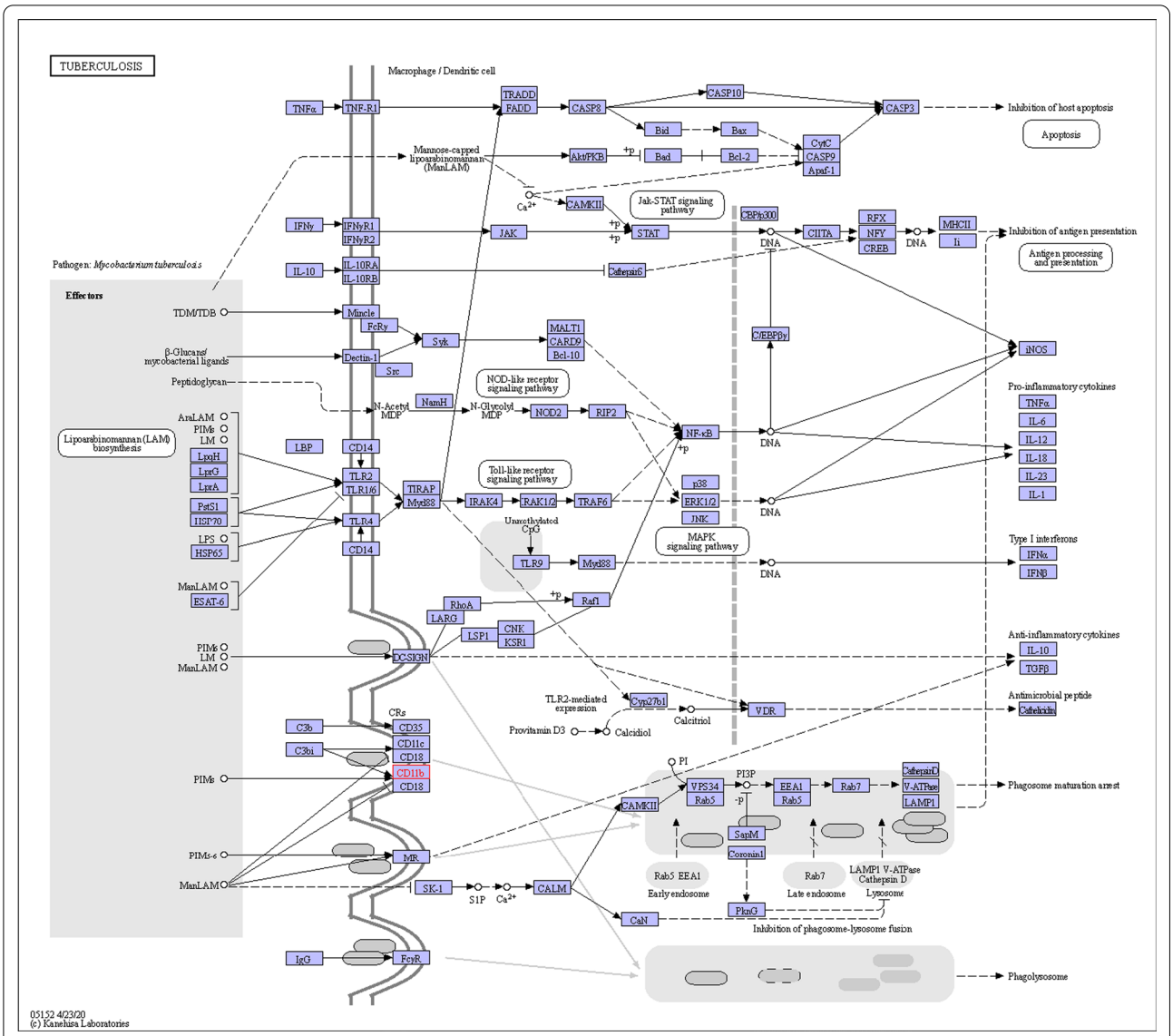

Fig. 8 KEGG pathways enriched by genes in cluster 4 and where the hub gene with large degree value, integrin subunit alpha M (ITGAM/CD11b) identified as a critical in the tuberculosis pathway (highlighted in red colour)

perinuclear area of cytoplasm are among the enriched cellular components. RNA polymerase II core promoter proximal region sequence-specific DNA binding, RNA polymerase II core promoter sequence-specific DNA binding, transcription factor activity and RNA polymerase II core promoter sequence-specific DNA binding were also enriched.

Chemokine signalling system, osteoclast differentiation, toll-like receptor signalling pathway, Jak-STAT signalling pathway, prolactin signalling pathway, thyroid hormone signalling pathway and TB are among the most enriched KEGG pathways associated with STAT1. Cluster 5 contains a large number of genes that are involved in the IFN signalling pathway and the defensive response. The immune system's IFN signalling pathway and cytokine signalling have also been strongly connected to STAT1 [41].

Cytokines regulate cell differentiation, proliferation and immunity by binding to cell membrane surface receptors and activating intracellular signalling pathways such as the JAK-STAT signalling network and the p53 signalling system. In vivo, signalling between cytokines and specific

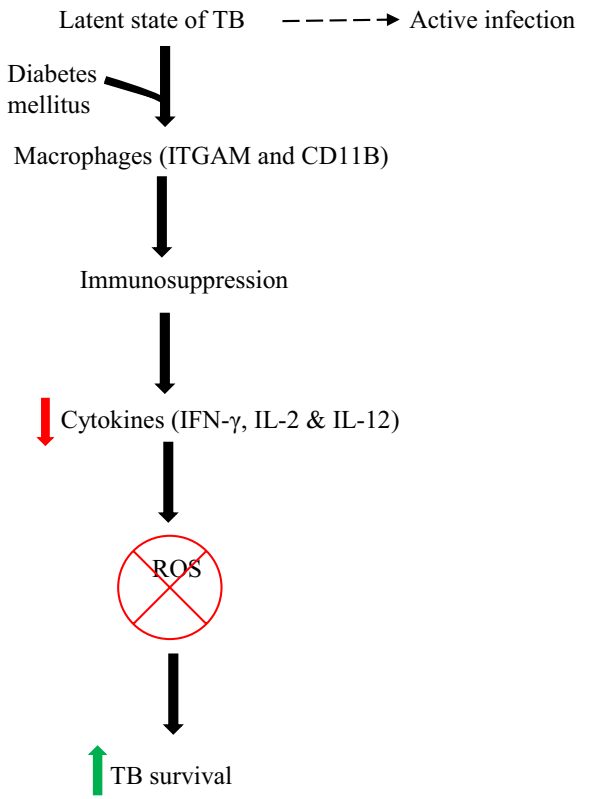

Fig. 9 Schematic representation that promotes TB infection during diabetes mellitus. TB is a latent illness that can progress to active TB (dotted arrow); in diabetic conditions, during immunosuppression state, the patient is more likely to acquire active tuberculosis (bold arrows)

cell subsets is critical for maintaining homeostasis. IFN signalling is important for the host immune defence response in the pathogenesis of tuberculosis, and IFN can boost the activity of native immune cells such as natural killer cells, cytotoxic lymphocyte cells, and macrophages. IFN, generated in large quantities by Th1 cells, can trigger the huge production of MIP-1 and RANTES, allowing chemotactic monocytes to phagocytose and remove MTB [54, 55].

Phosphorylation of STAT1 can boost transcription activation by downstream apoptotic factors in the early stages of TB infection. Non-phosphorylated STAT1 proteins have been shown to increase the expression of the anti-apoptotic protein McL-1, inhibit the phosphorylated kinase JAK1 of STAT1, inhibit CD95/CD95l-mediated apoptosis in macrophages and destroy the stability of the pro-apoptotic protein McL-1, according to research [56]. STAT1 binds to phosphotyrosine-containing peptide sequences and forms homologous dimers that activate the IFN-induced signalling cascade when phosphorylated. An IFN-γ activation region linked to the promoter activates IFN-induced early gene expression [57]. As a result, we believe STAT1 plays a critical role in the immune system's fight against TB. Our findings also demonstrated that cytokine and IFN signalling are the most important host defence responses to TB infection in diabetes.

5 Conclusions

Our current study found that the NF-kappa B signalling pathway, toll-like receptor signalling pathway, Jak-STAT signalling pathway and cytokine signalling in the immune system, especially the IFN signalling pathway, are extremely important for TB disease in T2DM conditions. As a result, molecules with significant relevance to these pathways ITGAM and STAT1 were identified as potential biomolecules in the host defence response to TB infection in diabetic conditions. Overall, our findings support the development of host-directed therapies in T2DM-TB that target cytokine/chemokine signalling pathways and diabetic complication pathways to reduce the morbidity and mortality associated with the common dual burden of communicable and non-communicable diseases.

Abbreviations

T2DM: Type 2 diabetes mellitus; TB: Tuberculosis; DEGs: Differentially expressed genes; Jak-STAT: Janus kinases—signal transducer and activator of transcription proteins; ITGAM: Integrin subunit alpha M; STAT1: Signal transducer and activator of transcription 1; TIMP4: Tissue inhibitor of matrix metalloproteinase 4; VEGF: Vascular endothelial growth factor; HbA1c: Glycated haemoglobin; GO: Gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI: Protein–protein interaction; DAVID: Database for Annotation, Visualization, and Integrated Discovery; STRING: Search tool for the retrieval of interacting genes/proteins database; BC: Betweenness centrality; CC: Closeness centrality; MCODE: Molecular cluster detection.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s43088-022-00257-5.

Additional file 1. Table S1. Data set of the differentially expression genes. **Table S2**. Network topology analysis. **Table S3**. List of large degree nodes. **Table S4**. The list of high BC nodes. **Table S5**. Network topology analysis of giant network. **Table S6**. Network topology analysis of backbone network.

Acknowledgements

The authors thank Nitte (Deemed to be University), Mangalore, India, for providing all the facilities to complete this work.

Author contributions

TSG and PG designed and performed the experiments; all authors analysed the data; PG prepared the figures; PKS and SKN supervised the work. TSG and PG wrote the first draft of the manuscript. All authors discussed results of the experiments, edited and approved the final version of the manuscript. All authors read and approved the final manuscript.

Funding
Funding not received for the study.

Availability of data and materials
All the data we generated in this paper are available in the body of the manuscript as supporting figures and tables. We do not have any ethical or legal consideration for not to make our data publicly available.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that there are no conflicts of interest.

Author details
[1]Central Research Laboratory, KS Hegde Medical Academy, Nitte (Deemed to Be University), Mangalore, Karnataka 575018, India. [2]Centre for Bioinformatics, Nitte (Deemed to Be University), Deralakatte, Mangalore, Karnataka 575018, India. [3]Department of Biochemistry, KS Hegde Medical Academy, Nitte (Deemed to Be University), Deralakatte, Mangalore 575018, Karnataka, India.

Received: 14 February 2022 Accepted: 19 May 2022
Published online: 04 June 2022

References

1. Magee M, Ali M, Prabhakaran D et al (2017) Integrated public health and health service delivery for noncommunicable diseases and comorbid infectious diseases and mental health. In: Prabhakaran D, Anand S, Gaziano TA et al (eds) Cardiovascular, respiratory, and related disorders, 3rd edn. The International Bank for Reconstruction and Development/The World Bank, Washington, DC
2. Pande T, Huddart S, Xavier W et al (2018) Prevalence of diabetes mellitus amongst hospitalized tuberculosis patients at an Indian tertiary care center: a descriptive analysis. PLoS ONE 13(7):e0200838. https://doi.org/10.1371/journal.pone.0200838
3. Restrepo BI (2016) Diabetes and tuberculosis. Microbiol Spectrum. https://doi.org/10.1128/microbiolspec.TNMI7-0023-2016
4. Whiting DR, Guariguata L, Weil C et al (2011) IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. Diabetes Res Clin Pract 94(3):311–321. https://doi.org/10.1016/j.diabres.2011.10.029
5. Al-Rifai RH, Pearson F, Critchley JA et al (2017) Association between diabetes mellitus and active tuberculosis: A systematic review and meta-analysis. PLoS ONE 12(11):e0187967. https://doi.org/10.1371/journal.pone.0187967
6. Martinez N, Kornfeld H (2014) Diabetes and immunity to tuberculosis. Eur J Immunol 44(3):617–626. https://doi.org/10.1002/eji.201344301
7. Berbudi A, Rahmadika N, Tjahjadi AI et al (2020) Type 2 Diabetes and its impact on the immune system. Curr Diabetes Rev 16(5):442–449. https://doi.org/10.2174/1573399815666191024085838
8. Kumar NP, Moideen K, Sivakumar S et al (2017) Tuberculosis-diabetes co-morbidity is characterized by heightened systemic levels of circulating angiogenic factors. J Infect 74(1):10–21. https://doi.org/10.1016/j.jinf.2016.08.021
9. Prada-Medina CA, Fukutani KF, Pavan Kumar N et al (2017) Systems immunology of diabetes-tuberculosis comorbidity reveals signatures of disease complications. Sci Rep 7(1):1999. https://doi.org/10.1038/s41598-017-01767-4
10. The Gene Ontology Consortium (2019) The gene ontology resource: 20 years and still going strong. Nucleic Acids Res 47(D1):D330–D338. https://doi.org/10.1093/nar/gky1055
11. Kanehisa M, Furumichi M, Tanabe M et al (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45(D1):D353–D361. https://doi.org/10.1093/nar/gkw1092
12. Jiao X, Sherman BT, da Huang W et al (2012) DAVID-WS: a stateful web service to facilitate gene/protein list analysis. Bioinformatics 28(13):1805–1806. https://doi.org/10.1093/bioinformatics/bts251
13. Szklarczyk D, Morris JH, Cook H et al (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res 45(D1):D362–D368. https://doi.org/10.1093/nar/gkw937
14. Gollapalli P, Manjunatha H, Shetty P (2021) Network topology analysis of essential genes interactome of Helicobacter pylori to explore novel therapeutic targets. Microbial Pathogenesis 158:105059. https://doi.org/10.1016/j.micpath.2021.105059
15. Miryala SK, Anbarasu A, Ramaiah S (2019) Systems biology studies in Pseudomonas aeruginosa PA01 to understand their role in biofilm formation and multidrug efflux pumps. Microb Pathog 136:103668. https://doi.org/10.1016/j.micpath.2019.103668
16. Miryala SK, Ramaiah S (2019) Exploring the multi-drug resistance in Escherichia coli O157:H7 by gene interaction network: a systems biology approach. Genomics 111(4):958–965. https://doi.org/10.1016/j.ygeno.2018.06.002
17 Su G, Morris JH, Demchak B et al (2014) Biological network exploration with Cytoscape 3. Curr Protoc Bioinformatics. 47:8–13. https://doi.org/10.1002/0471250953.bi0813s47
18. Miryala SK, Anbarasu A, Ramaiah S (2020) Role of SHV-11, a class a β-lactamase, gene in multidrug resistance among klebsiella pneumoniae strains and understanding its mechanism by gene network analysis. Microb Drug Resist 26(8):900–908. https://doi.org/10.1089/mdr.2019.0430
19. Zhang P, Tao L, Zeng X et al (2017) A protein network descriptor server and its use in studying protein, disease, metabolic and drug targeted networks. Brief Bioinform 18(6):1057–1070. https://doi.org/10.1093/bib/bbw071
20. Mitra K, Carvunis AR, Ramesh SK et al (2013) Integrative approaches for finding modular structure in biological networks. Nat Rev Genet 14(10):719–732. https://doi.org/10.1038/nrg3552
21. Ma CY, Chen YP, Berger B et al (2017) Identification of protein complexes by integrating multiple alignment of protein interaction networks. Bioinformatics 33(11):1681–1688. https://doi.org/10.1093/bioinformatics/btx043
22. Hossain SM, Mahboob Z, Chowdhury R et al (2016) Protein complex detection in ppi network by identifying mutually exclusive protein-protein interactions. Procedia Comput Sci 93:1054–1060. https://doi.org/10.1016/j.procs.2016.07.309
23. Broido AD, Clauset A (2019) Scale-free networks are rare. Nat Commun 10(1):1017. https://doi.org/10.1038/s41467-019-08746-5
24. Ashtiani M, Salehzadeh-Yazdi A, Razaghi-Moghadam Z et al (2018) A systematic survey of centrality measures for protein-protein interaction networks. BMC Syst Biol 12(1):80. https://doi.org/10.1186/s12918-018-0598-2
25. Rao VS, Srinivas K, Sujini GN et al (2014) Protein-protein interaction detection: methods and analysis. Int J Proteomics 2014:147648. https://doi.org/10.1155/2014/147648
26. Liu W, Ma L, Chen L (2019) Identification of essential proteins by using complexes and biological information on dynamic PPI network. Seventh Int Conf Adv Cloud Big Data (CBD). https://doi.org/10.1109/CBD.2019.00032
27. Ryndak MB, Laal S (2019) Mycobacterium tuberculosis primary infection and dissemination: a critical role for alveolar epithelial cells. Front Cell Infect Microbiol 9:299. https://doi.org/10.3389/fcimb.2019.00299
28. O'Shea MK, Tanner R, Müller J et al (2018) Immunological correlates of mycobacterial growth inhibition describe a spectrum of tuberculosis infection. Sci Rep 8(1):14480. https://doi.org/10.1038/s41598-018-32755-x
29. Sekaran TSG, Kedilaya VR, Kumari SN et al (2021) Exploring the differentially expressed genes in human lymphocytes upon response to ionizing radiation: a network biology approach. Radiat Oncol J 39(1):48–60. https://doi.org/10.3857/roj.2021.00045
30. Farhadian M, Rafat SA, Panahi B et al (2021) Weighted gene co-expression network analysis identifies modules and functionally enriched pathways in the lactation process. Sci Rep 11(1):2367. https://doi.org/10.1038/s41598-021-81888-z

31 Pisu D, Huang L, Narang V et al (2021) Single cell analysis of M tuberculosis phenotype and macrophage lineages in the infected lung. J Exp Med. 218(9):e20210615. https://doi.org/10.1084/jem.20210615
32. Hickman S, Izzy S, Sen P et al (2018) Microglia in neurodegeneration. Nat Neurosci 21(10):1359–1369. https://doi.org/10.1038/s41593-018-0242-x
33. Bose TO, Pham QM, Jellison ER et al (2013) CD11a regulates effector CD8 T cell differentiation and central memory development in response to infection with Listeria monocytogenes. Infect Immun 81(4):1140–1151. https://doi.org/10.1128/IAI.00749-12
34. Wang Y, Cheng X, Samma MK et al (2018) Differential cellular responses by oncogenic levels of c-Myc expression in long-term confluent retinal pigment epithelial cells. Mol Cell Biochem 443(1–2):193–204. https://doi.org/10.1007/s11010-017-3224-3225
35. Yim HC, Li JC, Pong JC et al (2011) A role for c-Myc in regulating anti-mycobacterial responses. Proc Natl Acad Sci U S A 108(43):17749–17754. https://doi.org/10.1073/pnas.1104892108
36. Stine ZE, Walton ZE, Altman BJ et al (2015) MYC, Metabolism, and Cancer. Cancer Discov 5(10):1024–1039. https://doi.org/10.1158/2159-8290.CD-15-0507
37. Roy S, Schmeier S, Kaczkowski B et al (2018) Transcriptional landscape of Mycobacterium tuberculosis infection in macrophages. Sci Rep 8(1):6758. https://doi.org/10.1038/s41598-018-24509-6
38. Chen K, Liu J, Liu S et al (2017) Methyltransferase SETD2-mediated methylation of STAT1 Is critical for interferon antiviral activity. Cell 170(3):492-506.e14. https://doi.org/10.1016/j.cell.2017.06.042
39. Zhang Y, Mao D, Roswit WT et al (2015) PARP9-DTX3L ubiquitin ligase targets host histone H2BJ and viral 3C protease to enhance interferon signaling and control viral infection. Nat Immunol 16(12):1215–1227. https://doi.org/10.1038/ni.3279
40. Layoun A, Goldberg AA, Baig A et al (2019) Regulation of protein kinase Cδ Nuclear import and apoptosis by mechanistic target of rapamycin complex-1. Sci Rep 9(1):17620. https://doi.org/10.1038/s41598-019-53909-5
41. Yi XH, Zhang B, Fu YR et al (2020) STAT1 and its related molecules as potential biomarkers in Mycobacterium tuberculosis infection. J Cell Mol Med 24(5):2866–2878. https://doi.org/10.1111/jcmm.14856
42. López-Hernández Y, Lara-Ramírez EE, Salgado-Bustamante M et al (2019) Glycerophospholipid metabolism alterations in patients with Type 2 diabetes mellitus and tuberculosis comorbidity. Arch Med Res 50(2):71–78. https://doi.org/10.1016/j.arcmed.2019.05.006
43. Maertzdorf J, Repsilber D, Parida SK et al (2011) Human gene expression profiles of susceptibility and resistance in tuberculosis. Genes Immun 12(1):15–22. https://doi.org/10.1038/gene.2010.51
44. Kumar NP, Moideen K, Nancy A et al (2019) Plasma chemokines are biomarkers of disease severity, higher bacterial burden and delayed sputum culture conversion in pulmonary tuberculosis. Sci Rep 9(1):18217. https://doi.org/10.1038/s41598-019-54803-w
45. Subrahmanian M, Marimuthu J, Sairam T et al (2020) In vitro ubiquitination of Mycobacterium tuberculosis by E3 ubiquitin ligase, MKRN1. Biotechnol Lett 42(8):1527–1534. https://doi.org/10.1007/s10529-020-02873-6
46. Carow B, Rottenberg ME (2014) SOCS3, a major regulator of infection and inflammation. Front Immunol 5:58. https://doi.org/10.3389/fimmu.2014.00058
47. Verma R, Balakrishnan L, Sharma K et al (2016) A network map of Interleukin-10 signaling pathway. J Cell Commun Signal 10(1):61–67. https://doi.org/10.1007/s12079-015-0302-x
48. Wang J, Teng JL, Zhao D et al (2016) The ubiquitin ligase TRIM27 functions as a host restriction factor antagonized by Mycobacterium tuberculosis PtpA during mycobacterial infection. Sci Rep 6:34827. https://doi.org/10.1038/srep34827
49. Bhattacharyya C, Majumder PP, Pandit B (2018) CXCL10 is overexpressed in active tuberculosis patients compared to M. tuberculosis-exposed household contacts. Tuberculosis (Edinb) 109:8–16. https://doi.org/10.1016/j.tube.2018.01.005
50. Sidahmed AM, León AJ, Bosinger SE et al (2012) CXCL10 contributes to p38-mediated apoptosis in primary T lymphocytes in vitro. Cytokine 59(2):433–441. https://doi.org/10.1016/j.cyto.2012.05.002
51. King BC, Kulak K, Krus U et al (2019) Complement component C3 Is highly expressed in human pancreatic islets and prevents β cell death via ATG16L1 interaction and autophagy regulation. Cell Metab 29(1):202–210. https://doi.org/10.1016/j.cmet.2018.09.009
52. Pal R, Ansari MA, Hameed S et al (2016) Diabetes mellitus as hub for tuberculosis infection: a snapshot. Int J Chronic Dis 2016:5981574. https://doi.org/10.1155/2016/5981574
53. Alam A, Imam N, Ahmed MM et al (2019) Identification and classification of differentially expressed genes and network meta-analysis reveals potential molecular signatures associated with tuberculosis. Front Genet 10:932. https://doi.org/10.3389/fgene.2019.00932
54 Churchyard G, Kim P, Shah NS et al (2017) What We Know About Tuberculosis Transmission: An Overview. J Infect Dis. 216(Suppl_6):S629–S635. https://doi.org/10.1093/infdis/jix362
55. Wang F, Mao L, Hou H et al (2016) The source of Mycobacterium tuberculosis-specific IFN-γ production in peripheral blood mononuclear cells of TB patients. Int Immunopharmacol 32:39–45. https://doi.org/10.1016/j.intimp.2016.01.012
56. Yao K, Chen Q, Wu Y et al (2017) Unphosphorylated STAT1 represses apoptosis in macrophages during Mycobacteriumtuberculosis infection. J Cell Sci 130(10):1740–1751. https://doi.org/10.1242/jcs.200659
57. O'Connell D, Bouazza B, Kokalari B et al (2015) IFN-γ-induced JAK/STAT, but not NF-κB, signaling pathway is insensitive to glucocorticoid in airway epithelial cells. Am J Physiol Lung Cell Mol Physiol 309(4):L348-359. https://doi.org/10.1152/ajplung.00099.2015

Publisher's Note