

# Bioinformatics analysis of RNA-seq data revealed critical genes in type 1 diabetes and molecular docking studies

**Prashanth G**

Department of General Medicine, Basaveshwara Medical College, Chitradurga 577501, Karnataka, India.

<https://orcid.org/0000-0003-2571-5734>

**Basavaraj Vastrad**

Department of Biochemistry, Basaveshwar College of Pharmacy, Gadag, Karnataka 582103, India.

<https://orcid.org/0000-0003-2202-7637>

**Anandkumar Tengli**

Department of Pharmaceutical Chemistry, JSS College of Pharmacy, Mysuru and JSS Academy of Higher Education & Research, Mysuru, Karnataka, 570015, India <https://orcid.org/0000-0001-8076-928X>

**Chanabasayya Vastrad** (✉ [channu.vastrad@gmail.com](mailto:channu.vastrad@gmail.com))

Biostatistics and Bioinformatics, Chanabasava Nilaya, Bharthinagar, Dharwad 580001, Karnataka, India. <https://orcid.org/0000-0003-3615-4450>

**Iranna Kotturshetti**

Department of Ayurveda, Rajiv Gandhi Education Society's Ayurvedic Medical College, Ron 562209, Karnataka, India. <https://orcid.org/0000-0003-1988-7345>

---

## Research Article

**Keywords:** bioinformatics, type 1 diabetes, differentially expressed genes, enrichment analysis, pathways

**Posted Date:** December 22nd, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-133026/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Type 1 diabetes (T1D) is a serious threat to childhood life and has fairly complicated pathogenesis. Profound attempts have been made to enlighten the pathogenesis, but the molecular mechanisms of T1D are still not well known.

**Methods:** To identify the candidate genes in the progression of T1D, Expression profiling by high throughput sequencing dataset GSE123658 was downloaded from Gene Expression Omnibus (GEO) database. The differentially expressed genes (DEGs) were identified, and gene ontology (GO) and pathway enrichment analyses were performed. The protein-protein interaction network (PPI), modules, target gene - miRNA regulatory network and target gene - TF regulatory network analysis were constructed and analyzed using HIPPIE, miRNet, NetworkAnalyst and Cytoscape. Finally, validation of hub genes was conducted by using ROC (Receiver operating characteristic) curve and RT-PCR analysis. Molecular docking studies were performed.

**Results:** A total of 284 DEGs were identified, consisting of 142 up regulated genes and 142 down regulated genes. The gene ontology (GO) and pathways of the DEGs include cell-cell signaling, vesicle fusion, plasma membrane, signaling receptor activity, lipid binding, signaling by GPCR and innate immune system. Four hub genes were identified and biological process analysis revealed that these genes were mainly enriched in cell-cell signaling, cytokine signaling in immune system, signaling by GPCR and innate immune system. ROC curve and RT-PCR analysis showed that EGFR, GRIN2B, POLR2A and PRKACA might be involved in the advancement of T1D. Molecular docking studies showed high docking score.

**Conclusions:** DEGs and hub genes identified in the present investigation help us understand the molecular mechanisms underlying the advancement of T1D, and provide candidate targets for diagnosis and treatment of T1D.

## Introduction

Type 1 diabetes (T1D) (insulin-dependent) is a core challenge for endocrine research around the world [1]. Approximately 5% to 10% of the childhood population is affected with T1D worldwide [2]. T1D affects the eyes, kidneys, heart, peripheral and autonomic nervous systems [3]. Pancreatic cells, particularly  $\beta$ -cells, play a key role in the occurrence and progression of T1D [4]. Although disease etiology still remains unknown. Genetic and environmental factors were associated with pathogenesis of T1D [5-6]. Nisticò et al [7] reported that up regulation of molecular biomarker CTLA-4 was linked with progression of T1D. Thus, identifying key molecular biomarkers is essential for early diagnosis, prevention, and treatment of T1D.

It worth a lot of money and time to identify disease related molecular biomarkers by experiment alone. With the wide application of microarray platform, there were huge genomics data deposited in public

databases [8]. The progression of computational tools gives us an alternative method to diagnose novel molecular biomarkers.

In this investigation, we employed the bioinformatics approach to discover the differentially expressed genes between T1D patients and healthy donors. Original microarray dataset GSE123658 was downloaded. 39 T1D patients' samples and 43 healthy donors' samples were analyzed in our investigation. Commonly altered DEGs were isolated from integrated data. Additionally, GO/ REACTOME pathway analysis, construction of protein–protein interaction network, modules, target gene - miRNA regulatory network and target gene - TF regulatory network analysis were performed to analyze these data. Four hub genes (EGFR, GRIN2B, POLR2A and PRKACA) were identified. ROC (Receiver operating characteristic) curve and RT-PCR analysis were used to verify clinically relevant hub genes. The aim of this investigation was to gain a better understanding of the underlying molecular mechanisms and to discover molecular biomarkers for T1D.

## Material And Methods

### Data resources

Expression profiling by high throughput sequencing dataset GSE123658 was downloaded from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>) [9]. The data was produced using a GPL18573 Illumina NextSeq 500 (Homo sapiens). The GSE123658 dataset contained data from 82 samples, including 39 T1D patients' samples and 43 healthy donors' samples.

### Identification of DEGs

The identification of DEGs between 39 T1D patients' samples and 43 healthy donors' samples was performed using Limma package in R bioconductor. To correct the discovery of statistically important molecular biomarkers and limitations of false-positives, we using the adjusted P-value and Benjamini and Hochberg false discovery rate method [10]. A  $|\log_2FC| > 1$  and  $P\text{-value} < 0.05$  were used as considered statistically significant. The volcano plot was implemented using ggplot2 package, and the heat map was established using gplots package in R language.

### Gene Ontology (GO) and pathway enrichment analyses of DEGs

Gene Ontology (GO) (<http://www.geneontology.org>) analysis is a routine analysis for annotating genes and determining biological component, including biological process (BP), cellular component (CC) and molecular function (MF) [11]. REACTOME (<https://reactome.org/>) [12] pathway database is applied for classification by correlating gene sets into their respective pathways. The ToppGene (ToppFun) (<https://toppgene.cchmc.org/enrichment.jsp>) [13] is a gene functional classification tool that objective to provide a extensive set of functional annotation tools for authors to recognize the biological explanation behind large lists of genes.  $P < 0.05$  was find statistically significant.

### PPI network construction and module analysis

The online Human Integrated Protein-Protein Interaction rEference (HIPPIE) (<http://cbdm.uni-mainz.de/hippie/>) [14] online database was used to predict the PPI network information. Analyzing the interactions and functions between DEGs may provide information about the mechanisms of generation and development of disease (PPI score > 0.4). Cytoscape 3.8.0 (<http://www.cytoscape.org/>) [15] is a bioinformatics platform for constructing and visualizing molecular interaction networks. The Network Analyzer Cytoscape plug-in was used to find hub genes were screened using highest node degree [16], betweenness centrality [17], stress centrality [18] and closeness centrality [19] methods and hub genes were further analyzed for pathway and GO enrichment analysis. The plug-in PEWCC1 (<http://apps.cytoscape.org/apps/PEWCC1>) [20] of Cytoscape was applied to detect densely connected regions in PPI networks. The PPI networks were constructed using Cytoscape and the most key module in the PPI networks was preferred using PEWCC1. The criteria for selection were set as follows: Max depth = 100, degree cut-off = 2, Node score cut-off = 0.2, PEWCC1 scores > 5, and K-score = 2.

### Prediction of key miRNAs

miRNet database (<https://www.mirnet.ca/>) [21] online database was used to predict miRNAs that targeted the DEGs associated with T1D. The DEGs were selected according to the screening criterion of P value < 0.05. The results were exported into the Cytoscape software for analysis. The target genes - miRNA regulatory network was constructed through network topology prosperities. The node degree was determined using the Network analysis plugin, and miRNAs with a node degree > 12.

### Prediction of key TFs

NetworkAnalyst database (<https://www.networkanalyst.ca/>) [22] online database was used to predict miRNAs that targeted the DEGs associated with T1D. The DEGs were selected according to the screening criterion of P value < 0.05. The results were exported into the Cytoscape software for analysis. The target genes - TF regulatory network was constructed through network topology prosperities. The node degree was determined using the Network analysis plugin, and TFs with a node degree > 12.

### Validation of hub genes

Receiver operating characteristic (ROC) curve analysis was implemented to calculate the sensitivity and specificity of the DEGs for T1D diagnosis using the R package by using the generalized linear model (GLM) in machine learning algorithms [23]. An area under the curve (AUC) value was determined and used to label the ROC effect. In RT-PCR analysis, total RNA was isolated from cultured cells of diabetics and normal using a TRI Reagent® (Sigma, USA). cDNA was synthesized using 2.0 µg of total RNA with the FastQuant RT kit (with gDNase; Tiangen Biotech Co., Ltd.). The relative mRNA expression was measured in the QuantStudio 7 Flex real-time PCR system (Thermo Fisher Scientific, Waltham, MA, USA). The relative expression levels were determined by the  $2^{-\Delta\Delta Ct}$  method and normalized to internal control  $\beta$ -actin [24]. All RT-PCR reactions were performed in triplicate. The primers used to explore mRNA expression of ten hub genes were shown in Table 1.

## Molecular Docking studies

The surflex-docking studies of the designed molecule, standard on overexpressed genes of PDB protein were performed using SYBYL-X 2.0 perpetual software. All the designed molecules and standard were outlined using ChemDraw Software, imported and saved in sdf. templet using openbabel free software. The protein structures of EGFR(epidermal growth factor receptor) and its co-crystallised protein of PDB code 2XYJ, 2XYX and GRIN2B its co-crystallisedprotein of PDB code and its co-crystallised protein of PDB code 5EWL, 6E7V wereextracted from Protein Data Bank [25-26]. Together with the TRIPOS force field, GasteigerHuckel (GH) charges were added to all designed molecules and standard for the structure optimization process.Additionally, using MMFF94s and MMFF94 algorithm methods, energy minimization was done.The protein preparation was carried out after incorporation of protein.The co-crystallized ligand and all water molecules were removed from the crystal structure; more hydrogen's were added and the side chain was set. TRIPOS force field was used for the minimization of structure. The compounds' interaction efficiency with the receptor was represented by the Surflex-Dock score in kcal / mol units. The interaction between the protein and the ligand, the best pose was incorporated into the molecular area. The visualization of ligand interaction with receptor is done by using discovery studio visualizer.

## Results

### Identification of DEGs

With a threshold of P-value <0.05 and absolute value of  $|\log_2FC| > 1$ , DEGs were identified from dataset. 284 DEGs were screened from the GSE123658 dataset, which consisted of 142 up regulated genes and 142 down regulated genes and only top ten up regulated genes and down regulated genes are listed in Table 2. Volcano plots were used to visualize differential expression of genes between the T1D group and healthy donors group (Fig.1). A heat map showed expression profiling of DEGs in the analysis result (Fig.2).

### Gene Ontology (GO) and pathway enrichment analyses of DEGs

To identify the pathways which had the most significant involvement with the genes identified, up regulated and down regulated genes are listed in Table 3 and Table 4. DEGs were submitted into ToppGene for GO and REACTOME pathway enrichment analysis. GO enrichment analysis revealed that in BP terms, the up regulated genes were mainly enriched in cell-cell signaling and reproductive process, whereas down regulated genes were mainly enriched in vesicle fusion and biological adhesion. In CC terms, up regulated genes were mainly enriched in integral component of plasma membrane and supramolecularfiber, whereas down regulated genes were mainly enriched in whole membrane and cell junction. In MF terms, up regulated genes were mainly enriched in signaling receptor activity and structural molecule activity, whereas down regulated genes were mainly enriched in lipid binding and GTPase binding. REACTOME pathway analysis demonstrated that up regulated genes were significantly

enriched in signaling by GPCR and muscle contraction, whereas down regulated genes were mainly enriched in innate immune system and cytokine signaling in immune system

### **PPI network construction and module analysis**

Interactions between the identified up and down regulated genes were reported by constructing a PPI network. In total, there were 5017 nodes and 8133 edges in the network (Fig.3). According to node degree, betweenness centrality, stress centrality and closeness centrality levels, the top four hub nodes were: EGFR (degree, 1343; betweenness, 0.514227; stress, 4.35E+08; closeness, 0.441893), GRIN2B (degree, 195; betweenness, 0.008701; stress, 14535992; closeness, 0.282164), POLR2A (degree, 512; betweenness, 0.095104; stress, 2E+08; closeness, 0.3738) and PRKACA (degree, 350; betweenness, 0.104658; stress, 44280174; closeness, 0.378052), and are listed in Table 5. Significant modules were subsequently constructed with 12 nodes and 23 edges for up regulated genes and 5 nodes and 10 edges for down regulated genes, which gained the highest PEWCC1 score (Fig.4). Subsequent functional enrichment analysis revealed that the genes in these modules were mainly enriched in cell-cell signaling and cytokine signaling in immune system.

### **Prediction of key miRNAs**

The regulatory relationships between the target genes and their miRNAs were established using Cytoscape, which showed that the single gene was regulated by multiple miRNAs is shown in Fig.4A. Subsequently, 106 miRNAs (ex, hsa-mir-4257) targeting GRIN2B, 83 miRNAs (ex, hsa-mir-4257) targeting EGFR, 83 miRNAs (ex, hsa-mir-564) targeting EGFR, 190 miRNAs (ex, hsa-mir-4300) targeting TLN1 and 139 miRNAs (ex, hsa-mir-5694) targeting IGF2R were verified in miRNet database are listed in Table 6. Integrating with the result of REACTOME pathway analysis, it was indicated that these key target genes - miRNA network was mainly involved in the signaling by GPCR and innate immune system.

### **Prediction of key TFs**

The regulatory relationships between the target genes and their TFs were established using Cytoscape, which showed that the single gene was regulated by multiple TFs is shown in Fig.4B. Subsequently, 19 TFs (ex, FOXD1) targeting RGS4, 16 TFs (ex, GATA2) targeting EYA1, 15 TFs (ex, STAT3) targeting PRKACA and 14 TFs (ex, TFAP2A) targeting GAB2 were verified in NetworkAnalyst database are listed in Table 6. Integrating with the result of REACTOME pathway analysis, it was indicated that these key target genes - TF network was mainly involved in the signaling by GPCR and innate immune system.

### **Validation of hub genes**

As these 4 genes are remarkably expressed in T1D, we executed a ROC curve analysis to calculate their sensitivity and specificity for the diagnosis of T1D. As shown in Fig. 5A, EGFR, GRIN2B, POLR2A and PRKACA achieved an AUC value of > 0.985, demonstrating that these genes have high sensitivity and specificity for T1D diagnosis. We further used RT-PCR to detect the mRNA expression of the hub gene. The 4 hub genes contain two up regulated genes (EGFR and GRIN2B) and two down regulated gene

(POLR2A and PRKACA). The RT-PCR data showed that although the trend of expression patterns of these 4 hub genes were consistent, among these up regulated genes, only EGFR and GRIN2B were significantly up regulated in T1D. In addition, the expression of POLR2A and PRKACA were reduced in T1D (Fig. 5B).

## Molecular Docking studies

The docking simulations are conducted in the present study is to identify the active site conformation and major interactions responsible for complex stability with the receptor binding sites. In type 1 diabetes mellitus over expression of genes are identified and their proteins of x-ray crystallographic structure are selected from PDB for docking studies. The drugs containing thiazolidindione ring pioglitazone are most commonly used either alone or in combination with other antidiabetic drugs. The docking studies of pioglitazone (as standard) and designed molecules containing the heterocyclic ring of thiazolidinedione have been carried out using Sybyl X 2.1 drug design software. The docking studies were performed to know the binding interaction of pioglitazone standard and designed molecules on identified over expressed genes of protein. The X- RAY crystallographic structure of two proteins from each over expressed genes of EGFR (epidermal growth factor receptor) and its co-crystallised protein of PDB code 2XYJ, 2XYX and GRIN2B (glutamate ionotropic receptor NMDA type of subunit 2B) and its co-crystallised protein of PDB code 5EWL, 6E7V respectively were selected for the docking. The research of designed molecules had been carried out to perceive the potential molecule. A total of 54 molecules were designed and the binding score greater than six are said to be good, few molecules and pioglitazone obtained C-score greater than 7 with one in each of EGFR and GRIN2B over expressed genes of PDB code 2XYJ and 6E7V respectively. The designed molecules The two over expressed gens such as EGFR (epidermal growth factor receptor) and GRIN2B (glutamate ionotropic receptor NMDA type of subunit 2B) and two proteins from each were selected for docking. The compounds obtained with a good binding score with one protein in each over expressed genes out of two are PIO, TSPZP24, TSPZP25, TBPZP5, TSPZP27, TSIO6, TBPZP46, TSPZP23, TBIO34, TSPZ15 and TSPZP24 with binding score of 8.396, 7.6243, 7.3803, 17.3313, 7.2212, 7.20417.2021, 7.1987, 7.0795 and 7.0591 with PDB code of 2XYJ and 6E7V respectively (Fig. 6). Compounds such as TSIO8, TBPZP54, TBPZP49, TBPZP51, TBIO36, TBPZ44, TSPZ18, TSPZ13, TBIO35, TSPZ17, TBPZP53, TBPZ45, TBPZP50, TBIO30, TSPZ16, TBIO36, TBPZ40, TBPZ39, TSPZP24 and TSPZP24, TBPZP48 and TSPZ15, TSIO6, TSPZP24, TSPZP26 and TBPZ43, TBPZ42, TSPZ15, TBIO34 with binding score of 6.9341, 6.9257, 6.8768, 6.8488, 6.7296, 6.6452, 6.5682, 6.5625, 6.542, 6.407, 6.4023, 6.3781, 6.3091, 6.254, 6.1603, 6.160, 6.0903, 6.0703, 6.0263 and 6.1225, 5.5282 and 6.4652, 6.3819, 6.3205, 6.208 and 6.5068, 6.3741, 6.256, 5.9959 with PDB code of 2XYJ and 2XYX and 5EWL and 6E7V respectively. The molecules obtained with less binding score with respect to all proteins is TSPZP19 4.7843, TBPZ38 4.6791, TBPZ41 4.5968, TBPZ39 3.9732, TBPZ41 3.7662, TSPZP26 3.6423 and TBPZ41 4.9979, TBIO32 4.9247 and TBPZP48 4.5557, TSPZP21 4.1394 the values are depicted in Table 7. The binding score of the predicted molecules are compared with that of the standard pioglitazone used in type 1 diabetes mellitus, the standard obtained good binding score with all proteins. The molecule TSPZP24 and TSPZP22 has highest binding score its interaction with protein 6E7V and XYJ, and hydrogen bonding and other bonding interactions with amino acids are depicted by 3D (Fig. 7) and 2D (Fig. 8).

## Discussion

Pathophysiological complications of T1D remains undetermined, hyperglycemia develops to play a key role [27]. Although T1D is rare compared to type 2 diabetics, it may affect in any part of the body, such as the eyes, kidneys, heart, peripheral and autonomic nervous systems [28]. Considering the poor prognosis of T1D, understanding the specific molecular biomarkers of the disease is important for initial diagnosis and therapy to increase survival rates. In this investigation, we examined the expression profiling by high throughput sequencing dataset GSE123658 including T1D group and healthy donors group to identify the molecular mechanism of T1D and seek some molecular biomarkers. Bioinformatics analysis of these biological factors is applied to explore genes that are favorable to treatment.

In total, 142 up and 142 down regulated genes were identified. Polymorphic ARMS2 gene was reported to be associated with retinopathy and coronary artery disease in T1D [29]. Polymorphic AS3MT gene has an important role in the development of T1D [30]. CRHR2 is a protein coding gene which was first reported aberrantly over expressed and plays important roles in hypertension [31], but this gene might be involved in development of T1D with hypertension. Garcia-Martínez et al [32] reported that decrease expression of RNF122 proteins correlated with attention deficit hyperactivity disorder, but this gene might be linked with progression of T1D with hyperactivity disorder. Previous investigation have reported that decrease expression autophagy-regulating TP53INP2 was associated in the development of T1D [33]. Polymorphic CRTC2 gene was found to be involved in the progression of type 2 diabetes [34], but this polymorphic gene might be responsible for advancement of T1D.

GO and REACTOME pathway enrichment analyses were performed to explore interactions among the up and down regulated genes. Mutation in MYH6 gene contributes to progression of congenital heart defects [35], but this gene might be liable for progression of T1D with congenital heart defects. AVP (arginine vasopressin) has previously reported and has been suggested to play an important role in the T1D [36]. Up regulation of GRIA4 is associated with abdominal aortic aneurysm [37], but this gene might be answerable for advancement of T1D with abdominal aortic aneurysm. Kochetova et al [38] showed the possible involvement of GRIN2B in the development of type 2 diabetes, but this gene might be linked with development of T1D. The DCC (DCC netrin 1 receptor) gene has been known to participate in diabetic nephropathy [39]. Ruiz de Azua et al [40] showed that RGS4 was up regulated in type 2 diabetes, but this gene might be key for T1D. Wang et al. [41] reported that polymorphic GJA1 gene was expressed in patients with atrial fibrillation, but this might be play important role in T1D with atrial fibrillation. The high mRNA expression levels of GREM1 [42] and EGFR (epidermal growth factor receptor) [43] were correlated with T1D. García et al [44] reported that TRHR (thyrotropin releasing hormone receptor) was essential for progression of hypertension, but this gene might be associated with development of T1D with hypertension. PTPRT (protein tyrosine phosphatase receptor type T) is a key regulator of obesity and insulin resistance [45], but this gene might be involved in progression of T1D. Recent studies have shown that polymorphic FCAMR (Fc fragment of IgA and IgM receptor) gene is associated with progression of atherosclerosis [46], but this gene might be crucial for progression of T1D with atherosclerosis. Altered expression of CCL19 was observed to be associated with the progression of diabetic nephropathy [47]

and may be considered to be a novel biomarker for T1D. Down-regulation of DSP (desmoplakin) has been identified in cardiomyopathy [48], but this gene might be linked with progression of T1D with cardiomyopathy. ITIH4 was identified as a key gene associated with progression of coronary heart disease [49], but this gene may be responsible for advancement of T1D with coronary heart disease. Miyashita et al [50] reported that the low expression of the GAB2 gene is correlated with Alzheimer disease progression, but this gene might be crucial in advancement of T1D with Alzheimer disease. Notably, recent investigation have proposed that IGF2R is linked with the advancement and progression of T1D [51]. Investigation have shown that MYADM (myeloid associated differentiation marker) can be used as biomarkers for hypertension [52], but this gene might be liable for advancement of T1D with hypertension. TPCN2 [53] and APOL1 [54] genes are a key factors that initiates a type 2 diabetes, but these genes might be responsible for development of T1D. MEFV (MEFV innate immunity regulator, pyrin) down regulation can enhance ischemic heart disease [55], but this gene might be key for progression of T1D with ischemic heart disease. A previous investigation reported that the expression of DTX4 was associated with obesity [56], but this gene might be linked with progression of T1D with obesity. These GO term and REACTOME pathway analyses indicated the possible direction of experimental validation.

PPI network, modules, target gene - miRNA regulatory network and target gene - TF regulatory network analysis demonstrated that both up- and down regulated genes interacted directly or indirectly. The more edges associated with genes, miRNAs and TFs indicated the more potential selection for the targets.

Baptista et al [57] and Parackova et al [58] suggested that KIF1A and STAT3 were involved in progression of T1D. Polymorphic GATA2 gene plays a critical role in coronary artery disease [59], but this gene might be responsible for advancement of T1D with coronary artery disease. POLR2A, PRKACA (protein kinase cAMP-activated catalytic subunit alpha), CIITA (class II major histocompatibility complex transactivator), GABARAP (GABA type A receptor-associated protein), DSP (desmoplakin), hsa-mir-4257, hsa-mir-564, hsa-mir-4300, hsa-mir-5694, RGS4, FOXD1, EYA1, TFAP2A and GAB2 might be considered as a novel biomarkers associated with the development of T1D.

In addition, we also performed validation of hub genes by ROC analysis and RT-PCR. Results showed that these hub genes differentiated T1D group from healthy donors group, and may be candidates for diagnostic biomarkers and new therapeutic target. Moreover, EGFR, GRIN2B, POLR2A and PRKACA are involved in progression of T1D.

In conclusion, the present investigation was designed to identify DEGs that may be involved in the progression of T1D. A total of 284 DEGs and 4 hub genes were identified and may be regarded as diagnostic biomarkers and new therapeutic target for T1D. However, further investigations are vital to enlighten the biological function of these genes in T1D.

## Declarations

## Acknowledgement

I thank Felipe Leal Valentim,INSERM/UPMC, Unité Immunologie-Immunopathologie-Immunothérapie, Paris, France very much, the author who deposited their microarray dataset, GSE123658, into the public GEO database.

### **Conflict of interest**

The authors declare that they have no conflict of interest.

### **Ethical approval**

This article does not contain any studies with human participants or animals performed by any of the authors.

### **Informed consent**

No informed consent because this study does not contain human or animals participants.

### **Availability of data and materials**

The datasets supporting the conclusions of this article are available in the GEO (Gene Expression Omnibus) (<https://www.ncbi.nlm.nih.gov/geo/>) repository. [(GSE123658) (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE123658>)]

### **Consent for publication**

Not applicable.

### **Competing interests**

The authors declare that they have no competing interests.

### **Author Contributions**

P. G - Methodology and validation

B. V - Writing original draft, and review and editing

A. T - Formal analysis and validation

C. V - Software and investigation

I. K - Supervision and resources

### **Authors**

Prashanth G

ORCID ID: 0000-0003-2571-5734

Basavaraj Vastrad ORCID ID: [0000-0003-2202-7637](https://orcid.org/0000-0003-2202-7637)

Anandkumar Tengli ORCID ID: [0000-0001-8076-928X](https://orcid.org/0000-0001-8076-928X)

Chanabasayya Vastrad ORCID ID: [0000-0003-3615-4450](https://orcid.org/0000-0003-3615-4450)

Iranna Kotturshetti ORCID ID: [0000-0003-1988-7345](https://orcid.org/0000-0003-1988-7345)

## References

1. Atkinson MA, Eisenbarth GS. Type 1 diabetes: new perspectives on disease pathogenesis and treatment. *Lancet*. 2001;358(9277):221-229. doi:10.1016/S0140-6736(01)05415-0
2. Costacou T. The Epidemiology of Cardiovascular Disease in Adults with Type 1 Diabetes. *Curr Diabetes Rev*. 2017;13(6):520-527. doi:2174/1573399812666160927122643
3. Silverstein J, Klingensmith G, Copeland K, et al. Care of children and adolescents with type 1 diabetes: a statement of the American Diabetes Association. *Diabetes Care*. 2005;28(1):186-212. doi:2337/diacare.28.1.186
4. Foulis AK, Farquharson MA, Meager A. Immunoreactive alpha-interferon in insulin-secreting beta cells in type 1 diabetes mellitus. *Lancet*. 1987;2(8573):1423-1427. doi:10.1016/s0140-6736(87)91128-7
5. Pociot F, Lernmark Å. Genetic risk factors for type 1 diabetes. *Lancet*. 2016;387(10035):2331-2339. doi:10.1016/S0140-6736(16)30582-7
6. Butalia S, Kaplan GG, Khokhar B, Rabi DM. Environmental Risk Factors and Type 1 Diabetes: Past, Present, and Future. *Can J Diabetes*. 2016;40(6):586-593. doi:10.1016/j.jcjd.2016.05.002
7. Nisticò L, Buzzetti R, Pritchard LE, et al. The CTLA-4 gene region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. *Belgian Diabetes Registry. Hum Mol Genet*. 1996;5(7):1075-1080. doi:10.1093/hmg/5.7.1075
8. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. *Science*. 2013;339(6127):1546-1558. doi:1126/science.1235122
9. Clough E, Barrett, T. The Gene Expression Omnibus Database. *Methods. Mol. Biol*. 2016,1418,93-110. doi:10.1007/978-1-4939-3578-9\_5
10. Winnepenninckx V, Lazar V, Michiels S, et al. Gene expression profiling of primary cutaneous melanoma and clinical outcome. *J Natl Cancer Inst*. 2006;98(7):472-482. doi:10.1093/jnci/djj10
11. Thomas PD. The Gene Ontology and the Meaning of Biological Function. *Methods Mol Biol*. 2017;1446:15-24. doi:10.1007/978-1-4939-3743-1\_2
12. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, Haw R, Jassal B, Korninger F, May B et al The Reactome Pathway Knowledgebase. *Nucleic Acids Res*. 2018;46(D1):D649–D655. doi:10.1093/nar/gkx1132
13. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res*. 2009;37(Web Server issue):W305-W311.

doi:1093/nar/gkp427

14. Alanis-Lobato G, Andrade-Navarro MA, Schaefer MH. HIPPIE v2.0: enhancing meaningfulness and reliability of protein-protein interaction networks. *Nucleic Acids Res.* 2017;45(D1):D408-D414. doi:1093/nar/gkw985
15. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13(11):2498-2504. doi:1101/gr.1239303
16. Przulj N, Wigle DA, Jurisica I. Functional topology in a network of protein interactions. *Bioinformatics.* 2004;20(3):340–348. doi:1093/bioinformatics/btg415
17. Nguyen TP, Liu WC, Jordán F. Inferring pleiotropy by network analysis: linked diseases in the human PPI network. *BMC Syst Biol.* 2011;5:179. Published 2011 Oct 31. doi:1186/1752-0509-5-179
18. Shi Z, Zhang B. Fast network centrality analysis using GPUs. *BMC Bioinformatics.* 2011;12:149. doi:1186/1471-2105-12-149
19. Fadhal E, Gamielien J, Mwambene EC. Protein interaction networks as metric spaces: a novel perspective on distribution of hubs. *BMC Syst Biol.* 2014;8:6. doi:1186/1752-0509-8-6
20. Zaki N, Efimov D, Berenguères J. Protein complex detection using interaction reliability assessment and weighted clustering coefficient. *BMC Bioinformatics.* 2013;14:163. doi:1186/1471-2105-14-163
21. Fan Y, Xia J (2018) miRNet-Functional Analysis and Visual Exploration of miRNA-Target Interactions in a Network Context. *Methods Mol Biol* 1819:215-233. doi:1007/978-1-4939-8618-7\_10
22. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J (2019) NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res* 47:W234-W241. doi:1093/nar/gkz240
23. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Müller M (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:77. doi:1186/1471-2105-12-77
24. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>(Delta Delta C(T)) Method. *Methods* 25:402–408. doi:1006/meth.2001.1262
25. Gong Z, Xie Z, Qiu J, Wang G. Synthesis, Biological Evaluation and Molecular Docking Study of 2-Substituted-4,6-Diarylpyrimidines as  $\alpha$ -Glucosidase Inhibitors. *Molecules.* 2017;22(11):1865. Published 2017 Oct 30. doi:3390/molecules22111865
26. Grewal AS, Kharb R, Prasad DN, Dua JS, Lather V. Design, synthesis and evaluation of novel 3,5-disubstituted benzamide derivatives as allosteric glucokinase activators. *BMC Chem.* 2019;13(1):2. doi:1186/s13065-019-0532-8
27. Salvi GE, Kandylaki M, Troendle A, Persson GR, Lang NP. Experimental gingivitis in type 1 diabetics: a controlled clinical and microbiological study. *J Clin Periodontol.* 2005;32(3):310-6. doi:1111/j.1600-051X.2005.00682.x.

28. Twetman S, Petersson GH, Bratthall D. Caries risk assessment as a predictor of metabolic control in young Type 1 diabetics. *Diabet Med.* 2005;22(3):312-5. doi:1111/j.1464-5491.2005.01419.x.
29. Toni M, Hermida J, Toledo E, Goñi MJ, Díez Goñi N. Role of CFH and ARMS2 polymorphisms in retinopathy and coronary artery disease in type 1 diabetes. *An Sist Sanit Navar.* 2012 ;35(3):425-32. . doi:23938/ASSN.0098.
30. Drobná Z, Del Razo LM, García-Vargas GG, Sánchez-Peña LC, Barrera-Hernández A, Stýblo M, Loomis D. Environmental exposure to arsenic, AS3MT polymorphism and prevalence of diabetes in Mexico. *J Expo Sci Environ Epidemiol.* 2013 ;23(2):151-5. doi:1038/jes.2012.103.
31. Wang LA, Nguyen DH, Mifflin SW. CRHR2 (Corticotropin-Releasing Hormone Receptor 2) in the Nucleus of the Solitary Tract Contributes to Intermittent Hypoxia-Induced Hypertension. *Hypertension.* 2018;72(4):994-1001. doi:1161/HYPERTENSIONAHA.118.11497
32. Garcia-Martínez I, Sánchez-Mora C, Soler Artigas M, et al. Gene-wide Association Study Reveals RNF122 Ubiquitin Ligase as a Novel Susceptibility Gene for Attention Deficit Hyperactivity Disorder. *Sci Rep.* 2017 Jul 14;7(1):5407. doi:1038/s41598-017-05514-7.
33. Sala D, Ivanova S, Plana N, et al. Autophagy-regulating TP53INP2 mediates muscle wasting and is repressed in diabetes. *J Clin Invest.* 2014;124(5):1914-1927. doi:1172/JCI72327
34. Keshavarz P, Inoue H, Nakamura N, Yoshikawa T, Tanahashi T, Itakura M. Single nucleotide polymorphisms in genes encoding LKB1 (STK11), TORC2 (CRTC2) and AMPK alpha2-subunit (PRKAA2) and risk of type 2 diabetes. *Mol Genet Metab.* 2008;93(2):200-209. doi:1016/j.ymgme.2007.08.125
35. Granados-Riveron JT, Ghosh TK, Pope M, et al. Alpha-cardiac myosin heavy chain (MYH6) mutations affecting myofibril formation are associated with congenital heart defects. *Hum Mol Genet.* 2010;19(20):4007-4016. doi:1093/hmg/ddq315
36. McKenna K, Morris AD, Ryan M, et al. Renal resistance to vasopressin in poorly controlled type 1 diabetes mellitus. *Am J Physiol Endocrinol Metab.* 2000;279(1):E155-E160. doi:1152/ajpendo.2000.279.1.E155
37. Yue J, Zhu T, Yang J, et al. CircCBFB-mediated miR-28-5p facilitates abdominal aortic aneurysm via LYPD3 and GRIA4. *Life Sci.* 2020;253:117533. doi:1016/j.lfs.2020.117533
38. Kochetova OV, Avzaletdinova DS, Korytina GF, Morugova TV, Mustafina OE. The association between eating behavior and polymorphisms in GRIN2B, GRIK3, GRIA1 and GRIN1 genes in people with type 2 diabetes mellitus. *Mol Biol Rep.* 2020;47(3):2035-2046. doi:1007/s11033-020-05304-x
39. Jiao X, Zhang D, Hong Q, et al. Netrin-1 works with UNC5B to regulate angiogenesis in diabetic kidney disease. *Front Med.* 2020;14(3):293-304. doi:1007/s11684-019-0715-7
40. Ruiz de Azua I, Scarselli M, Rosemond E, et al. RGS4 is a negative regulator of insulin release from pancreatic beta-cells in vitro and in vivo. *Proc Natl Acad Sci U S A.* 2010;107(17):7999-8004. doi:1073/pnas.1003655107
41. Wang P, Qin W, Wang P, et al. Genomic Variants in NEURL, GJA1 and CUX2 Significantly Increase Genetic Susceptibility to Atrial Fibrillation. *Sci Rep.* 2018;8(1):3297. doi:1038/s41598-018-21611-7

42. Faienza MF, Ventura A, Delvecchio M, et al. High Sclerostin and Dickkopf-1 (DKK-1) Serum Levels in Children and Adolescents With Type 1 Diabetes Mellitus. *J Clin Endocrinol Metab.* 2017;102(4):1174-1181. doi:1210/jc.2016-2371
43. Galán M, Kassan M, Choi SK, et al. A novel role for epidermal growth factor receptor tyrosine kinase and its downstream endoplasmic reticulum stress in cardiac damage and microvascular dysfunction in type 1 diabetes mellitus. *Hypertension.* 2012;60(1):71-80. doi:1161/HYPERTENSIONAHA.112.192500
44. García SI, Porto PI, Dieuzeide G, et al. Thyrotropin-releasing hormone receptor (TRHR) gene is associated with essential hypertension. *Hypertension.* 2001;38(3 Pt 2):683-687. doi:1161/01.hyp.38.3.683
45. Feng X, Scott A, Wang Y, et al. PTPRT regulates high-fat diet-induced obesity and insulin resistance. *PLoS One.* 2014;9(6):e100783. Published 2014 Jun 20. doi:1371/journal.pone.0100783
46. Ward-Caviness CK, Neas LM, Blach C, et al. A genome-wide trans-ethnic interaction study links the PIGR-FCAMR locus to coronary atherosclerosis via interactions between genetic variants and residential exposure to traffic. *PLoS One.* 2017;12(3):e0173880. doi:1371/journal.pone.0173880
47. Sun J, Wang J, Lu W, et al. MiR-325-3p inhibits renal inflammation and fibrosis by targeting CCL19 in diabetic nephropathy. *Clin Exp Pharmacol Physiol.* 2020;10.1111/1440-1681.13371. doi:1111/1440-1681.13371
48. Yang Z, Bowles NE, Scherer SE, et al. Desmosomal dysfunction due to mutations in desmoplakin causes arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Circ Res.* 2006;99(6):646-655. doi:1161/01.RES.0000241482.19382.c6
49. Xu H, Shang Q, Chen H, et al. ITIH4: A New Potential Biomarker of "Toxin Syndrome" in Coronary Heart Disease Patient Identified with Proteomic Method. *Evid Based Complement Alternat Med.* 2013;2013:360149. doi:1155/2013/360149
50. Miyashita A, Arai H, Asada T, et al. GAB2 is not associated with late-onset Alzheimer's disease in Japanese. *Eur J Hum Genet.* 2009;17(5):682-686. doi:1038/ejhg.2008.181
51. McCann JA, Xu YQ, Frechette R, Guazzarotti L, Polychronakos C. The insulin-like growth factor-II receptor gene is associated with type 1 diabetes: evidence of a maternal effect. *J Clin Endocrinol Metab.* 2004;89(11):5700-5706. doi:1210/jc.2004-0553
52. Sun L, Lin P, Chen Y, et al. miR-182-3p/Myadm contribute to pulmonary artery hypertension vascular remodeling via a KLF4/p21-dependent mechanism. *Theranostics.* 2020;10(12):5581-5599. doi:7150/thno.44687
53. Fan Y, Li X, Zhang Y, et al. Genetic Variants of TPCN2 Associated with Type 2 Diabetes Risk in the Chinese Population. *PLoS One.* 2016;11(2):e0149614. doi:1371/journal.pone.0149614
54. Chan GC, Divers J, Russell GB, et al. FGF23 Concentration and APOL1 Genotype Are Novel Predictors of Mortality in African Americans With Type 2 Diabetes. *Diabetes Care.* 2018;41(1):178-186. doi:2337/dc17-0820

55. Hermansson C, Lundqvist A, Wasslavik C, Palmqvist L, Jeppsson A, Hultén LM. Reduced expression of NLRP3 and MEFV in human ischemic heart tissue. *Biochem Biophys Res Commun.* 2013;430(1):425-428. doi:10.1016/j.bbrc.2012.11.070
56. Wang Z, Dai Z, Pan Y, Wu S, Li Z, Zuo C. E3 ubiquitin ligase DTX4 is required for adipogenic differentiation in 3T3-L1 preadipocytes cell line. *Biochem Biophys Res Commun.* 2017;492(3):419-424. doi:10.1016/j.bbrc.2017.08.083
57. Baptista FI, Pinto MJ, Elvas F, Almeida RD, Ambrósio AF. Diabetes alters KIF1A and KIF5B motor proteins in the hippocampus. *PLoS One.* 2013;8(6):e65515. doi:10.1371/journal.pone.0065515
58. Parackova Z, Vrabcova P, Zentsova I, et al. Enhanced STAT3 phosphorylation and PD-L1 expression in myeloid dendritic cells indicate impaired IL-27/R $\alpha$  signaling in type 1 diabetes. *Sci Rep.* 2020;10(1):493. doi:10.1038/s41598-020-57507-8
59. Muiya NP, Wakil S, Al-Najai M, et al. A study of the role of GATA2 gene polymorphism in coronary artery disease risk traits. *Gene.* 2014;544(2):152-158. doi:10.1016/j.gene.2014.04.064

## Tables

Due to technical limitations, Tables 1-7 are only available as a download in the supplementary files section.

## Figures

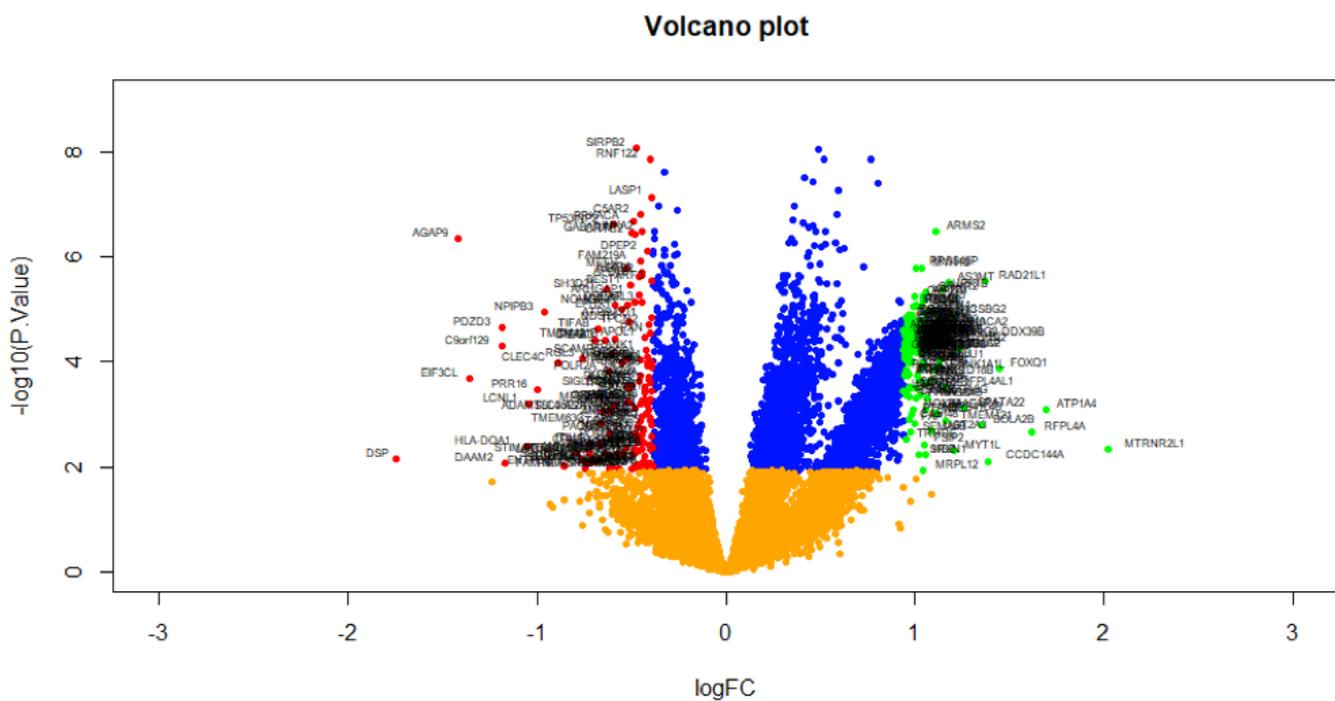


Figure 1

Volcano plot of differentially expressed genes. Genes with a significant change of more than two-fold were selected. Green dot represented up regulated significant genes and red dot represented down regulated significant genes.

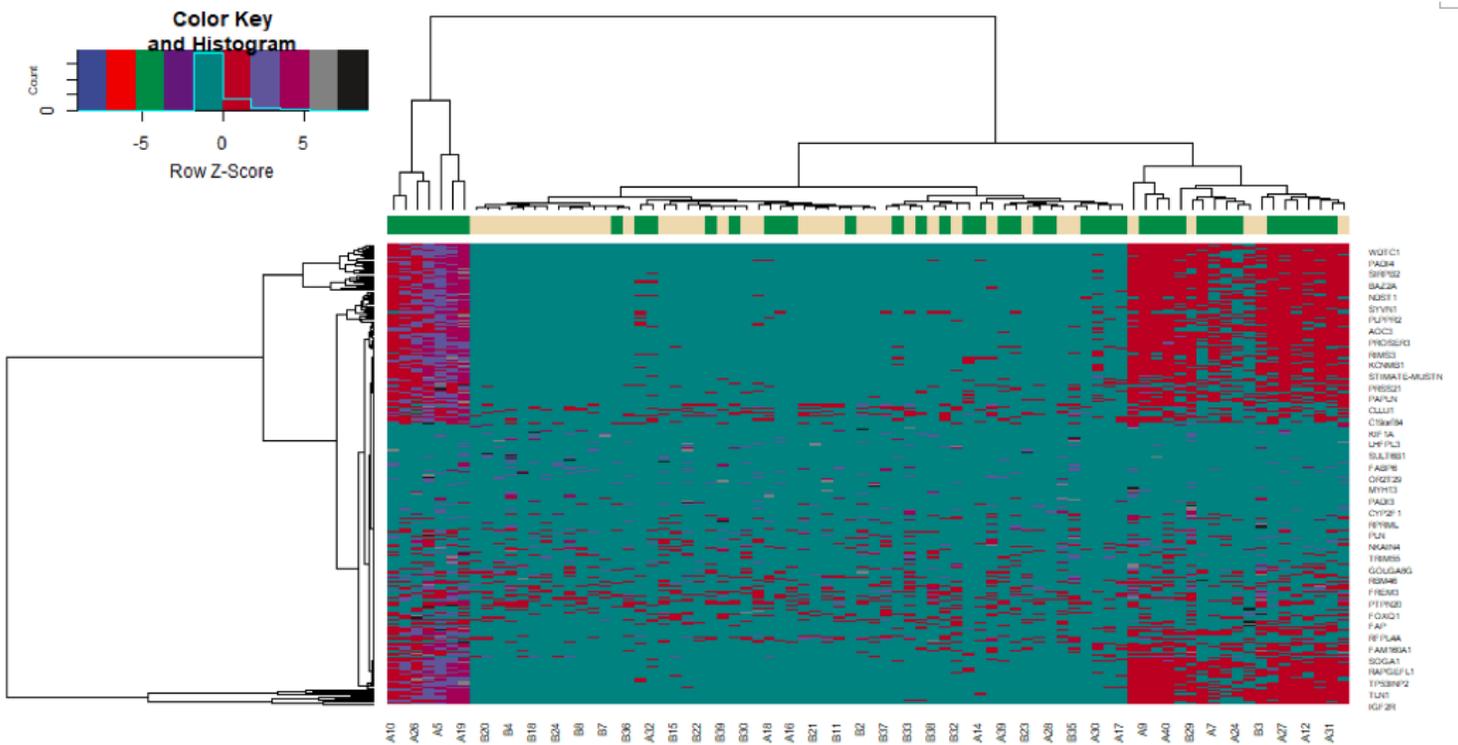
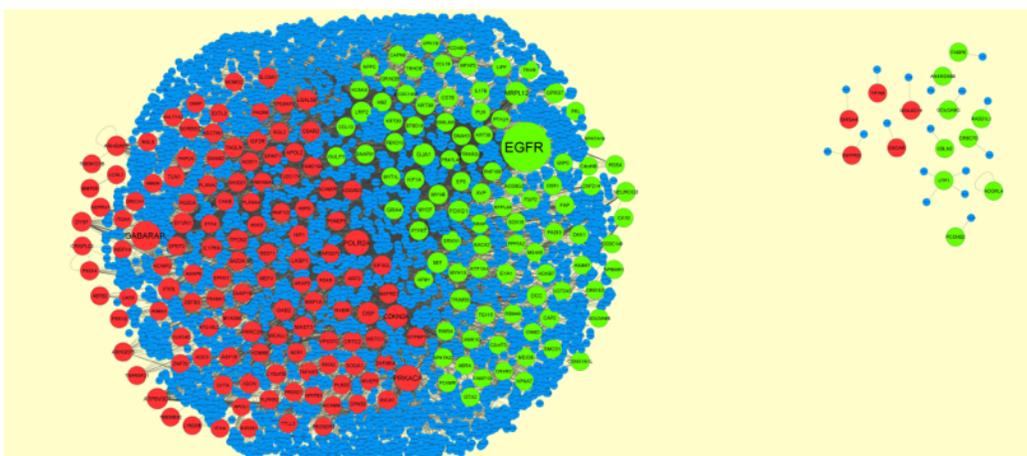


Figure 2

Heat map of differentially expressed genes. Legend on the top left indicate log fold change of genes. (A1 – A43 = healthy donors' samples; B1 – B14 = T1D samples)

A



B

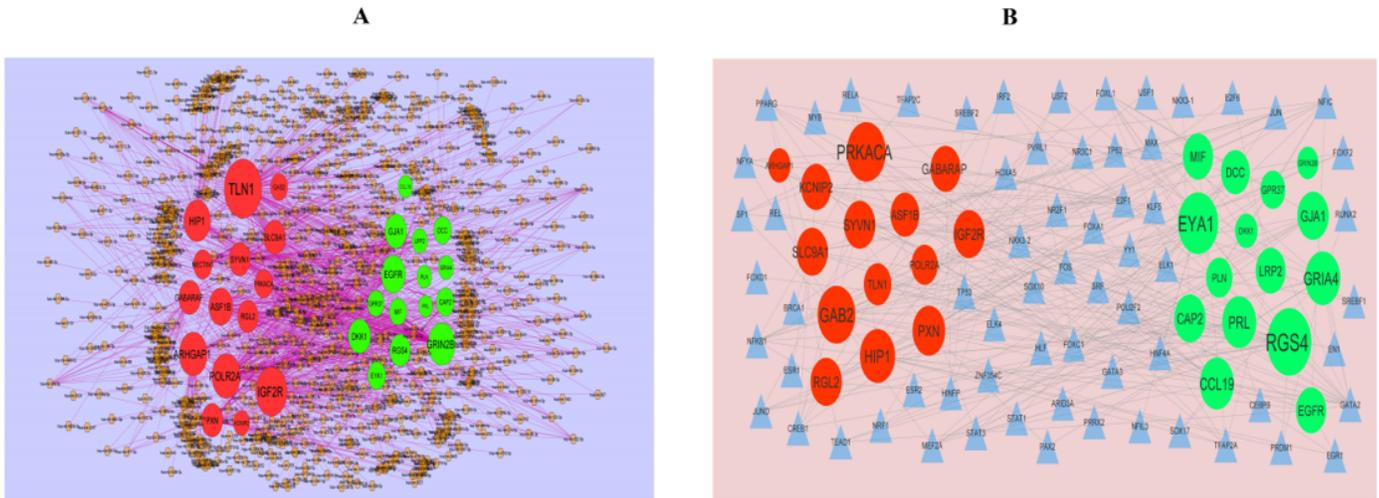


C



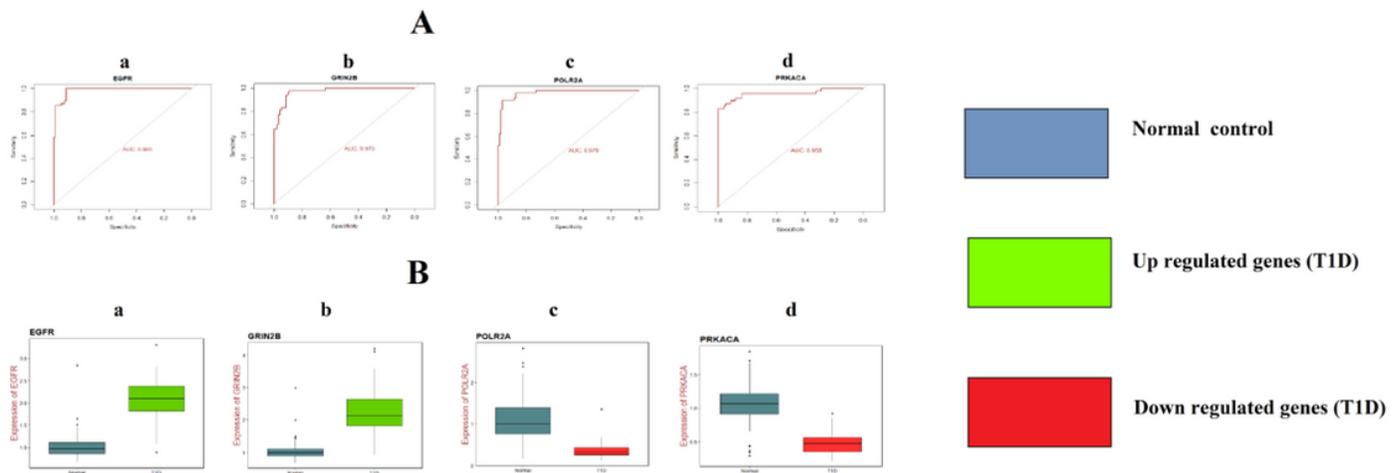
Figure 3

PPI network and the most significant modules of DEGs. (A) The PPI network of DEGs was constructed using Cytoscape (B) The most significant module was obtained from PPI network with 12 nodes and 23 edges for up regulated genes (C) The most significant module was obtained from PPI network with 5 nodes and 10 edges for up regulated genes. Up regulated genes are marked in green; down regulated genes are marked in red



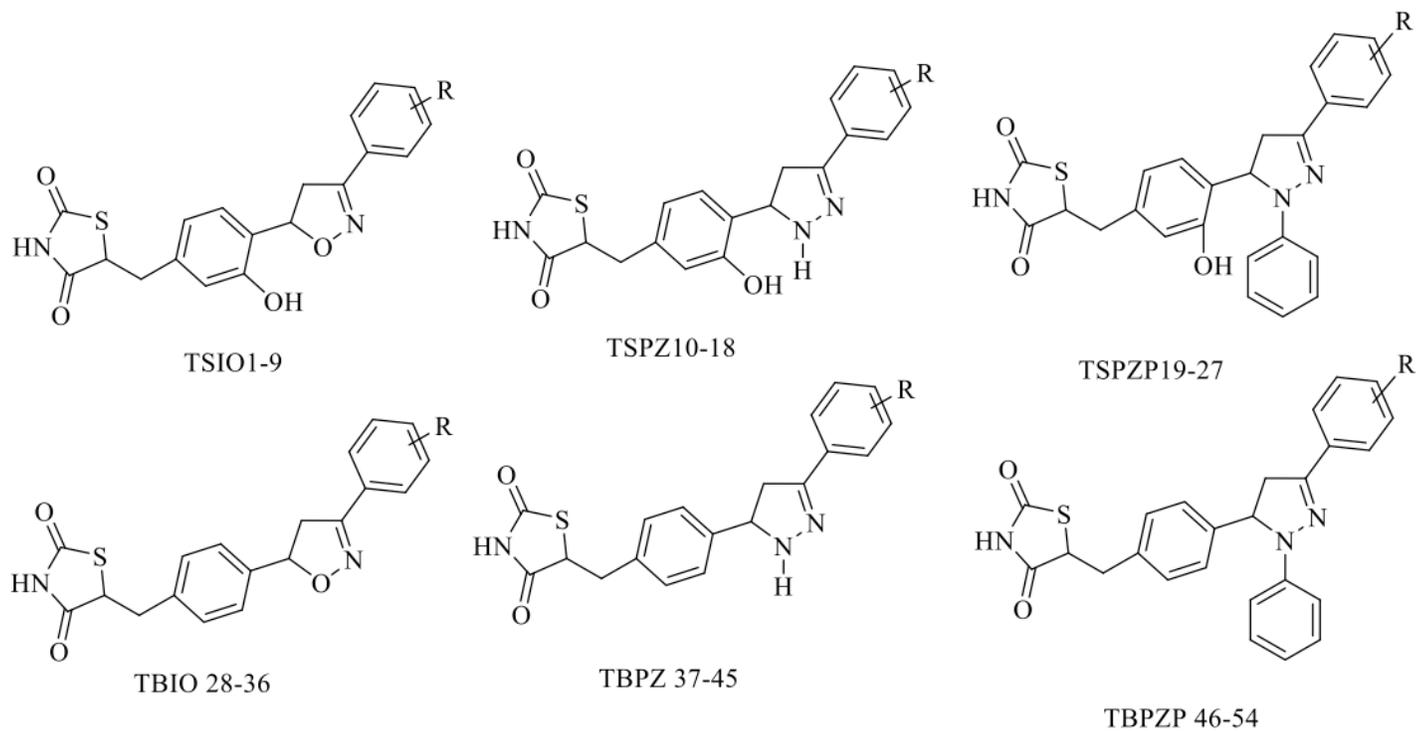
**Figure 4**

(A) Target gene - miRNA regulatory network between target genes and miRNAs (B) Target gene - TF regulatory network between target genes and TFs. Up regulated genes are marked in green; down regulated genes are marked in red; The chocolate color triangle nodes represent the key miRNAs; the blue color triangle nodes represent the key miRNAs



**Figure 5**

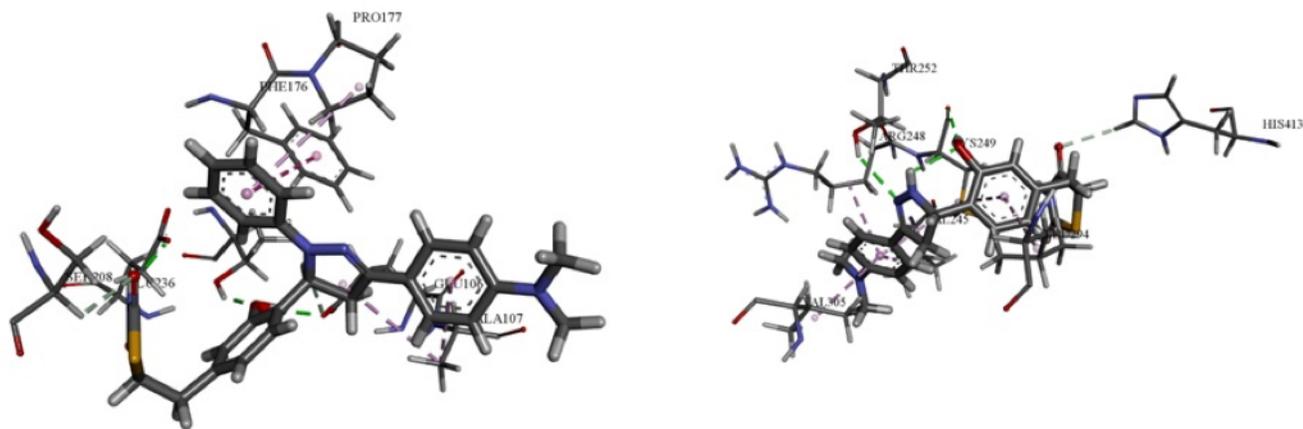
(A) ROC curve and (B) RT-PCR analyses of hub genes. a) EGFR, b) GRIN2B, c) POLR2A d) PRKACA



Aldehydes R= 1: p-Cl 2: p-Br 3: p-OH 4: p-F 5: p-I 6: p-N(CH<sub>3</sub>)<sub>2</sub> 7: p-OCH<sub>3</sub> 8: p-NH<sub>2</sub> 9: H

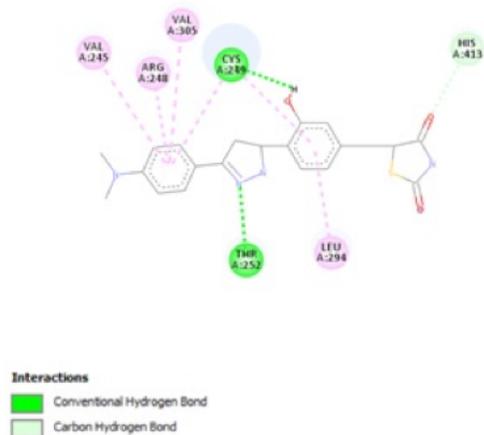
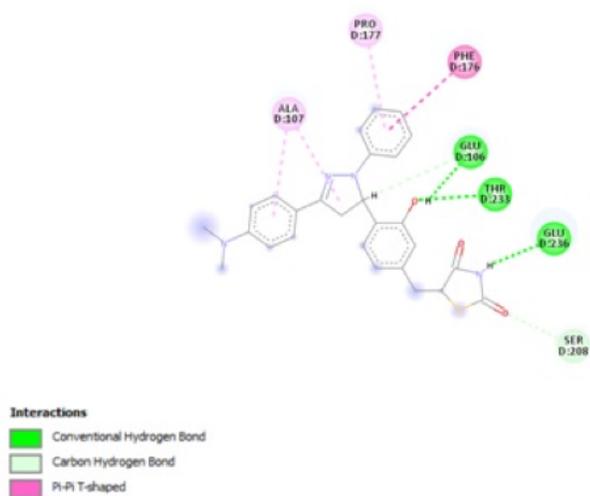
**Figure 6**

Structures of Designed Molecules



**Figure 7**

3D Binding of Molecule PIO&TSPZP24with 2XYJ&6E7V



**Figure 8**

2D Binding of Molecule SALISO 2&SALPYR 12with 6E7V&2XYJ

## Supplementary Files

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

- [Tables.docx](#)