



Prioritizing Candidate Genes for Type 2 Diabetes Mellitus using Integrated Network and Pathway Analysis

Tejaswini Prakash and Nallur B Ramachandra *

Genetics and Genomics Lab, Department of Studies in Genetics and Genomics, University of Mysore, Manasagangothri, Mysuru – 570 006, Karnataka, India

Abstract

Background: Type 2 Diabetes Mellitus (T2DM) has emerged as a major threat to global health that fosters life-threatening clinical complications, taking a huge toll on our society. More than 65 million Indians suffer from T2DM, making it one of the leading causes of death. T2DM and associated complications have to be constantly monitored and managed which reduces the overall quality of life and increases socio-economic burden. Therefore, it is crucial to develop specific treatment and management strategies. In order to achieve this, it is essential to understand the underlying genetic causes and molecular mechanisms.

Methods: Integrated gene network and ontology analyses facilitate prioritization of plausible candidate genes for T2DM and also aid in understanding their mechanistic pathways. In this study, T2DM-associated genes were subjected to sequential interaction network and gene set enrichment analysis. High ranking network clusters were derived and their interrelation with pathways was assessed.

Results: About 23 significant candidate genes were prioritized from 615 T2DM-associated genes which were overrepresented in pathways related to insulin resistance, type 2 diabetes, signaling cascades such as insulin receptor signaling pathway, PI3K signaling, IGF1R signaling pathway, ERBB signaling pathway, MAPK signaling pathway and their regulatory mechanisms.

Conclusion: Of these, two tyrosine kinase receptor genes-*EGFR* and *IGF1R* were identified as common nodes and can be considered to be significant candidate genes in T2DM.

Avicenna J Med Biotech 2022; 14(3): 239-246

Keywords: Gene ontology, Hub genes identification, *In silico* analysis, Text mining, Type 2 diabetes mellitus

Introduction

Type 2 Diabetes Mellitus (T2DM) is a major disorder of the metabolic and endocrine system which has gained the status of "global epidemic" owing to the increasing number of diagnosed cases worldwide ¹. According to the 2019 report by International Diabetes Federation (IDF), 463 million people in the age group of 20-64 years were diagnosed with T2DM worldwide, which is approximately 8.8% of the adult population. This is further predicted to increase by 60-70% by 2030. In addition, 374 million individuals are said to be in pre-diabetic stage, displaying Impaired Glucose Tolerance (IGT). Each year, 3.7 million deaths occur as a consequence of T2DM and high blood glucose, making it the 8th leading cause of death. Globally, India ranks second with 77 million individuals diagnosed with T2DM. It is further predicted to increase to 101 million

by 2030 and 134.2 million by 2045 ². Besides being an important health concern on its own, T2DM can also lead to severe complications of the vascular, renal and ophthalmic systems manifesting as Coronary Artery Diseases (CAD), Peripheral Artery Disease (PAD), diabetic nephropathy, diabetic retinopathy and diabetic neuropathy. In view of these debilitating consequences, immense efforts have been made towards understanding the molecular mechanisms of T2DM using high-throughput technologies, paving way for novel treatment and disease-management strategies. Currently, genome and transcriptome sequencing with bioinformatics analyses have been used widely to successfully identify numerous genes and functional pathways associated with pathophysiology of T2DM. However, defining the causal molecular players and their mecha-

nism of action from the many identified gene sets is challenging and time consuming. *In silico* analysis using interaction network approach provides a practical solution to reduce noise and prioritize significant candidate genes for functional experiments.

System biology and protein interaction network approaches are being increasingly used to study complex diseases such as T2DM, CAD, Autism Spectrum Disorders (ASD) and cancers³⁻⁸. The foundation for these approaches is the knowledge that protein molecules interact with each other to perform particular functions that are generally represented as interaction networks. Ding *et al*⁹ used differential gene expression profile of pancreatic β -cells integrated with Protein-Protein Interaction (PPI) network and functional analysis to identify *HNF1A*, *STAT3*, *SERPING1*, *ANPEP* and Glucocorticoid Receptor (GR) as crucial genes involved in the development of T2DM. Similarly, Lin *et al*¹⁰ identified 36 differentially expressed genes between diabetic and non-diabetic pancreatic islets that were significantly enriched in PI3K-Akt signaling pathway, pathways in cancer, cytokine-cytokine receptor interaction and Rheumatoid arthritis. Furthermore, 10 important hub genes were established of which *IL6*, *MMP3*, *MMP1* and *IL11* were regarded potential biomarkers for T2DM diagnosis. Vastrad *et al*¹¹ utilized expression profiles of T2DM and normal control datasets along with PPI network and functional enrichment to identify plausible candidates for T2DM and their mechanistic pathways. The study also predicts potential miRNA and transcription factors involved in regulatory functions, identifying *JUN*, *RELA*, *U2AF2*, *VCAM1*, *FNI*, *CDK1*, *ADRB2*, *TK1*, *ACTA2* and *A2M* as potential biomarkers for T2DM. In a similar study utilizing expression profiles of diabetic obese and non-diabetic obese samples, Vastrad *et al*¹² also identified key genes such as *FLNA*, *MYH9*, *CLTC*, *ERBB2*, *DCTN1*, *TCF4*, *VIM*, *LRRK2*, *CAVI* and *IFI16* as novel, probable biomarkers for progression of obesity associated T2DM.

There is abundant literature on integrating differential gene expression profiles with network analysis and functional enrichment to describe molecular mechanisms of T2DM. However, there are numerous T2DM-associated genes identified through candidate gene studies, Genome Wide Association Studies (GWAS) and Exome/Genome Sequencing (WES/WGS). Studies aimed at leveraging these T2DM-associated genes to establish their mechanistic roles in disease development and progression would be insightful. Prioritization of potential candidates from established disease-associated genes using network based approaches is rapidly gaining popularity as it provides time and cost efficient alternative to biological validation of all the T2DM-associated genes identified thus far¹³⁻¹⁵. In view of this, we employed integrated network analysis, gene set enrichment and pathway interrelation approaches to discern possible candidates from T2DM-associated genes.

Materials and Methods

Compilation of gene set

T2DM associated genes were extracted from two publicly available databases: Type 2 Diabetes Knowledge Portal by the Genetics of Type 2 Diabetes Consortium (GoT2D)^{16,17} and T-HOD database¹⁸. T2D knowledge portal lists predicted effector genes for T2DM using a heuristic scoring method based on genetic, regulatory and perturbation evidences derived from large genomic studies. Genes predicted to have "causal", "strong" and "moderate" link to T2DM risk were considered for present study. T-HOD is a literature based database that provides information on candidate genes for hypertension, obesity and diabetes. Additionally, DisGeNET¹⁹ was also used to retrieve curated genes for T2DM. Genes from all three sources were collated and a list of non-redundant, unique genes was compiled.

Construction of protein-protein interaction network and hub genes identification

PPI network was constructed using STRING v11.0 (Search Tool for the Retrieval of Interacting Genes/Proteins) (<https://string-db.org/>)²⁰. The compiled list of genes was used as input with *Homo sapiens* as the organism of interest. Experiments, databases and co-expression were used as active interaction sources, with a median confidence score of >0.4. Rest of the parameters were set to default. The resultant protein interaction network was visualized in Cytoscape v3.8.0.0²¹. Significant hub genes from the large network were identified using cytoHubba application²². Top 10 genes evaluated employing five topological analysis methods namely, Degree, MCC, MNC, EPC and Ec-Centricity were selected for further analysis. The intersecting genes derived using these five algorithms encodes core proteins and may represent key candidate genes with important biological functions.

Gene set enrichment and pathway interrelation analysis

Functional annotation was performed using the ontologies and pathways function on Enrichr²³⁻²⁵ to categorize the top scoring genes into biological processes, molecular functions, cellular components and pathways. Enrichment terms relevant to T2DM and associated phenotypes with a significant p-value of <0.05 were considered. ClueGO²⁶ was used on Cytoscape for pathway interrelation analysis using KEGG annotations. Right-sided hypergeometric test with Benjamini-Hochberg method for p-value correction was employed.

Results

Compilation of gene set

The T-HOD database yielded a total of 497 T2DM specific genes while 79 predicted effector genes were extracted from GoT2D knowledge portal. A total of 160 curated genes were retrieved from DisGeNET database. All three gene sets were combined and dupli-

cate gene symbols were removed to obtain 615 unique, non-redundant genes used to build PPI network.

Construction of protein-protein interaction network and hub genes identification

PPI network constructed from 615 seed genes on STRING database yielded a complex network of 762 interactions among 605 nodes (Figure 1). The resultant edges were found to be more than the expected interactions indicating that the queried proteins are at least partially biologically connected as a group. Importing the PPI network to Cytoscape and analyzing using de-

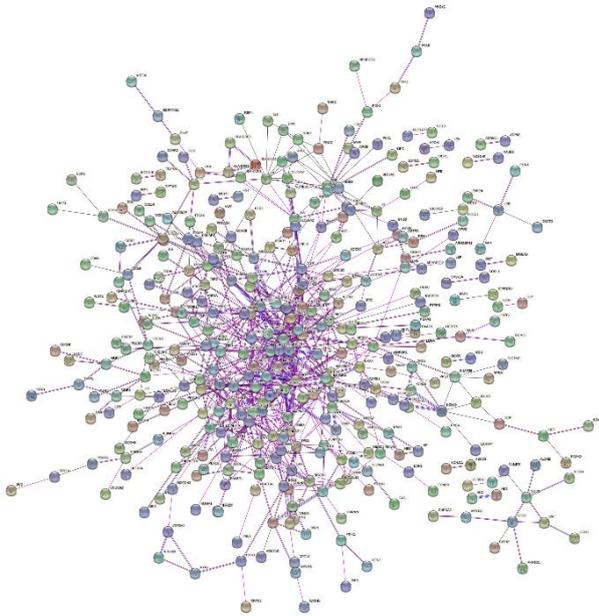


Figure 1. Protein interaction network constructed on STRING database using T2DM-associated genes.

gree, MCC, MNC, EPC and EcCentricity methods on cytoHubba resulted in identification of 23 top scoring hub genes (Figure 2A-E). *IGF1R* and *EGFR* were observed to be intersections of the five algorithms as shown in the Venn plot (Figure 3) and may serve as possible candidates for further studies.

Gene set enrichment and pathway interrelation analysis

Functional association of the top scoring hub genes with biological processes, molecular function, cellular components and pathways performed on Enrichr revealed significant enrichment in pathways for insulin resistance, type 2 diabetes, signaling cascades such as insulin receptor signaling pathway, PI3K signaling, IGF1R signaling pathway, ERBB signaling pathway, MAPK signaling pathway and their regulatory mechanisms (Table 1). Additionally, the hub genes were also overrepresented in pathways related to vascular complications associated with diabetes such as diabetic cardiomyopathy, AGE-RAGE signaling pathway in diabetic complications, fluid shear stress and atherosclerosis, cellular response to lipid and atherosclerosis (Table 2). The interrelation between these pathways and 23 hub genes identified was explored using Cytoscape ClueGO and is depicted in figure 4.

Discussion

The steady increase in prevalence of T2DM over the last few decades has led to diverse efforts towards delineating its molecular mechanism and pathophysiology. Numerous causal and susceptibility genes associated with T2DM have been identified by several genetic studies over the years. However, owing to the complex etiology of T2DM, unraveling the precise molecular pathways involved in disease development and pro-

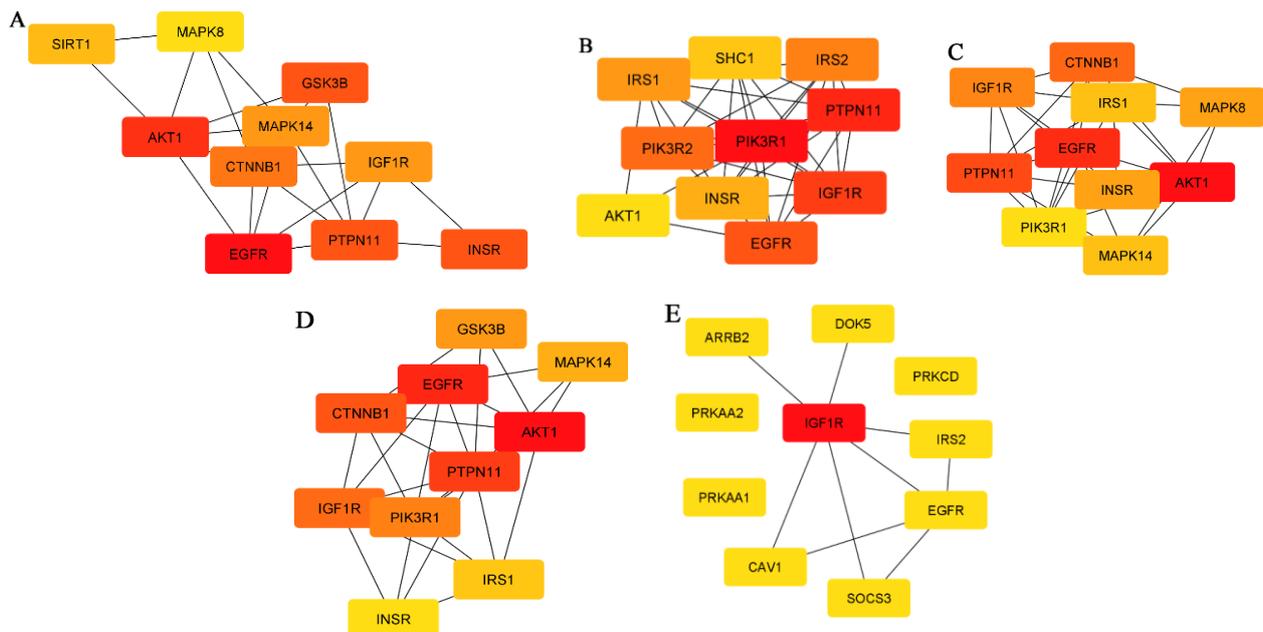


Figure 2. Top ranking gene clusters derived from cytoHubba. Algorithms used: A) Degree, B) MCC, C) MNC, D) EPC, E) EcCentricity.

Network Analysis for T2DM Candidate Genes

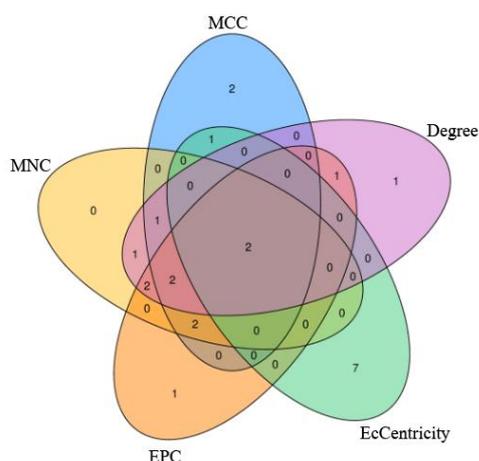


Figure 3. Venn plot depicting intersection of all five algorithms used on cytoHubba, with each algorithm represented by different colours: dark blue-MCC; purple-Degree; light blue-EcCentricity; orange-EPC; yellow-MNC. Number of genes found in common across algorithms is indicated. Two hub genes, *EGFR* and *IGF1R*, are found to be common across all the algorithms.

gression has been challenging. In this view, complex diseases are progressively being studied using systemic network approaches combined with functional analysis to decipher the mechanism of action of genes involved [5,7].

In recent years, gene ontology and network based approaches have been employed extensively for T2DM biomarker identification using enormous data from linkage studies, GWAS and more recently-NGS based genome and transcriptome studies. PPI networks constructed using genes associated with T2DM and T2DM-related phenotypes are used to identify hub genes that possibly have a role to play in disease pathophysiology. In this study, we implemented integrated PPI network and hub gene analysis on 615 T2DM associated genes, along with gene set enrichment and pathway interrelation analysis to identify 23 genes namely, *AKT1*, *ARRB2*, *CAVI*, *CTNNB1*, *DOK5*, *EGFR*, *ESR1*, *GSK3B*, *IGF1R*, *INSR*, *IRS1*, *IRS2*, *MAPK14*, *MAPK8*, *PIK3R1*, *PIK3R2*, *PRKAA1*, *PRKAA2*, *PRKCD*, *PTPN11*, *SHC1*, *SIRT1* and *SOCS3* that are plausible high-risk

Table 1: Gene set enrichment and pathway analysis of 23 hub genes identified. Gene ontology categories: pathways, biological processes, molecular function, cellular component

| GO term | p-value | Adj p-value | Associated genes |
|---|----------|-------------|---|
| Pathway | | | |
| Insulin resistance | 1.72E-24 | 2.73E-22 | <i>GSK3B</i> ; <i>PRKAA1</i> ; <i>PRKAA2</i> ; <i>IRS1</i> ; <i>INSR</i> ; <i>PRKCD</i> ; <i>IRS2</i> ; <i>PIK3R2</i> ; <i>PTPN11</i> ; <i>PIK3R1</i> ; <i>SOCS3</i> ; <i>MAPK8</i> ; <i>AKT1</i> |
| Insulin signaling pathway | 8.27E-21 | 4.38E-19 | <i>GSK3B</i> ; <i>SOCS3</i> ; <i>PRKAA1</i> ; <i>MAPK8</i> ; <i>PRKAA2</i> ; <i>IRS1</i> ; <i>SHC1</i> ; <i>INSR</i> ; <i>AKT1</i> ; <i>IRS2</i> ; <i>PIK3R2</i> ; <i>PIK3R1</i> |
| Type II diabetes mellitus | 1.97E-16 | 2.84E-15 | <i>SOCS3</i> ; <i>MAPK8</i> ; <i>IRS1</i> ; <i>INSR</i> ; <i>PRKCD</i> ; <i>IRS2</i> ; <i>PIK3R2</i> ; <i>PIK3R1</i> |
| PI3K-AKT signaling pathway | 2.48E-12 | 1.88E-11 | <i>GSK3B</i> ; <i>PRKAA1</i> ; <i>PRKAA2</i> ; <i>IRS1</i> ; <i>INSR</i> ; <i>AKT1</i> ; <i>PIK3R2</i> ; <i>PIK3R1</i> ; <i>EGFR</i> ; <i>IGF1R</i> |
| MAPK signaling pathway | 2.77E-08 | 7.59E-08 | <i>MAPK8</i> ; <i>INSR</i> ; <i>AKT1</i> ; <i>ARRB2</i> ; <i>MAPK14</i> ; <i>EGFR</i> ; <i>IGF1R</i> |
| Biological process | | | |
| Regulation of protein kinase B signaling | 1.34E-18 | 1.46E-15 | <i>IRS1</i> ; <i>INSR</i> ; <i>AKT1</i> ; <i>IRS2</i> ; <i>PIK3R2</i> ; <i>PTPN11</i> ; <i>PIK3R1</i> ; <i>ARRB2</i> ; <i>ESR1</i> ; <i>SIRT1</i> ; <i>EGFR</i> ; <i>IGF1R</i> |
| Insulin receptor signaling pathway | 5.41E-17 | 2.94E-14 | <i>GSK3B</i> ; <i>IRS1</i> ; <i>SHC1</i> ; <i>INSR</i> ; <i>AKT1</i> ; <i>IRS2</i> ; <i>PIK3R2</i> ; <i>PIK3R1</i> ; <i>IGF1R</i> |
| Cellular response to insulin stimulus | 1.10E-14 | 2.39E-12 | <i>GSK3B</i> ; <i>IRS1</i> ; <i>SHC1</i> ; <i>INSR</i> ; <i>AKT1</i> ; <i>IRS2</i> ; <i>PIK3R2</i> ; <i>PIK3R1</i> ; <i>IGF1R</i> |
| Phosphatidylinositol 3-kinase signaling | 3.44E-10 | 1.56E-08 | <i>IRS1</i> ; <i>AKT1</i> ; <i>PIK3R2</i> ; <i>PIK3R1</i> ; <i>IGF1R</i> |
| Insulin-like growth factor receptor signaling pathway | 4.36E-10 | 1.89E-08 | <i>IRS1</i> ; <i>AKT1</i> ; <i>PIK3R1</i> ; <i>IGF1R</i> |
| Regulation of phosphatidylinositol 3-kinase signaling | 1.80E-09 | 6.75E-08 | <i>INSR</i> ; <i>IRS2</i> ; <i>PIK3R1</i> ; <i>SIRT1</i> ; <i>EGFR</i> ; <i>IGF1R</i> |
| ERBB signaling pathway | 1.08E-08 | 3.25E-07 | <i>SHC1</i> ; <i>AKT1</i> ; <i>PTPN11</i> ; <i>PIK3R1</i> ; <i>EGFR</i> |
| Regulation of insulin receptor signaling pathway | 1.92E-07 | 3.79E-06 | <i>IRS1</i> ; <i>PRKCD</i> ; <i>PTPN11</i> ; <i>SIRT1</i> |
| Molecular function | | | |
| Insulin-like growth factor receptor binding | 1.32E-09 | 4.75E-08 | <i>IRS1</i> ; <i>SHC1</i> ; <i>INSR</i> ; <i>PIK3R1</i> |
| Protein serine/threonine kinase activity | 2.76E-09 | 7.46E-08 | <i>GSK3B</i> ; <i>PRKAA1</i> ; <i>MAPK8</i> ; <i>PRKAA2</i> ; <i>PRKCD</i> ; <i>AKT1</i> ; <i>MAPK14</i> ; <i>EGFR</i> |
| Protein kinase binding | 5.57E-08 | 1.00E-06 | <i>GSK3B</i> ; <i>IRS1</i> ; <i>SHC1</i> ; <i>CAVI</i> ; <i>PRKCD</i> ; <i>PTPN11</i> ; <i>ESR1</i> ; <i>SIRT1</i> |
| Insulin-like growth factor binding | 1.32E-04 | 5.47E-04 | <i>INSR</i> ; <i>IGF1R</i> |
| Cellular component | | | |
| Protein kinase complex | 4.64E-08 | 2.60E-06 | <i>IRS1</i> ; <i>INSR</i> ; <i>IGF1R</i> |

Table 2. Enriched pathways related to vascular complications

| GO term | p-value | Adj p-value | Associated genes |
|--|-----------------|-----------------|--|
| Pathway | | | |
| Fluid shear stress and atherosclerosis | 2.19E-14 | 2.32E-13 | <i>PRKAA1; MAPK8; PRKAA2; CAV1; AKT1; CTNNB1; PIK3R2; PIK3R1; MAPK14</i> |
| Diabetic cardiomyopathy | 6.91E-13 | 6.11E-12 | <i>GSK3B; MAPK8; IRS1; INSR; PRKCD; AKT1; PIK3R2; PIK3R1; MAPK14</i> |
| AGE-RAGE signaling pathway in diabetic complications | 1.26E-09 | 5.03E-09 | <i>MAPK8; PRKCD; AKT1; PIK3R2; PIK3R1; MAPK14</i> |
| Biological process | | | |
| <i>Cellular response to lipid</i> | <i>4.31E-06</i> | <i>4.63E-05</i> | <i>MAPK8; IRS1; MAPK14; ESR1; EGFR</i> |

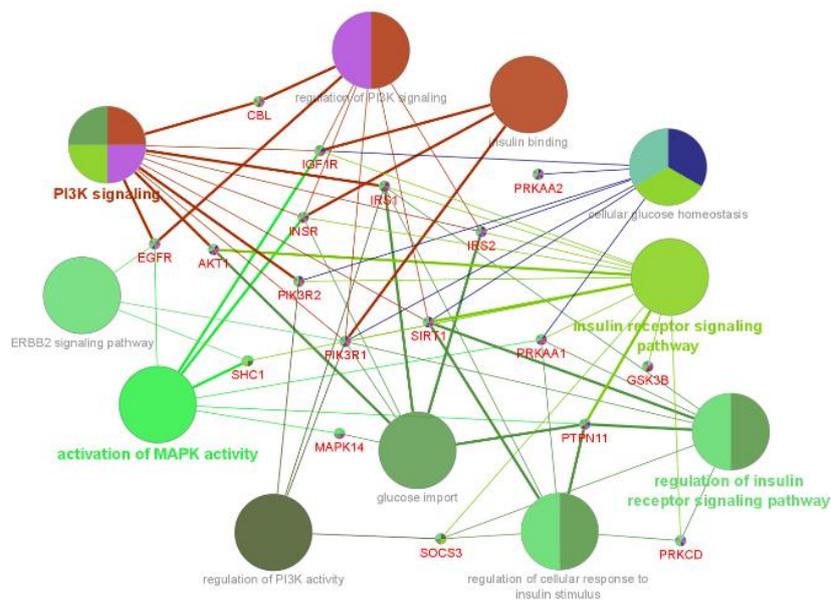


Figure 4. Pathway interrelation analysis of genes derived from top ranking gene networks. The enriched pathways are represented by larger nodes, and genes by smaller nodes. The corresponding edges indicate crosstalk between the enriched pathways and genes.

candidates for T2DM.

Of these, two tyrosine kinase receptor genes-*EGFR* and *IGF1R* were identified as common nodes by all five algorithms of cytoHubba and can be considered to be significant candidate genes in T2DM. Epidermal Growth Factor Receptor (EGFR) is a transmembrane glycoprotein that holds a key role as signaling molecule in the regulation of functions such as cellular growth, migration, proliferation and survival. Abnormal expression and signaling of EGFR is associated with diabetes, cardiovascular disease and cancer²⁷. There is growing body of research that suggests EGFR may play dual roles in the pancreas wherein it is integral for the regulation of pancreatic β cell mass and also mediate development of detrimental pathologies such as pancreatic fibrosis²⁸. Elevated EGFR signaling and increased receptor tyrosine kinase activity results in the dysfunction of diabetic micro-and macro-vasculature^{29,30}. EGFR also mediates detrimental pathways involving Endoplasmic Reticulum (ER)-stress,

Reactive Oxygen Species (ROS) and Renin-Angiotensin-Aldosterone-System (RAAS) systems, leading to diabetic cardiac myopathy. On the other hand, EGFR is believed to be beneficial in the non-diabetic heart, with key roles in cardiac development and survival of adult myocardium³¹. Insulin-like Growth Factor-1 Receptor (IGF1R) facilitates the action of Insulin-like Growth Factor-1 (IGF1) by activating the PI3K/AKT and MAPK signaling pathways. IGF1R signaling is important in several physiological functions during cellular growth, differentiation and survival in normal and diseased conditions such as cancer and diabetes-induced cardiovascular dysfunction³². IGF1 has been demonstrated by several independent studies to have a role in diabetic vascular diseases such as, atherosclerosis, angiogenesis, hypertension and restenosis³³. In diabetic conditions, Advanced Glycation End-products (AGEs) are formed as a result of prolonged hyperglycemia. Vascular proliferative changes that are induced due to hyperglycemia are reported in human mono-

cytes as a consequence of AGE-induced IGF1 synthesis³⁴. Studies on diabetic APOE-null mice have demonstrated that AGEs promote atherosclerotic lesion formation by enhancing proinflammatory pathways³⁵. Two distinct hallmarks of atherosclerosis—Vascular Smooth Muscle Cell (VSMC) differentiation and migration leading to atherosclerotic plaque formation, and destabilization and rupture of the plaque as a result of VSMC apoptosis are also intricately governed by differential expression of IGF1 and IGF1R³⁶⁻³⁸.

Pathway interrelation analysis on ClueGO identified crosstalk between hub genes and molecular mechanisms which are generally known to underlie T2DM pathogenesis. Functional enrichment analyses on Enrichr identified several crucial processes and signaling pathways that are key factors for disease development. In particular, over-representation was seen in biological processes related to insulin receptor signaling, PI3K signaling, response to insulin stimulus, IGFR-signaling process and their regulation; molecular functions such as IGFR binding, serine/threonine kinase activity and protein kinase binding.

The past few decades have seen abundant efforts towards understanding the role of these signaling pathways in T2DM. Circulating insulin, insulin and insulin-like growth factor receptors, and their corresponding signal transduction hold pivotal roles in glucose and energy metabolism. Under physiological conditions, insulin signaling cascade gets triggered by the binding of insulin to its receptors (IR), which undergo autophosphorylation and recruit IR substrates. Downstream signaling partners are then activated through kinases such as FOXO, AMPK, GSK3, ERK, mTOR, PI3K and AKT³⁹⁻⁴². Dysfunction of these pathways could lead to abnormal glucose homeostasis and subsequent insulin resistance, which is the primary cause of T2DM.

Of these, PI3K plays an important role in insulin action by activating the PI3K/AKT cascade. The PI3K/AKT signaling pathway gets activated by postprandial insulin secretion, leading to increased utilization of glucose and reduced gluconeogenesis in muscle and liver. Improper functioning of PI3K/AKT insulin signaling pathway results in insulin resistance, leading to reduced functioning of β -cells⁴³. Enrichment analysis also revealed significant enrichment of genes in pathways such as diabetic cardiomyopathy, AGE-RAGE signaling pathway in diabetic complications, cellular response to lipid, fluid shear stress and atherosclerosis indicating possible involvement of these genes in development of diabetic cardiovascular complications. In our previous work⁴⁴, we have described in detail the mechanistic action of these processes leading to the development of coronary artery disease. Thus, as established by pathway interrelation analysis, a complex interplay of genes is involved in the functioning of these pathways. Therefore, any disruptive change in these genes and pathways may trigger a series of

pathological outcomes.

Conclusion

In the current study, we employed sequential network and pathway analysis to identify 23 plausible high risk candidate genes for T2DM and its involvement in functional processes leading to disease development. Two genes, *EGFR* and *IGF1R* can be considered significant candidates for further exploration and understanding with respect to T2DM. Thus, it can be derived that the onset of T2DM and its progression might be a consequence of disrupted downstream processes caused by any perturbations in the normal functioning of these identified genes. The study is aided by ontology based gene filtration and prioritization. However, these results serve as preliminary data for subsequent research on T2DM etiology. As *in silico* analyses alone cannot provide substantial evidence, these results need to be validated through appropriate wet-lab experiments.

Acknowledgement

We are grateful to Department of Studies in Genetics and Genomics, University of Mysore for providing facility to conduct this work. We also thank members of the Genetics and Genomics lab, Department of Studies in Genetics and Genomics for their support and encouragement. We thank Indian Council of Medical Research (ICMR) for providing fellowship to Mrs. Tejaswini Prakash.

Conflict of Interest

The authors declare no competing financial interests.

References

1. World Health Organization. Global Report on Diabetes. WHO Press 2016.
2. International Diabetes Federation. IDF Diabetes Atlas, 9th edition. 2019.
3. Doncheva NT, Kacprowski T, Albrecht M. Recent approaches to the prioritization of candidate disease genes. *Wiley Interdiscip Rev Syst Biol Med* 2012;4(5):429-42.
4. Tan X, Zhang X, Pan L, Tian X, Dong P. Identification of key pathways and genes in advanced coronary atherosclerosis using bioinformatics analysis. *Biomed Res Int* 2017;2017:4323496.
5. Kaimal V, Sardana D, Bardes EE, Gudivada RC, Chen J, et al. Integrative systems biology approaches to identify and prioritize disease and drug candidate genes. *Methods Mol Biol* 2011;700:241-59.
6. Chen BS, Yang SK, Lan CY, Chuang YJ. A systems biology approach to construct the gene regulatory network of systemic inflammation via microarray and databases mining. *BMC Med Genomics* 2008;1:46-67.
7. Nair J, Ghatge M, Kakkar VV, Shanker J. Network analysis of inflammatory genes and their transcriptional regu-

- lators in coronary artery disease. *PLoS One* 2014;9(4): e94328.
8. Hofree M, Shen JP, Carter H, Ideker T. Network-based stratification of tumor mutations. *Nat Methods* 2013;10(11):1108-15.
 9. Ding L, Fan L, Xu X, Fu J, Xue Y. Identification of core genes and pathways in type 2 diabetes mellitus by bioinformatics analysis. *Mol Med Rep* 2019;20(3):2597-608.
 10. Lin Y, Li J, Wu D, Wang F, Fang Z, Shen G. Identification of hub genes in type 2 diabetes mellitus using bioinformatics analysis. *Diabetes Metab Syndr Obes* 2020;13: 1793-801.
 11. Vastrad B, Vastrad C. Bioinformatics analysis of potential key genes and mechanisms in type 2 diabetes mellitus. *bioRxiv*; 2021.
 12. Vastrad B, Tengli A, Vastrad C, Kotturshetti I. Bioinformatics analysis of key genes and pathways for obesity associated type 2 diabetes mellitus as a therapeutic target. *bioRxiv* 2020.
 13. Marín M, Esteban FJ, Ramírez-Rodrigo H, Ros E, Sáez-Lara MJ. An integrative methodology based on protein-protein interaction networks for identification and functional annotation of disease-relevant genes applied to channelopathies. *BMC Bioinformatics* 2019;20(1):565-77.
 14. Sahrawat TR, Dwivedi J. Investigating the relationship between diabetes and alzheimer's disease: A network systems biology approach. *J Sci Sustainable Development* 2020;7(1):1-11.
 15. Saxena A, Wahi N, Kumar A, Mathur SK. Functional interactomes of genes showing association with type-2 diabetes and its intermediate phenotypic traits point towards adipo-centric mechanisms in its pathophysiology. *Biomolecules* 2020;10(4):601-11.
 16. Fuchsberger C, Flannick J, Teslovich TM, Mahajan A, Agarwala V, Gaulton KJ, et al. The genetic architecture of type 2 diabetes. *Nature* 2016;536(7614):41-7.
 17. Flannick J, Fuchsberger C, Mahajan A, Teslovich TM, Agarwala V, Gaulton KJ, et al. Sequence data and association statistics from 12,940 type 2 diabetes cases and controls. *Sci Data* 2017;4:170179.
 18. Dai HJ, Wu JCY, Tsai RTH, Pan WH, Hsu WL. T-HOD: a literature-based candidate gene database for hypertension, obesity and diabetes. *Database (Oxford)* 2013;2013: bas061.
 19. Piñero J, Queralt-Rosinach N, Bravo À, Deu-Pons J, Bauer-Mehren A, Baron M, et al. DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database (Oxford)* 2015;2015: bav028.
 20. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019;47(D1):D607-613.
 21. Shannon P, Markiel A, Ozier O, Baliga NS, Wang AT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13(11):2498-504.
 22. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* 2014;8(4): S11.
 23. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Vaz Meirelles, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 2013;14:128-141.
 24. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 2016;44:W90-W97.
 25. Xie Z, Bailey A, Kuleshov MV, Clarke DJB., Evangelista JE, Jenkins SL, et al. Gene set knowledge discovery with Enrichr. *Curr Protoc* 2021;1(3):e90.
 26. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 2009;25(8): 1091-3.
 27. Dhomen NS, Mariadason J, Tebbutt N, Scott AM. Therapeutic targeting of the epidermal growth factor receptor in human cancer. *Crit Rev Oncol* 2012;17(1):31-50.
 28. Blaine SA, Ray KC, Branch KM, Robinson PS, Whitehead RH, Means AL. Epidermal growth factor receptor regulates pancreatic fibrosis. *Am J Physiol Gastrointest Liver Physiol* 2009;297(3):G434-441.
 29. Babel RA, Dandekar MP. A review on cellular and molecular mechanisms linked to the development of diabetes complications. *Curr Diabetes Rev* 2021;17(4):457-73.
 30. Barrick CJ, Yu M, Chao HH, Threadgill DW. Chronic pharmacologic inhibition of EGFR leads to cardiac dysfunction in C57BL/6J mice. *Toxicol Appl Pharmacol* 2008;228(3):315-25.
 31. Shraim BA, Moursi MO, Benter IF, Habib AM, Akhtar S. The role of epidermal growth factor receptor family of receptor tyrosine kinases in mediating diabetes-induced cardiovascular complications. *Front Pharmacol* 2021; 12:701390.
 32. Yuan J, Yin Z, Tao K, Wang G, Gao J. Function of insulin-like growth factor 1 receptor in cancer resistance to chemotherapy. *Oncol Lett* 2018;15(1):41-7.
 33. Delafontaine P, Song YH, Li Y. Expression, regulation, and function of IGF-1, IGF-1R, and IGF-1 binding proteins in blood vessels. *Arterioscler Thromb Vasc Biol* 2004;24(3):435-44.
 34. Kirstein M, Aston C, Hintz R, Vlassara H. Receptor-specific induction of insulin-like growth factor I in human monocytes by advanced glycosylation end product-modified proteins. *J Clin Invest* 1992;90(2):439-46.
 35. Kislinger T, Tanji N, Wendt T, Qu W, Lu Y, Ferran LJ, et al. Receptor for advanced glycation end products mediates inflammation and enhanced expression of tissue factor in vasculature of diabetic apolipoprotein E-null mice. *Arterioscler Thromb Vasc Biol* 2001;21:905-10.
 36. Patel VA, Zhang QJ, Siddle K, Soos MA, Goddard M, Weissberg PL, et al. Defect in insulin-like growth factor-1 survival mechanism in atherosclerotic plaque-derived

Network Analysis for T2DM Candidate Genes

- vascular smooth muscle cells is mediated by reduced surface binding and signaling. *Circ Res* 2001;88(9):895-902.
37. Bennett MR, Evan GI, Schwartz SM. Apoptosis of human vascular smooth muscle cells derived from normal vessels and coronary atherosclerotic plaques. *J Clin Invest* 1995;95(5):2266-74.
 38. Okura Y, Brink M, Itabe H, Scheidegger KJ, Kalangos A, Delafontaine P. Oxidized low-density lipoprotein is associated with apoptosis of vascular smooth muscle cells in human atherosclerotic plaques. *Circulation* 2000;102:2680-6.
 39. Petersen MC, Shulman GI. Mechanisms of insulin action and insulin resistance. *Physiol Rev* 2018;98:2133-23.
 40. Boucher J, Kleinridders A, Kahn CR. Insulin receptor signaling in normal and insulin-resistant states. *Cold Spring Harb Perspect Biol* 2014;6(1):a009191.
 41. Haeusler RA, McGraw TE, Accili D. Biochemical and cellular properties of insulin receptor signalling. *Nat Rev Mol Cell Biol* 2018;19:31-44.
 42. Copps KD, White MF. Regulation of insulin sensitivity by serine/threonine phosphorylation of insulin receptor substrate proteins IRS1 and IRS2. *Diabetologia* 2012;55(10):2565-82.
 43. Saltiel AR, Kahn CR. Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* 2001;414(6865):799-806.
 44. Prakash T, Ramachandra NB. Integrated network and gene ontology analysis identifies key genes and pathways for coronary artery diseases. *Avicenna J Med Biotechnol* 2021;13(1):15-23.