Identification of Perturbed Pathways Rendering Susceptibility to Tuberculosis in Type 2 Diabetes Mellitus Patients Using BioNSi Simulation of Integrated Network of Implicated Human Genes

Jyoti Rani
University of Delhi
Anasuya Bhargav
Academy of Scientific and Innovative Research (AcSIR)
Malabika Datta
CSIR – Institute of Genomics and Integrative Biology (CSIR-IGIB)
Urmi Bajpai
University of Delhi
Srinivasan Ramachandran (ramuigib@gmail.com)
CSIR – Institute of Genomics and Integrative Biology (CSIR-IGIB)

Research Article

Keywords: Type 2 Diabetes, Tuberculosis, text mining, differential expression, simulation, BioNSi

Posted Date: September 3rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-863821/v1

Abstract

Adaptive immune response of the Th1 arm is the main defense against tuberculosis (TB). However, in Type 2 Diabetes Mellitus (T2DM) patients, chronic hyperglycemia and inflammation underlie susceptibility to TB and results in poor TB control. The molecular pathways causing susceptibility of diabetics to tuberculosis is not fully understood. Here, an integrative pathway-based approach is used to investigate the perturbed pathways in T2DM patients rendering susceptibility to TB. We obtained 36 genes implicated in the Type 2 diabetes associated tuberculosis (T2DMTB) from literature. Gene expression analysis on T2DM patients' data (GSE28168) showed that *DEFA1* is differentially expressed at $P_{adj} < 0.05$. The genes *CAMP*, *CD14*, *CORO1A*, *LAMP1*, *TLR4*, *IL17F* and *SOCS3* were differentially expressed in T2DM patients at P value < 0.05. 7 microRNAs associated with these T2DMTB genes were obtained from NetworkAnalyst and verified for their literature evidences. The hsa-miR-146a microRNA was differentially expressed at $P_{adj} < 0.05$. The human host TB susceptibility genes *TNFRSF10A*, *MSRA*, *GPR148*, *SLC37A3*, *PXK*, *PROK2*, *REV3L*, *PGM1*, *HIST3H2A*, *PLAC4*, *LETM2*, *EMP2* and were also differentially expressed at $P_{adj} < 0.05$. We included all these genes and added the remaining 28 genes from the T2DMTB set and the rest of differentially expressed genes at $P_{adj} < 0.05$ in STRING and obtained a well-connected network with high confidence score greater than 0.7. From this network we extracted the KEGG pathways at FDR < 0.05 and retained only Diabetes and TB pathways among the disease pathways. The network was simulated with BioNSi using gene expression data from GSE26168. The Necroptosis pathway showed the maximum perturbations in T2DM patients, followed by NOD-like receptor signaling, Toll-like receptor signaling, NF-kappa-B signaling and MAPK signaling. These pathways likely underlie susceptibility to TB in T2DM patients.

Introduction

Patients with Type 2 Diabetes Mellitus (T2DM) are reportedly at risk of acquiring Tuberculosis (TB) [1 2 3]. T2DM develops due to insulin resistance. It is on the rise affecting an increasing number of people every year worldwide [4]. Majority of diabetic patients are of T2DM type [5]. 10% of deaths due to TB cases are linked to diabetes in low and middle-income countries [4]. Drugs used to treat TB are reported to cause diabetic associated conditions. Rifampicin reportedly causes hyperglycaemic conditions and Isoniazid (INH) is prone to cause peripheral neuropathy, which may worsen diabetic neuropathy [3]. Harries et al., 2016 suggested a collaborative care approach for DM and TB patients to tackle this comorbidity [6]. Tuberculosis co-morbidity is associated with poor glycaemic control in diabetic patients [7]. A mice model study points to the impact of chronic hyperglycemia in impairment of initiation of adaptive immunity thereby resulting in higher load of *Mycobacterium tuberculosis* in the lungs [8].

Diabetic patients are 3.1 times more susceptible to TB infection compared to non-diabetic individuals [9] and showed lower cellular immunity [10] relative to non-diabetic control. Because cellular immunity and macrophages offer defense against TB infection, reduced cellular immunity and impairment of Th1 immune response could underlie susceptibility to TB [8 11]. Enhancement of susceptibility to tuberculosis

in diabetic patients is related to hyperglycemia and insulin resistance as well as indirect effects on macrophage and lymphocyte functions [12 13]. Kumar et al., 2013 observed that heightened pro-inflammatory cytokines (IL1B, IL6 and IL18) in diabetic patients result in poor control of tuberculosis infection [12 14]. Later, they reported significant reduced expression of cytotoxic markers perforin, granzyme B and CD107a in $CD8^+$ T cells whereas in NK cells expression of CD107a was decreased in T2DMTB patients [15]. Therefore, altered repertoire of $CD8^+$ T and NK cells expressing cytotoxic molecules evidently underlie T2DMTB comorbidity [15]. Kirtimaan Syal et al., 2015 observed that the Vitamin A receptor (RXR) expression is negatively regulated in T2DM patients and the Tryptophan-aspartate containing coat protein (TACO) gene expression is positively regulated, which could predispose T2DM patients to TB [16]. Hui-Qi Qu et al., 2012 identified *HK2* and *CD28* as potential genes for association of TB in diabetic patients [17].

Gene expression regulation is also achieved post transcriptionally through microRNAs (miRNAs). Lattore et al., selected 4 miRNAs, namely, hsa-miR-150, hsa-miR-21, hsa-miR-29c and hsa-miR-194 dysregulated in TB versus latent tuberculosis infection or TB versus healthy controls and described a whole blood derived miRNA signature enabling diagnosis of TB with over 90% sensitivity and over 87% specificity [18]. These investigations have offered valuable data. The molecular dynamics underlying both T2DM and TB are complex. The studies of both diseases present challenges and a key question arises regarding the susceptibility of human host towards developing comorbidities. In the context of computational disease modelling, it was realized some years ago that studies of groups of patients with common patterns in disease is more tractable than deciphering individual dynamics [19]. It was also realized that computational modelling can benefit from machine learning, agent-based modelling, network modeling and stochastic simulations. Over the years interest has been gaining in computational modeling due to its data integrative power. Computational models offer integrated description of disease thereby offering a platform for learning and generate hypothesis for therapeutics. [20 21]

In the present work we have used a computational approach to address T2DMTB comorbidity by using an integrative method through collection of the genes implicated in T2DMTB comorbidity, and miRNAs implicated in T2DM and in TB from the published literature. This list was pooled with the differentially expressed host genes susceptible to TB, differentially expressed miRNA, and the rest of differentially expressed genes. This data was used to identify the network at high confidence score and the pathways involved. Subsequently, we investigated the probable cause of susceptibility to tuberculosis in patients with Type 2 Diabetes using pathway simulations and available gene expression data from T2DM patients.

Results

We obtained a total of 36 genes associated with T2DMTB comorbidity along with their evidences in the literature (**Supplementary Table S1**). Heme Oxygenase 1 (HMOX1) for example, a mediator of antioxidant and expressed in lung tissue, can distinguish latent from active TB [22] and it is reportedly high in patients

with TB and T2DM. Tumor Necrosis Factor Alpha (TNF) offers protection from mycobacterial infection [23] and it is reported to be lower in patients with T2DM. Interleukin 22 (IL22) is reported to protect against the development of metabolic disorder and it is decreased in diabetics with TB [24]. The increased plasma level of HMOX1 and decreased plasma levels of TNF and of IL22 result in increased susceptibility of T2DMTB comorbidity. [22 24]

Gene enrichment analysis

The GO enrichment data includes 26 unique GOs (**Supplementary Table S2**). We searched the literature using the GO terms. The host immune response to Mtb infection consists of inflammation, cytokine activity and cellular response. The inflammatory cytokines (IL6, IFNG, TNF-α, IL1B) were observed to be positively regulated in mice models of T2DM with Mtb infection. The cytokines IFNγ, TNFα, and IL17A, considered to offer protection against tuberculosis are all positively regulated in diabetic individuals [12 25]. However, it appears that chronic inflammation underlying type 2 diabetes results in poor control of tuberculosis infection. High Glucose concentrations positively regulate TLR2, TLR4 and activate NF-kappa B to induce pro-inflammatory cytokines. Apparently, the production of nitric oxide exhibiting antibacterial activity also is affected in T2DM patients [26]. Overall, it is apparent that even though the T2DM patients are able to mount an immune response against Mtb infection, the infection is not controlled compared to healthy people.

Differential expression

Our goal was to identify whether the T2DMTB genes were differentially expressed between the T2DM patients without reported tuberculosis and normal controls (GSE26168). A total of 705 genes were differentially expressed in the blood sample of diabetic patients at Padj < 0.05 (**Supplementary Table S3**). Among 8 genes associated with T2DMTB, only *DEFA1* gene was differentially expressed at Padj < 0.05. *CAMP, CD14, CORO1A, LAMP1, TLR4, IL17F*, and *SOCS3* were differentially expressed at P Value < = 0.05 **(Supplementary Table S4)**. The expression changes of T2DMTB genes in diabetic patients (literature data) and T2DM patients (GSE26168) are shown in Table 1. Most T2DMTB genes are not differentially expressed in T2DM patients. A few genes are changing in expression in the same direction in T2DM patients as in T2DMTB patients, namely, *DEFA1*, *SOCS3* and *CORO1A*. The genes whose expressions are changing in opposite directions are *CAMP*, *IL17F*, *LAMP1*, *TLR4* and *CD14*. The remaining genes are not differentially expressed in T2DM patients without tuberculosis. *CORO1A* codes for the protein Coronin 1A, which is retained by mycobacterial phagosomes thereby preventing fusion with lysosomes [27]. This mechanism prevents intracellular killing of Mtb [28]. The expression of this gene is positively regulated in T2DM patients without tuberculosis in parallel with literature reports. *DEFA1* coding for Defensin Alpha 1 is important for host defense against Mtb [29]. Its expression is positively regulated in T2DM patients and also parallels the evidence in the literature report. On the other hand, *CAMP* is also positively regulated contrary to literature report. Same is the case with *LAMP1* and *IL17F*. *LAMP1* coding for Lysosomal Associated Membrane Protein 1 is selectively excluded in the process of inhibiting fusion of Mtb

phagosomes with lysosomes [30]. *IL17F* coding for Interleukin 17F is part of Th17 immune response. It is reported that Th17 recall responses can be targeted to improve vaccine design against TB [31]. In the case of *SOCS3*, macrophage-specific SOCS3 appears necessary to ensure a balanced inflammatory immune response during experimental TB [32] thereby attributing a positive role in protection against Mtb. Taken together even though the gene expression environment in T2DM patients is favorable for colonization of Mtb, host immune response molecules like Defensin and CAMP are positively regulated in T2DM patients without tuberculosis.

Human host TB Susceptible genes

Next, we examined whether in the T2DM patients the host genes rendering susceptibility to TB [33] were differentially expressed. Out of 275 genes, 12 genes (*TNFRSF10A, MSRA, GPR148, SLC37A3, PXK, PROK2, REV3L, PGM1, HIST3H2A, PLAC4, LETM2*, and *EMP2*) were differentially expressed at *Padj* < = 0.05 (Table 2). Although the magnitude of differential expression is low (FC in the range 0.72–1.49), the changes are highly statistically significant. 9 genes were positively regulated and 3 genes *GPR148*, *PLAC4* and *EMP2* were negatively regulated in T2DM patients compared to normal controls (**Supplementary Figure S1**). The data suggest that a small minority (4.36%) of host TB susceptible genes were differentially expressed in T2DM patients of which majority (75%) are positively regulated. It is apparent that these alterations in host TB susceptible genes expression are favorable for colonization of Mtb.

Network Analysis

i) Collection of deregulated miRNAs in T2DM and TB patients

Our premise is that deregulated miRNAs could affect the regulation of target genes, which may lead to T2DMTB co-morbidity. We obtained a total 7 miRNAs (hsa-miR-146a-5p, hsa-miR-155-5p, hsa-miR-125b-5p, hsa-miR-21-5p, hsa-miR-27a-3p, hsa-miR-29b-3p and hsa-miR-223-3p) from the T2DMTB genes-miRNAs network using NetworkAnalyst [34] (**Supplementary Figure S2**). These miRNAs were further examined for evidence of dysregulation in T2DM and in TB conditions in the literature. The results are shown in Supplementary Table S5. It is evident that these miRNAs have their role in the regulation of the immune system.

ii) Differential Expression

Microarray expression analysis revealed the differential expression of above-mentioned miRNAs in T2D Patients. With the criteria of selection on the basis of *Padj* < = 0.05, only hsa-miR-146a was differentially expressed (Fig. 2). This microarray analysis corroborated a previous study of down-regulation of miR-146a in diabetic condition and in children with TB.

iii) Gene regulators of and genes regulated by miR-146a

With the aim to elucidate the mechanism, we collected the gene regulators of and genes regulated by hsa-miR-146a using text mining with the help of pubmed.mineR. The data is displayed in (**Supplementary Table S6**). *SOD2* among the negative regulators of hsa-miR-146a was differentially expressed at Padj < 0.05. *STAT3 a*nd *TGFB1* among the positive regulators were differentially expressed in T2DM patients at P < 0.05. Among the genes negatively regulated by hsa-miR-146, *MMP9* and *SOD2* were differentially expressed at Padj < 0.05 whereas *TLR4*, *CD86*, *FBXL10*, were differentially expressed at P < 0.05. *PAK1*, *STAT1* among the genes positively regulated by hsa-miR-146a were differentially expressed at P < 0.05 in T2DM patients.

v) Biological Network Simulation:

Resulted pathways (**Supplementary Table S7**) further merged and simulated using Biological Network Simulator (BioNSi) [35], a cytoscape plug-in. BioNSi provides modeling of biological networks and discrete-time simulation. The initial states of the genes were assigned using expression values. BioNSi, set the initial state of the genes between 0–9, where 0 is minimum and 9 is maximum state based on the expression values. The simulation was performed in normals and diabetic conditions. The aim to merge and simulate the enriched pathways is to understand the role of miR-146a in the T2DMTB co-morbidity.

We attempted to elucidate the root pathway for susceptibility of tuberculosis in diabetic patients through BioNSi simulations. From the above integrated analysis, we note that the down-regulation of miRNA-146a in diabetic patients could underlie the mechanism conferring TB susceptibility in T2DM patients. As genes and their regulators are networked in cellular systems, we carried out two simulations, one each with the corresponding expression values from non-diabetics and from diabetic patients. Typically, simulations reach a steady state depending on the network connections and the expression values of genes in the network. The exhibited differences in the pattern of the simulation contours could identify the pathways most disturbed due to the changes in the expression values of the genes in the disease conditions. The extent of the disturbance can be quantified through Euclidean distance and we considered a minimum Euclidean distance of 2.

We obtained 26 genes from non-diabetics with a minimum Euclidean distance of 2. They also reached a steady state either prior to or coincident with diabetics (Table 3a). Under the same criteria, in diabetics, 14 genes had a minimum Euclidean distance of 2 and reached steady state either prior to or coincident with non-diabetics (Table 3b). Genes other than cut-off are given in **Supplementary Table S8a** and **S8b.** It is evident that the most altered pathway is the Necroptosis pathway (Fig. 3) followed by NOD-like receptor signaling pathway, Toll-like receptor signaling pathway, NF-kappa-B signaling pathway and MAPK signaling pathway (**Supplementary Figure S3**) As several of these pathways are associated with inflammation, these results uphold the current understanding that chronic inflammation in Type 2 Diabetes Mellitus patients likely underlies susceptibility to Tuberculosis. Recently, Pajuelo et al., 2018 [36] have reported that macrophage necroptosis is exploited by *M. tuberculosis*, which includes NAD +

depletion, depolarization of mitochondria and impaired ATP synthesis. The results obtained here suggest that the necroptic pathway is already active in T2DM patients and therefore once they are infected the active necroptic pathway could enable the dissemination of *M. tuberculosis*. The NOD-like receptors and Toll-like receptors are pattern recognition receptors that recognize pathogens and serve in innate immunity [37]. The interferon regulatory factor-5 (IRF) transcription factor plays important role in innate immune responses via the TLR4-MyD88 pathway [38]. The expression of IRF5 is constitutive in macrophages, B cells and dendritic cells. The alterations observed in IRF5 pathway through TLR4 signaling suggest its contribution in the pathogenesis of TB in the diabetic condition.

Discussion:

Investigations of T2DMTB patients' have revealed the role of chronic inflammation in poor TB control [39]. In this work we sought to investigate the perturbed pathways in T2DM patients with potential to render them susceptible to TB. To this end, we carried out integrative analysis with multiple gene sets including 36 T2DMTB genes, 12 differentially expressed host genes in T2DM patients susceptible to TB, 705 differentially expressed genes in T2DM patients, and one differentially expressed microRNA associated with T2DM and with TB.

Compared to the gene expression changes in T2DMTB patients, only 8 out of 36 genes showed altered gene expression in T2DM patients. One of these genes, *DEFA1* expression is significantly altered whereas the remaining 7 genes expression alterations were statistically significant only at $P < 0.05$. These data clearly show that the expression changes of a lot more genes were observed when T2DM patients have TB. We closely examined for identifiable patterns informative of warning signal for *M. tuberculosis* infection in T2DM patients. Genes with same direction of expression changes compared with T2DM patients were *DEFA1, SOCS3*, and *CORO1A*. The remaining genes *CAMP*, *IL17F*, *LAMP1*, *TLR4* and *CD14* exhibited change in expression in opposite direction. It is noteworthy that the defensin gene *DEFA1* and the cathelicidin antimicrobial peptide gene *CAMP* are positively regulated in T2DM patients signifying that the defense arm is active and therefore could offer protection against microbial infections. These genes are reported as vital in defense against tuberculosis [28 29 30 31]. Along same lines, the positive regulation of *SOCS3* is apparently helpful in modulating the inflammation towards fighting TB [32]. On the other hand, the up-regulation of *CORO1A* gene expression in T2DM patients is indicative of setting a favorable environment for colonization of *M. tuberculosis*. Given this mixed signature of changes in gene expression it is hard to predict susceptibility to TB with high statistical confidence. We can only educate T2DM patients about their potential susceptibility to *M. tuberculosis* infection. We therefore attempted to use pathways approach in order to derive potential signals of susceptibility to *M. tuberculosis* infection. To this end, we used the BioNSi simulation software with gene expression data input from T2DM patients and normal controls in order to shortlist the most perturbed pathways.

The BioNSi software uses a state updation function with starting values scaled in 0–9 range from the gene expression data. Simulations were carried out for T2DM patients and normals separately. The end

points correspond to reaching the steady state. The differences between the gene expression states of T2DM patients and of normal controls were quantified by Euclidean distance. We considered a minimum of Euclidean distance 2 for electing the perturbed genes.

We observed 26 genes from normals and 14 genes from T2DM patients meeting our criteria. After mapping these genes to pathways and ranking these pathways in descending order of scores, the topmost pathways were Necroptosis pathway, NOD-like receptor signaling, Toll-like receptor signaling, NF-KB signaling and MAPK signaling pathway. Recently, Pajuelo et al. reported that Neroptosis pathway enable dissemination of *M. tuberculosis* [40]. Therefore, the positive regulation of the Necroptosis pathway offers a favorable environment for susceptibility to TB. Our analysis is informative regarding monitoring the perturbation of necroptic pathway in gene expression data from blood of T2DM patients. The necrosome dependent necrotic cell death is defined as necroptosis and also reported as pathogenic for pulmonary diseases [41]. Necrosome consists of receptor interacting protein kinase 1 (RIPK1), receptor interacting protein kinase 2 (RIPK2), receptor interacting protein kinase 3 (RIPK3) and mixed lineage kinase domain-like (MLKL) proteins. It is reported that alterations in the Necroptosis pathway may lead to variation in host immune responses against tuberculosis and disease severity [42]. Further, the alterations in mitochondrial permeability likely lead to primary necroptosis [43]. The Necroptosis of macrophages allow bacterial pathogens to evade the immune response [44]. Therefore, taken together the activation of Necroptosis pathway in T2DM patients could lead to easy survival of the bacteria thereby rendering them more prone to active tuberculosis. The hsa-miR-146a is down-regulated in T2DM patients thereby linking its role in chronic inflammation [45] through TLR receptor signaling and the NF-kB signaling pathways.

Although we obtained useful insights through this integrative approach and simulations, our inferences are associated with caveats. First, we used the gene expression dataset from the blood tissue. Most of the measurements pointing to the involvement of genes in the comorbid condition were carried out using blood samples. However, whether the signature obtained in this work would match the gene expression patterns in the lungs of T2DM patients remains to be investigated. Second, we have used only one dataset of T2DM available publicly because the gene expression measurements are from blood in this dataset only. Considering this limitation, we have used only statistically significantly differentially expressed genes, network of genes at high confidence scores and derived pathways at FDR < 0.05. But investigations in multiple datasets are ideally desired for a globally robust leads in unraveling the susceptibilities of T2DM patients to tuberculosis. Nonetheless, we envision that the results reported in this manuscript could guide further investigation in order to enable development of gene expression based diagnostic signatures and future therapeutics for preventing or controlling TB infection in T2DM patients.

Methods

All text analytics were carried out in R 3.4.3 with the CRAN package pubmed.mineR version 1.0.10 [46]. Diabetes abstracts were extracted from the PubMed database.

1. Data Collection:

All type 2 diabetes abstracts, were downloaded from PubMed and then subjected to an R package pubmed.mineR [46] to extract the gene disease associations

2. Entity recognition and their relationships

We used sentence tokenization function of pubmed.mineR to extract sentences with co-occurrence of the two entities: genes and disease terms. Mention of genes included gene symbols including official symbols recommended by the Human Gene Nomenclature Committee's (HGNC) [47], other symbols (aliases, previous symbols), alternative names from UniProt and gene names from HGNC using *official_fn(), altnamesfun(), prevsymbol_fn()*, *alias_fn()* and *names()* functions of pubmed.mineR. False matches arise due to use of identical acronyms by authors signifying other meanings. These were manually identified and discarded. The extracted sentences were examined for 'proof of association' or 'of evidence' in terms of relationships between the entities. In cases where the relationship was not clear, the entire abstract was examined along with full text wherever available from PubMed Central. We have analyzed the trends of resulting genes to mark their significance and their annotation with Gene Ontologies (GOs) by using DAVID (Database for Annotation, Visualization and Integrated Discovery) [48]. The schematic representation of the methodology is given in **Figure 1**. These genes are termed T2DMTB genes.

3. Web server development:

The fully annotated data of T2DMTB genes was plugged into T2DiACoD [49].

4. Gene Enrichment Analysis:

To infer potential biological significance of the reported genes, we have used DAVID (Database for Annotation, Visualization and Integrated Discovery) [48]. The output table was obtained with the following information: Annotation Cluster, Enrichment Score, Category, Term, Count, % (involved genes/total genes), Genes, List Total, Pop Hits, Pop Total, Fold Enrichment, Bonferroni, Benjamini and FDR, P-Value (calculated by Fishers' exact test) as well as FDR <= 0.05 were considered strongly enriched in the annotation categories. Gene ontology enrichment was prepared using EnrichmentMap [50], a cytoscape plugin.

5. Differential Expression:

The microarray gene expression (GSE26168) data for blood samples from T2DM patients were collected from NCBI GEO [51]. The GSE26168 is a super-series and composed of two subseries, GSE21321 and GSE26167. This data contains mRNA and miRNA profiles of T2DM patients without tuberculosis. From GSE21321 subseries, 8 controls, 7 impaired fasting glucose, and 9 diabetic mRNA samples, whereas 10 control, 7 impaired fasting glucose and 9 diabetic miRNA human samples were studied. For mRNAs, the

illumina platform and for miRNAs, the miRCURY LNA microRNA array platform was used. This illumina dataset was processed using the lumi package [52]. Normalization was performed using a modified Z-score method based on median absolute deviation (MAD) [53]. For differential expression, the limma [54] package was used. The differentially expressed mRNAs and miRNAs were selected on the basis of P-values < 0.05 and Padj < 0.05.

6. Integrative Analysis

i) *Collection of deregulated miRNAs in Blood from T2DM and TB patients*

We sought to identify the miRNAs targeting T2DMTB genes to understand their role in the pathogenicity of the co-morbidity. We used NetworkAnalyst [34] to obtain gene-miRNA interactions. It uses TarBase and mirTarBase databases for information retrieval. We extracted common miRNAs deregulated in T2DM as well as in TB in blood. Further literature mining was performed for the evidence of deregulated miRNAs in both conditions.

ii) Differential Expression

The deregulated miRNAs from the literature (**Supplementary Table S5**) were further examined for differential expression in microarray expression of miRNAs from blood samples of T2DM patients (GSE26168).

iii) Identification of regulators of miRNA-146a & the gene targets of miRNA-146a

To investigate the role of the differentially expressed hsa-miRNA-146a in the disease, the regulators of miRNA-146a and the gene targets regulated by hsa-miRNA-146a were extracted using text mining from PubMed. Literature mining using pubmed.mineR was performed to extract the human host regulators and targets of miRNA-146a.

iv) STRING Analysis

All 36 T2DMTB genes were included. Among the genes identified from other approaches, only the genes differentially expressed were included among the following datasets: between diabetic patients and normals (GSE26168), host genes susceptible to TB (12 genes), regulators of miR-146a (3 genes), genes targets of miR-146a (8 genes). The invariant genes in Human genomes [55] were also included and the entire gene set was interrogated using STRING [56]. The connections between genes with score >= 0.7 only were considered.

iv) Pathways Enrichment

The STRING output provided a list of 120 significant pathways (FDR < 0.05). Our goal was to identify the pathways with potential to render susceptibility to Tuberculosis in diabetic patients. We therefore

excluded 54 disease pathways from the list. Pathways information for genes not output by STRING were manually searched for their pathways and then included in Biological Network Simulation (BioNSi).

vi) Biological Network Simulation

BioNSi [35], a biological network simulator, was used to simulate the pathways in normals and T2DM patients. The BioNSi is a discrete biological network simulator tool. Pathways were imported and merged into a large network for simulation. The average MAD Z-score of normalized expression values for normals as well as T2DM patients were provided as input to BioNSi. Based on the expression values of the genes a state value in the range 0-9 is assigned. Subsequently, the connections between the regulators of and targets of hsa-miR-146a were added manually (**Supplementary Table S2**). 29 nodes (genes) in the merged network of pathways were connected to miR-146a. These connections included the differentially expressed 3 regulators and 8 targets of miR-146a. The initial state of miR-146a in diabetics was set to '6', while in normals it was set to '9' using the scaling criteria of gene expression values of BioNSi.

The initial states assigned based on gene expression values is updated during the course of simulation using a state updation function. We used the default set up in the BioNSi tool for this purpose. Authors of BioNSi have also provided a Python script to calculate the differences between the simulation contours for a given gene and rank them in terms of the differences in the state values during the course of simulation. This information was used to trace back to pathways and we selected the highly perturbed pathways between the disease state and the normal state as the one in which multiple genes in the same pathway showed differences in the states.

Declarations

ACKNOWLEDGEMENT

We thank our lab members, for their useful feedback. JR acknowledges ICMR (Indian Council of Medical Research) for Senior Research Fellowship (File No. BIC/11(21)/2015). SR's work is supported in part by a grant from ICMR (No. BIC/5(09)/Indo-Russian/2016).

AUTHORS' CONTRIBUTION

The study was designed by S.R. and J.R. Text mining was done by was done by J.R. Network modelling and simulation was done by J.R, A.B. Data analysis was performed by J.R, A.B, U.B, and S.R. Manuscript written by J.R., A.B., U.B., and S.R. All authors have read and approved the manuscript for publication

CONFLICT OF INTEREST

All authors declare no conflicts of interests.

References

1. Chaudhry, L. A., Essa, E. B., Al-Solaiman, S. & Al-Sindi, K. Prevalence of diabetes type-2 & pulmonary tuberculosis among Filipino and treatment outcomes: A surveillance study in the Eastern Saudi Arabia. *Int. J. Mycobacteriology*, **1**, 106–109 (2012).
2. Kapur, A. & Harries, A. D. The double burden of diabetes and tuberculosis - public health implications. *Diabetes Res. Clin. Pract*, **101**, 10–19 (2013).
3. Niazi, A. K. & Kalra, S. Diabetes and tuberculosis: a review of the role of optimal glycemic control. *J. Diabetes Metab. Disord*, **11**, 28 (2012).
4. Faurholt-Jepsen, D. *et al.* Diabetes is a risk factor for pulmonary tuberculosis: a case-control study from Mwanza, Tanzania. *PloS One*, **6**, e24215 (2011).
5. Agarwal, A. K., Ginisha, G., Preeti, G., Dwivedi, S. & Swamai, P. The association between diabetes and tuberculosis may be the next challenge for global tuberculosis control worldwide. *Indian J. Endocrinol. Metab*, **20**, 732–733 (2016).
6. Harries, A. D. *et al.* Addressing diabetes mellitus as part of the strategy for ending TB. *Trans. R. Soc. Trop. Med. Hyg*, **110**, 173–179 (2016).
7. Mustafa, A. S. Recombinant and synthetic peptides to identify Mycobacterium tuberculosis antigens and epitopes of diagnostic and vaccine relevance. *Tuberc. Edinb. Scotl*, **85**, 367–376 (2005).
8. Martens, G. W. *et al.* Tuberculosis susceptibility of diabetic mice. *Am. J. Respir. Cell Mol. Biol*, **37**, 518–524 (2007).
9. Jeon, C. Y. & Murray, M. B. Diabetes mellitus increases the risk of active tuberculosis: a systematic review of 13 observational studies. *PLoS Med*, **5**, e152 (2008).
10. Karachunskiĭ, M. A., Gergert, V. I. & Iakovleva, O. B. [Specific features of cellular immunity of pulmonary tuberculosis in patients with diabetes mellitus].*Probl. Tuberk*.59–60(1997).
11. Tsukaguchi, K. *et al.* [Case study of interleukin-1 beta, tumor necrosis factor alpha and interleukin-6 production peripheral blood monocytes in patients with diabetes mellitus complicated by pulmonary tuberculosis]. *Kekkaku*, **67**, 755–760 (1992).
12. Kumar, N. P. *et al.* Type 2 diabetes mellitus coincident with pulmonary tuberculosis is associated with heightened systemic type 1, type 17, and other proinflammatory cytokines. *Ann. Am. Thorac. Soc*, **10**, 441–449 (2013).
13. Lopez-Lopez, N. *et al.* Type-2 diabetes alters the basal phenotype of human macrophages and diminishes their capacity to respond, internalise, and control Mycobacterium tuberculosis. *Mem. Inst. Oswaldo Cruz*, **113**, e170326 (2018).
14. Dinarello, C. A. Proinflammatory cytokines., **118**, 503–508 (2000).
15. Kumar, N. P. *et al.* Type 2 diabetes mellitus is associated with altered CD8(+) T and natural killer cell function in pulmonary tuberculosis., **144**, 677–686 (2015).
16. Syal, K., Srinivasan, A. & Banerjee, D. V. D. R. RXR, Coronin-1 and Interferonγ Levels in PBMCs of Type-2 Diabetes Patients: Molecular Link between Diabetes and Tuberculosis. *Indian J. Clin. Biochem. IJCB*, **30**, 323–328 (2015).

17. Qu, H. Q. *et al.* Host susceptibility to tuberculosis: insights from a longitudinal study of gene expression in diabetes. *Int. J. Tuberc. Lung Dis. Off. J. Int. Union Tuberc. Lung Dis*, **16**, 370–372 (2012).
18. Latorre, I. *et al.* A novel whole-blood miRNA signature for a rapid diagnosis of pulmonary tuberculosis. *Eur. Respir. J*, **45**, 1173–1176 (2015).
19. Tegnér, J. N. *et al.* Computational disease modeling - fact or fiction? *BMC Syst. Biol*, **3**, 56 (2009).
20. Sancristóbal, B., Tastekin, I. & Dierssen, M. Computational Models. in Molecular-Genetic and Statistical Techniques for Behavioral and Neural Research 105–131(Elsevier, 2018). doi:10.1016/B978-0-12-804078-2.00005-2.
21. Parton, A., McGilligan, V., O'Kane, M., Baldrick, F. R. & Watterson, S. Computational modelling of atherosclerosis. *Brief. Bioinform*, **17**, 562–575 (2016).
22. Andrade, B. B. *et al.* Heightened plasma levels of heme oxygenase-1 and tissue inhibitor of metalloproteinase-4 as well as elevated peripheral neutrophil counts are associated with TB-diabetes comorbidity., **145**, 1244–1254 (2014).
23. Pavan Kumar, N. *et al.* Circulating biomarkers of pulmonary and extrapulmonary tuberculosis in children. *Clin. Vaccine Immunol. CVI*, **20**, 704–711 (2013).
24. Kumar, N. P. *et al.* Type 2 diabetes - Tuberculosis co-morbidity is associated with diminished circulating levels of IL-20 subfamily of cytokines. *Tuberc. Edinb. Scotl*, **95**, 707–712 (2015).
25. Trentini, M. M., de Oliveira, F. M., Kipnis, A. & Junqueira-Kipnis, A. P. The Role of Neutrophils in the Induction of Specific Th1 and Th17 during Vaccination against Tuberculosis. *Front. Microbiol*, **7**, 898 (2016).
26. Herrera, M. T., Gonzalez, Y., Hernández-Sánchez, F., Fabián-San Miguel, G. & Torres, M. Low serum vitamin D levels in type 2 diabetes patients are associated with decreased mycobacterial activity. *BMC Infect. Dis*, **17**, 610 (2017).
27. Liu, C. H., Liu, H. & Ge, B. Innate immunity in tuberculosis: host defense vs pathogen evasion. *Cell. Mol. Immunol*, **14**, 963–975 (2017).
28. Jayachandran, R. *et al.* Survival of mycobacteria in macrophages is mediated by coronin 1-dependent activation of calcineurin., **130**, 37–50 (2007).
29. Gonzalez-Curiel, I. *et al.* Differential expression of antimicrobial peptides in active and latent tuberculosis and its relationship with diabetes mellitus. *Hum. Immunol*, **72**, 656–662 (2011).
30. Via, L. E. *et al.* Arrest of mycobacterial phagosome maturation is caused by a block in vesicle fusion between stages controlled by rab5 and rab7. *J. Biol. Chem*, **272**, 13326–13331 (1997).
31. Monin, L. *et al.* Immune requirements for protective Th17 recall responses to Mycobacterium tuberculosis challenge. *Mucosal Immunol*, **8**, 1099–1109 (2015).
32. Schmok, E. *et al.* Suppressor of Cytokine Signaling 3 in Macrophages Prevents Exacerbated Interleukin-6-Dependent Arginase-1 Activity and Early Permissiveness to Experimental Tuberculosis. *Front. Immunol*, **8**, 1537 (2017).

33. Kumar, D. *et al.* Genome-wide analysis of the host intracellular network that regulates survival of Mycobacterium tuberculosis., **140**, 731–743 (2010).
34. Xia, J., Gill, E. E. & Hancock, R. E. W. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. *Nat. Protoc*, **10**, 823–844 (2015).
35. Yeheskel, A., Reiter, A., Pasmanik-Chor, M. & Rubinstein, A. Simulation and visualization of multiple KEGG pathways using BioNSi. F1000Research 6, 2120(2017).
36. Pajuelo, D. *et al.* NAD + Depletion Triggers Macrophage Necroptosis, a Cell Death Pathway Exploited by Mycobacterium tuberculosis. *Cell Rep*, **24**, 429–440 (2018).
37. Killick, K. E. *et al.* Receptor-mediated recognition of mycobacterial pathogens. *Cell. Microbiol*, **15**, 1484–1495 (2013).
38. Ban, T., Sato, G. R. & Tamura, T. Regulation and role of the transcription factor IRF5 in innate immune responses and systemic lupus erythematosus. *Int. Immunol*, **30**, 529–536 (2018).
39. Restrepo, B. I. Diabetes and Tuberculosis.*Microbiol. Spectr.***4**, (2016).
40. Pajuelo, D., Gonzalez-Juarbe, N. & Niederweis, M. NAD hydrolysis by the tuberculosis necrotizing toxin induces lethal oxidative stress in macrophages. *Cell. Microbiol*, **22**, e13115 (2020).
41. Mizumura, K., Maruoka, S., Gon, Y., Choi, A. M. K. & Hashimoto, S. The role of necroptosis in pulmonary diseases. *Respir. Investig*, **54**, 407–412 (2016).
42. Stutz, M. D. *et al.* Necroptotic signaling is primed in Mycobacterium tuberculosis-infected macrophages, but its pathophysiological consequence in disease is restricted. *Cell Death Differ*, **25**, 951–965 (2018).
43. Dubey, R. K. Assuming the role of mitochondria in mycobacterial infection. *Int. J. Mycobacteriology*, **5**, 379–383 (2016).
44. Robinson, N. *et al.* Type I interferon induces necroptosis in macrophages during infection with Salmonella enterica serovar Typhimurium. *Nat. Immunol*, **13**, 954–962 (2012).
45. Balasubramanyam, M. *et al.* Impaired miR-146a expression links subclinical inflammation and insulin resistance in Type 2 diabetes. *Mol. Cell. Biochem*, **351**, 197–205 (2011).
46. Rani, J., Shah, A. B. R. & Ramachandran, S. pubmed.mineR: an R package with text-mining algorithms to analyse PubMed abstracts. *J. Biosci*, **40**, 671–682 (2015).
47. Povey, S. *et al.* The HUGO Gene Nomenclature Committee (HGNC). *Hum. Genet*, **109**, 678–680 (2001).
48. Dennis, G. *et al.* DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*, **4**, 3 (2003).
49. Rani, J. *et al.* T2DiACoD: A Gene Atlas of Type 2 Diabetes Mellitus Associated Complex Disorders. *Sci. Rep*, **7**, 6892 (2017).
50. Isserlin, R., Merico, D., Voisin, V. & Bader, G. D. Enrichment Map - a Cytoscape app to visualize and explore OMICs pathway enrichment results. *F1000Research*, **3**, 141 (2014).

51. Barrett, T. *et al.* NCBI GEO: mining millions of expression profiles–database and tools. *Nucleic Acids Res*, **33**, D562–566 (2005).
52. Du, P., Kibbe, W. A. & Lin, S. M. lumi: a pipeline for processing Illumina microarray. *Bioinforma. Oxf. Engl*, **24**, 1547–1548 (2008).
53. Mandić-Rajčević, S. & Colosio, C. Methods for the Identification of Outliers and Their Influence on Exposure Assessment in Agricultural Pesticide Applicators: A Proposed Approach and Validation Using Biological Monitoring. *Toxics* **7**, (2019).
54. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*, **43**, e47 (2015).
55. Pathak, A. K., Jainarayanan, A. K. & Brahmachari, S. K. *Invariant Genes in Human Genomes*. http://biorxiv.org/lookup/doi/10.1101/739706 (2019) doi:10.1101/739706.
56. Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*, **45**, D362–D368 (2017).

Tables

Due to technical limitations, table 1 is only available as a download in the Supplemental Files section.

Table 2
TB susceptible genes differentially expressed in GSE26168 with *Padj* < = 0.05. FC calculated in T2DM vs normal

Genes	P.Value	adj.P.Val	Fold Change
TNFRSF10A	0.00031	0.037	1.46
MSRA	0.00036	0.039	1.24
GPR148	0.00047	0.040	0.82
SLC37A3	0.00062	0.042	1.35
PXK	0.00087	0.043	1.27
PROK2	0.0013	0.047	1.49
REV3L	0.0014	0.048	1.39
PGM1	0.0014	0.048	1.28
HIST3H2A	0.0015	0.048	1.49
PLAC4	0.0015	0.048	0.82
LETM2	0.0016	0.048	1.18
EMP2	0.0017	0.049	0.72

Table 3

a: Genes from non-diabetics with a minimum Euclidean distance of 2 reached a steady state either prior to or coincident with diabetics and mapped to KEGG pathways.

Pathway	Genes (Euclidean Distance)	Score[1]
Necroptosis	*GLUD1 (4), TRPM7 (4), DNM1L (4), RIPK3 (4), GLUL (4), TICAM1 (4), SMPD1 (4), PGAM5 (4), RIPK1-TRAF6 (4), MLKL (4), TICAM2 (3), TICAM1-TICAM2 (3), TLR4 (2)*	13
NOD-like receptor signaling pathway	*ATG12 (5), TRPM7 (4), DNM1L (4), RIPK3 (4), TICAM1 (4), RIPK1-TRAF6 (4), NLRP1 (2), TLR4 (2), ATG16L1 (2)*	9
Toll-Like receptor signaling pathway	*TICAM1 (4), IRAK1-TRAF6 (4), RIPK1-TRAF6 (4), TICAM2 (3), TICAM1-TICAM2 (3), TIRAP-MyD88 (3), TLR4 (2)*	7
NF-kappa-B signaling pathway	*TICAM1 (4), IRAK1-TRAF6 (4), RIPK1-TRAF6 (4), TICAM2 (3), TICAM1-TICAM2 (3), TLR4 (2)*	6
[1]: Score is the number of genes in a given pathway with minimum Euclidean distance of 2.		

Table 3

b: Pathway distribution of genes from diabetics exhibited differences and reached steady state before or together with non-diabetic.

Pathway	Genes (Distance)	Score[1]
Toll-like receptor signaling pathway	*IRAK1 (5), TRAF6 (4), IRF5 (3), IRAK4 (3), MyD88 (3), IRAK1-IRAK4 (3)*	7
MAPK signaling pathway	*IRAK1 (5), TRAF6 (4), IRAK4 (3), MyD88 (3), IRAK1-IRAK4 (3), BAK1 (2)*	6
NF-Kappa B signaling pathway	*IRAK1 (5), TRAF6 (4), IRAK4 (3), MyD88 (3), IRAK1-IRAK4 (3)*	6
NOD-like receptor signaling pathway	*TRAF6 (4), IRAK4 (3), MyD88 (3), IRAK1-IRAK4 (3), GSDMDC1 (3)*	5
[1]: Score is the number of genes in a given pathway with minimum Euclidean distance of 2.		

Figures

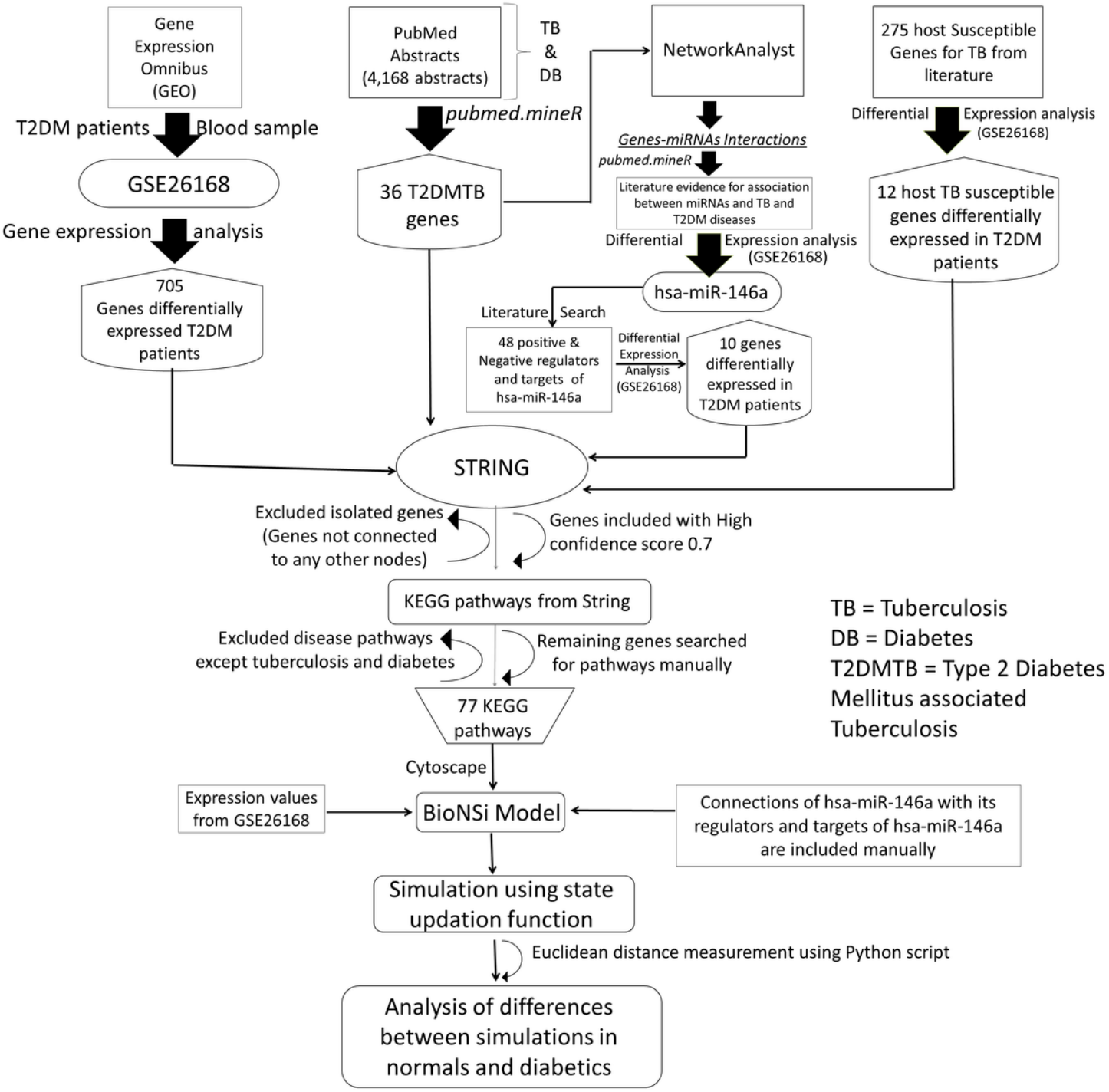

Figure 1

Schematic representation of the methodology. The genes for the final network model are assembled from the multiple sources as per the work flow including text mining and differential expression analysis.

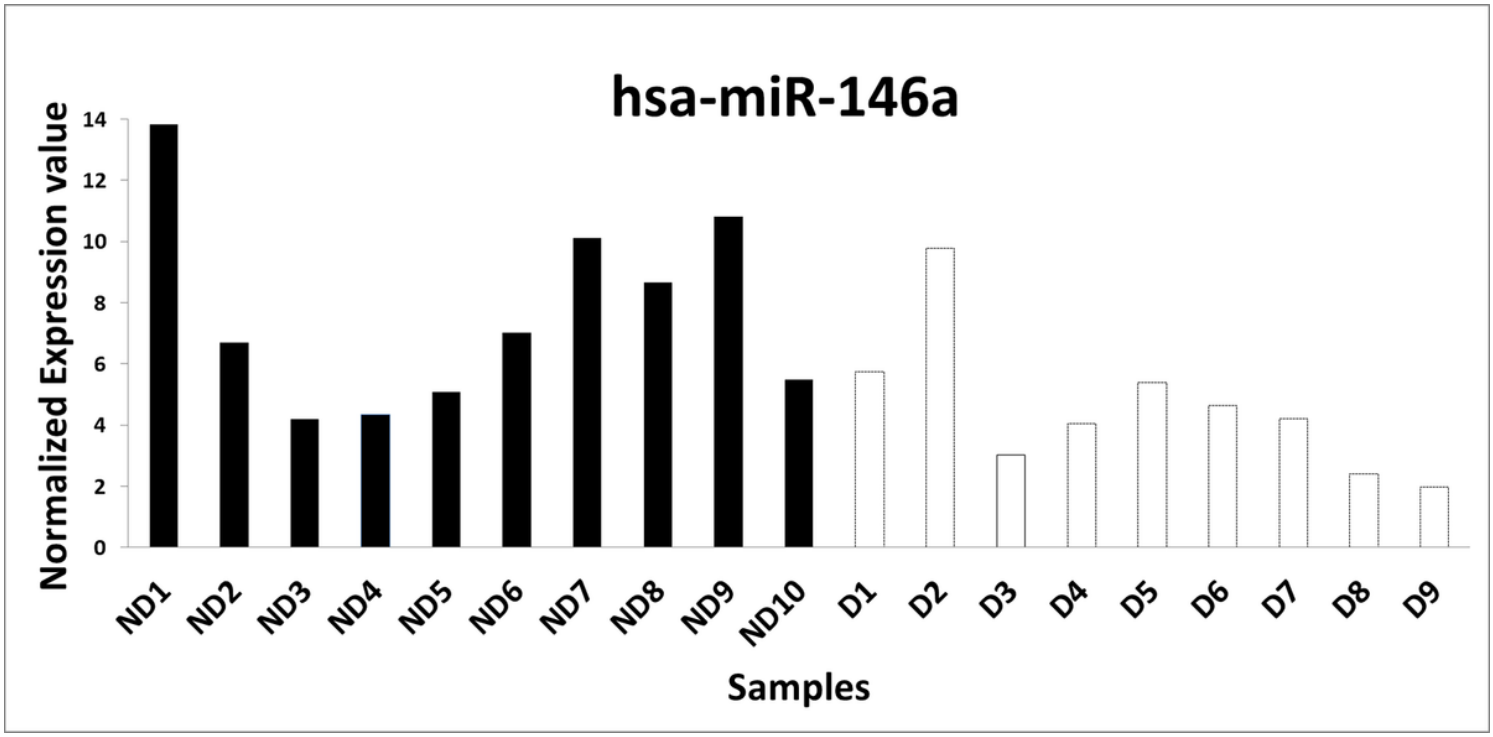

Figure 2

Differential expression of hsa-miR-146a in T2D blood samples from GSE26168. Filled bars represent non-diabetics, empty bars represent T2DM patients.

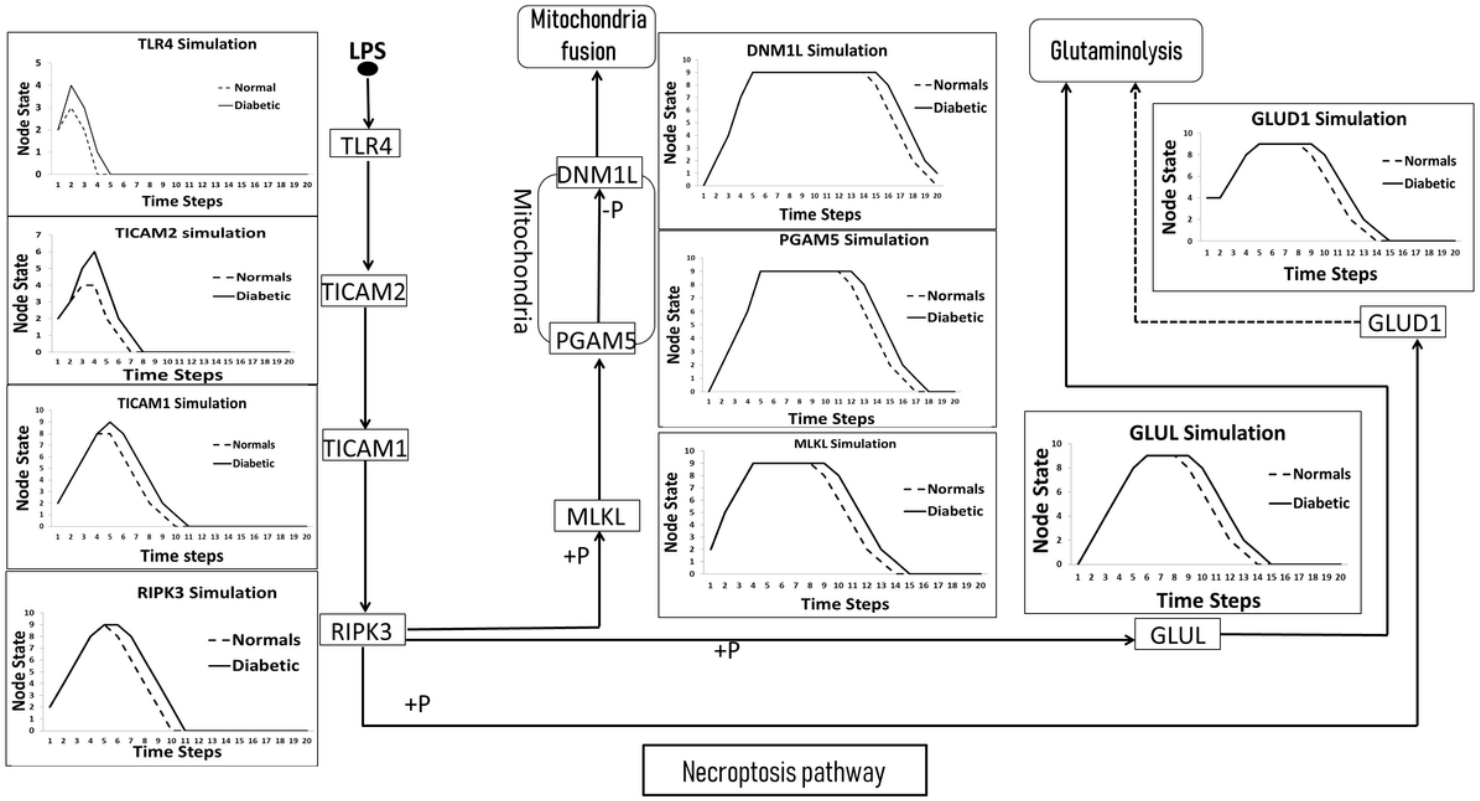

Figure 3

Schematic representation of the most perturbed pathway, the Necroptosis pathway (has04217) during simulations using BioNSi. The simulation graphs are shown in square boxes beside each protein in the pathway. The +P and –P signifies phosphorylation and de-phosphorylation respectively. The regular arrows represent the regular route of metabolic flow, broken arrows represent alternate route.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryfiguresandtables.docx
- SupplementaryTableS3DiffExpressedGenes.xlsx
- SupplementaryTableS7STRINGPATHWAYS.xlsx
- Table1.jpg