

## Construction of Gene Network in Type 2 Diabetes with Genome-Wide Association Study

Ming Zheng<sup>1, a \*</sup> and Mugui Zhuo<sup>1, b \*</sup>

<sup>1</sup>Guangxi Colleges and Universities Key Laboratory of Professional Software Technology, Wuzhou University, Wuzhou, China

<sup>a</sup>370505375@qq.com, <sup>b</sup>756456050@qq.com

**Keywords:** Diabetes mellitus; Type 2; Gene network; Pathway

**Abstract.** A genome-wide association analysis (GWAS) based on the research results, the local gene networks a preliminary establishment of susceptibility genes of type 2 diabetes mellitus, for future gene-gene interactions, provides reference for the research on gene environment interaction. 65 candidate genes selected from the study by GWAS of susceptibility genes of type 2 Diabetes Association, it selects 2 or more pathways target genes from 27, and through the selection of screening, intermediate gene pathway, construct local gene network, and use Ai CS6 software to draw network diagram. It consists of the following steps: (1) building a network of basic information of genes and pathways obtained by the KEGG and Genecards database; (2) establish a primary gene network; (3) selecting the middle gene; (4) the gene level two connection; (5) construction and mapping network local gene network. A total of 1 genes including the gene and a number of intermediate genes, as well as a series of linked genes and intermediate molecules, were obtained. The network contains 20 biological pathways. Local gene network conclusion this study obtained a more intuitive and comprehensive display of beta cell dysfunction and insulin resistance in susceptible relationship between genes, which can provide reference for study of molecular pathogenesis of type 2 diabetes in depth.

### Introduction

Type 2 diabetes[1] is a complex disease etiology, involving a variety of risk factors, from the point of view of genetics, is the interaction of multiple minor genes cause disease, but each individual genes are not enough to completely explain the cause. Early studies based on candidate genes brought us some research results, based on genome-wide genetic variation within and then produce (mainly single nucleotide polymorphism) association analysis (GWAS), for complex diseases (such as diabetes, hypertension, coronary heart disease and obesity etc.) susceptibility gene is a new research. The tools and methods, the method can be used to detect multiple genes at a time, and do not need to establish a hypothesis, but the results cannot fully explain some causes, therefore still need other technical support, so that it can have a broader development space. Since the GWAS of type 2 diabetes in 2007 the first since the world has carried out more than 30 studies found a series of susceptibility loci, mainly divided into insulin resistance and impaired beta cell function 2 susceptibility loci, in addition, there is a considerable part of unknown function gene. SNP mutations at these loci may influence the incidence of diabetes. Since the beginning of the study of gene networks in 1960s, has now entered the stage of complex gene regulatory networks, and has become a hot research at home and abroad, there have been a variety of models have been established, in the field of medicine, in addition to carrying out a tumor in a series of gene regulatory network study, to study hypertension gene network, researchers also begin to explain the relationship between the use of a variety of diseases such as diabetes and coronary heart disease gene network, association etc.. The development of gene network technology makes the research of complex diseases gene regulatory networks become possible in the candidate gene, GWAS research methods of susceptibility genes on the basis of comprehensive research results, the use of network technology to carry out multi gene-gene association is a new strategy, to explore the causes of complex diseases at the molecular

level and the mechanism. In this study, based on the results of the GWAS study, we selected some of the susceptibility loci to construct a local gene network to further explore the pathogenesis of type 2 diabetes[2].

## **Method and Gene Data**

**Gene Data.** To obtain the target genes are selected from 2007 up to now in the research of GWAS in type 2 diabetes with significant genes, a total of 65, according to the cause of diabetes mellitus, can be divided into impaired beta cell function (such as TCF7L2, HHEX, SLC30A8 and CDKN2A)[3], insulin resistance (such as IRS1, ADAMTS9, PTPRD and SPRY2) [4]and the function is no clear genes (such as ADCY5, HMGA2, DUSP9, AP3S2, PEPD and ANK1)[5]. Genetic and biological pathway data from KEGG (Kyoto Encyclopedia of Genes and Genomes, KEGG) database[6] (<http://www.kegg.com/jp>) and Genecards database[7] (<http://www.genecards.org>), path number and gene mapping to the KEGG database for.

**Methods.** Gets the 65 target gene pathway information by KEGG database and Genecards database, the genes and pathways to collect information in the table, and does not involve the biological pathways involved only one of the genes to be discarded, for a total of 27 primary target genes. 27 target genes were selected to access finishing to the table, screening for each gene and each pathway pathway, will involve 2 or more than 2 genes were screened out, give priority to the other pathway, gene pathways involved in many genes, through these pathways are connected, and as a connection based on gene. The primary gene connection is a simple form of linking the target genes, and some of them are relatively independent. In the two gene connection, the target genes are further connected through the intermediate genes. The pathway target genes were involved in the further biological pathway map provided by KEGG database, the target gene tagging in KEGG pathway, in the center of the pathway genes play the key roles in the pathway, can be connected to other target genes and pathways. These genes were annotated, and the relationship between diabetes and diabetes was obtained through literature review.

The phosphatidylinositol 3- kinase (phosphoinositide, 3-kinase, PI3K)[8]as an example, in the KEGG database, the PI3K gene is involved in 60 biological pathways, including PI3K-Akt[9] signaling pathway, Jak-STAT signaling pathway, TNF signaling pathway, insulin signaling pathway and type 2 diabetes[10], in a variety of metabolic processes were related to the expression of PI3K has correlation PI3K signaling pathway and glucose production, insulin secretion and signal transduction of literature reports, in the liver, PI3K signaling pathway mediated by glucose, PI3K signaling pathway is related to insulin sensitivity. Through the above process is often selected genes in various biological processes must be able to participate in the molecular biological pathway most, through the intermediate gene mediated to susceptibility genes connected. Intermediate gene selection is completed, respectively the encoding of the target gene, genes and pathways, the genes and pathways were involved in the pathway of axis list, gene tagging, in order to visually display the corresponding relationship between genes and pathways, and intuitive expression connections between genes. Because some genes are involved in multiple pathways, and there is some overlap between them, there may be a variety of connections between genes. Calculation of all connections, these connections can be expanded to establish a channel for connecting line, the gene network nodes, the network will be called two gene network, also known as the local gene network model of complex traits, many network only a tiny part of the accord with biological law.

The path was expanded to the target gene and the intermediate gene as the framework, using Adobe Illustrator Cs6 software to draw network diagram. As for diabetes, impaired beta cell function and insulin resistance as the two basic ways to extend all the gene pathway, due to a variety of connections between genes, genes and their products we will choose is connected with diabetes is most relevant to the way into the line connection, other have connections, but direct connection there is no contact with diabetes, for example, many genes are associated with the occurrence of cancer, it

will no longer be considered. On the basis of KEGG's access database, the related pathways were developed, and all the genes, gene expression products and intermediate molecules were introduced.

**Results**

Through the preliminary screening of the target gene, the pathways involved, we obtain an initial target gene based on network connection diagram, as shown in Figure 1, the primary connection includes 27 genes, gene line numbers for involving KEGG numbers pathway, corresponding relationship between the number and the path table 1 shows. Among the 27 target genes, WFS1, UBE2E2, DUSP9 and RASGRP1 genes belong to 2 relatively isolated pathways: Protein processing endoplasmic reticulum (hsa04141) and MAPK signaling pathway (hsa04010).

On the basis of the selection process of gene gene screening and above, a total of 27 target genes (G1-G27) and 10 (G28-G37) into the intermediate gene to construct local gene network, the number of target genes and pathways, corresponding relationship between the numbers shown in Table 1 and table 2. In 37 genes and 20 pathways as the basis, we construct a 2 gene network connection, as shown in Figure 2, the line between the 2 genes and expressed genes in the same pathway, does not represent the functional association between genes, the connection is just one of 1 forms of many connections. The purpose is all scattered in the gene set to the 1 network, so it is possible to network access is not real or not the most accurate, need to be further revised and improved, at the same time were compared and analyzed with other possible construction of local gene network access, to further confirm this or confirmed type 2 diabetes local gene network structure. The results can be seen from Fig. 1 as below:

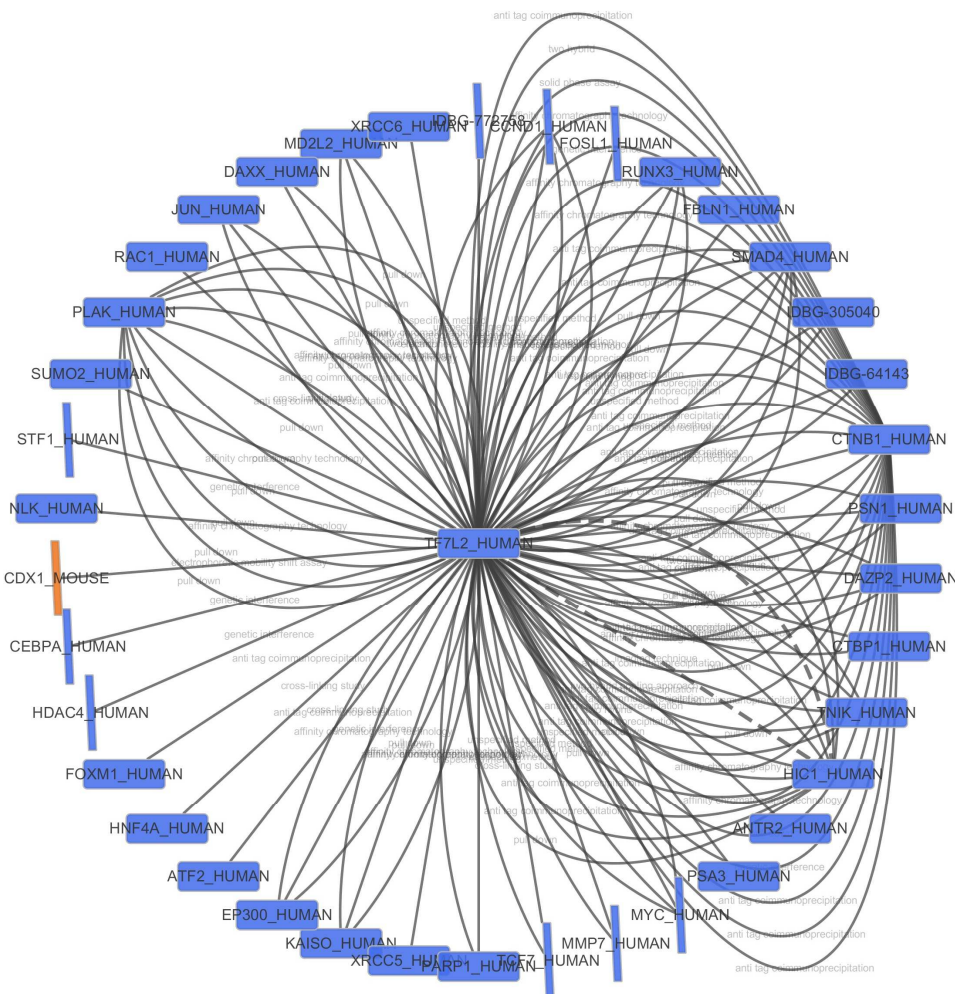


Figure 1. The gene regulatory network of the type 2 diabetes

## Summary

To build a local gene network using the theory of molecular biology, may be able to explain the relationship between the target gene at the molecular level, however, such a way of building is still in the stage of theoretical exploration, limited to the author's knowledge level, the local gene network still has some shortcomings, it also needs the molecular epidemiology of experiments to test and verify, through animal experiments and molecular biology experiments, upregulation of the expression of certain genes, expression of other genes in order to observe the affected, to further define the relationship between each gene, the local gene network original modify, expand and improve, to explore the mechanism of the occurrence and development of complex diseases from the molecular level.

## Acknowledgements

This work was supported by grants from The National Natural Science Foundation of Chi-na (No. 61502343, No. 61373051, and No. 61402423), China Postdoctoral Science Foundation funded(No. 2016M590260), the Guangxi Natural Science Foundation (No. 2015GXNSFBA139262), the Science Research Funds for the Guangxi Universities (No. KY2015ZD122), Guangxi Colleges and Universities Key Laboratory of Professional Software Technology, Wuzhou University.

## References

- [1] MOK Y, TAN C W, WONG H S, et al. Obstructive sleep apnoea and Type 2 diabetes mellitus: are they connected?[J]. Singapore Med J, 2017, 58(4): 179-183.
- [2] CHUNG C M, LIN T H, CHEN J W, et al. Common quantitative trait locus downstream of RETN gene identified by genome-wide association study is associated with risk of type 2 diabetes mellitus in Han Chinese: a Mendelian randomization effect[J]. Diabetes Metab Res Rev, 2014, 30(3): 232-240.
- [3] CORELLA D, COLTELL O, SORLI J V, et al. Polymorphism of the Transcription Factor 7-Like 2 Gene (TCF7L2) Interacts with Obesity on Type-2 Diabetes in the PREDIMED Study Emphasizing the Heterogeneity of Genetic Variants in Type-2 Diabetes Risk Prediction: Time for Obesity-Specific Genetic Risk Scores[J]. Nutrients, 2016, 8(12).
- [4] FERNANDES J C, RODRIGUES ALVES A P, MACHADO-NETO J A, et al. IRS1/beta-Catenin Axis Is Activated and Induces MYC Expression in Acute Lymphoblastic Leukemia Cells[J]. 2016.
- [5] KNIGGE A, KLOTING N, SCHON M R, et al. ADCY5 gene expression in adipose tissue is related to obesity in men and mice[J]. PLoS One, 2015, 10(3): e0120742.
- [6] KANEHISA M. KEGG Bioinformatics Resource for Plant Genomics and Metabolomics[J]. Methods Mol Biol, 2016, 1374: 55-70.
- [7] FISHILEVICH S, ZIMMERMAN S, KOHN A, et al. Genic insights from integrated human proteomics in GeneCards[J]. Database (Oxford), 2016, 2016.
- [8] JAYSON G C, KOHN E C, KITCHENER H C, et al. Ovarian cancer[J]. Lancet, 2014, 384(9951): 1376-1388.
- [9] LOPEZ-CARDONA A P, PEREZ-CEREZALES S, FERNANDEZ-GONZALEZ R, et al. CB1 cannabinoid receptor drives oocyte maturation and embryo development via PI3K/Akt and MAPK pathways[J]. Faseb j, 2017.
- [10] CHEN J, ZHONG M C, GUO H, et al. SLAMF7 is critical for phagocytosis of haematopoietic tumour cells via Mac-1 integrin[J]. Nature, 2017.