

RESEARCH ARTICLE

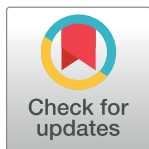
Integrative analysis of super enhancer SNPs for type 2 diabetes

Weiping Sun¹*, Sihong Yao²*, Jielong Tang³, Shuai Liu³, Juan Chen¹, Daqing Deng¹, Chunging Zeng^{3*}

1 Department of Geriatrics, the First People's Hospital of Xiangtan City, Xiangtan, PR, China, **2** Department of Clinical Medicine, Jishou University School of Medicine, Jishou, PR, China, **3** Department of Endocrinology, the Fifth Affiliated Hospital of Guangzhou Medical University, Guangzhou, PR, China

* These authors contributed equally to this work.

* zcp193@163.com



OPEN ACCESS

Citation: Sun W, Yao S, Tang J, Liu S, Chen J, Deng D, et al. (2018) Integrative analysis of super enhancer SNPs for type 2 diabetes. PLoS ONE 13 (1): e0192105. <https://doi.org/10.1371/journal.pone.0192105>

Editor: Margaret M. DeAngelis, University of Utah, UNITED STATES

Received: June 8, 2017

Accepted: January 18, 2018

Published: January 31, 2018

Copyright: © 2018 Sun et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This study was partially supported or benefited by grants from technological innovation project of Foshan City (2017AG100102), and medical technology research fund of Guangdong province (A2017575).

Competing interests: The authors have declared that no competing interests exist.

Abstract

Clinical studies in type 2 diabetes (T2D) primarily focused on the single nucleotide polymorphisms (SNPs) located in protein-coding regions. Recently, the SNPs located in noncoding regions have also been recognized to play an important role in disease susceptibility. The super enhancer is a cluster of transcriptional enhancers located in noncoding regions. It plays a critical role in cell-type specific gene expression. However, the exact mechanism of the super enhancer SNPs for T2D remains unclear. In this study, we integrated genome-wide association studies (GWASs) and T2D cell/tissue-specific histone modification ChIP-seq data to identify T2D-associated SNPs in super enhancer, followed by comprehensive bioinformatics analyses to further explore the functional importance of these SNPs. We identified several interesting T2D super enhancer SNPs. Interesting, most of them were clustered within the same or neighboring super enhancers. A number of SNPs are involved in chromatin interactive regulation and/or potentially influence the binding affinity of transcription factors. Gene Ontology (GO) analysis showed a significant enrichment in several well-known signaling pathways and regulatory process, *e.g.* WNT signaling pathway, which plays a key role in T2D metabolism. Our results highlighted the potential functional importance of T2D super enhancer SNPs, which may yield novel insights into the pathogenesis of T2D.

Introduction

Type 2 diabetes (T2D) is a long-term metabolic disorder which is characterized by hyperglycemia and the aberrant metabolism of fat and protein due to the deficient action of insulin [1]. Over the past two decades, the increasing prevalence of T2D has become a serious public health burden with over 4.9 million deaths in 2014 throughout the world [2], and it is estimated that 439 million people would suffer T2D by 2030 [3]. T2D has become one of the leading causes of death, disability, and the rising cost of health care. Compelling evidence showed that the risk of T2D is strongly influenced by genetic factors [4]. Recently, the successive waves

of genome-wide association studies (GWASs) have inferred more than 80 robust T2D single nucleotide polymorphisms (SNPs), but in aggregate, these SNPs explained only a small fraction of the T2D heritability [5].

Previous studies have shown that SNPs in protein-coding region are associated with hundreds of diseases by affecting the expression of their target genes [6]. However, over 90% of GWAS-identified variants lie within noncoding regions and the mechanisms of how these SNPs contribute to disease susceptibility remain unclear [7]. Super enhancer, a cluster of transcriptional enhancers, overlapped with noncoding regions (*e.g.* DNA methylation valleys and locus control regions) which plays an important role in cell-type specific gene expression [8]. A recent study showed that the SNPs associated with specific diseases are particularly enriched in the super enhancers of disease-relevant cell/tissue types [9]. For example, the SNPs associated with Alzheimer's disease are enriched in super enhancers of human brain tissue, and super enhancers in B and T cells are enriched for SNPs which are associated with rheumatoid arthritis [9]. However, the role of cell/tissue-type specific super enhancer SNPs for T2D is still unknown.

In this study, we integrated T2D GWAS meta-analysis dataset with T2D cell/tissue-specific histone modification ChIP-seq data to identify T2D super enhancer SNPs, followed by comprehensive bioinformatics analyses to further elucidate the potential functional importance of these SNPs in the pathogenesis of T2D. Our results may yield novel insights into the genetic basis of T2D.

Materials and methods

T2D GWAS datasets

We retrieved 792 potential T2D SNPs at GWAS significance threshold (p -value $< 5 \times 10^{-8}$) from Diabetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium (<http://diagram-consortium.org/downloads.html>). To our knowledge, it is the standard and the largest T2D meta-analysis study which conducted in 34,840 cases and 114,981 controls of European descent. Next, we used proxy SNAP (<https://www.broadinstitute.org/mpg/snap/ldsearch.php>) to identify SNPs in LD with retrieved T2D SNPs. The search was based on genotype data from the 1000 Genomes Project with the Utah residents with European ancestry (CEU) population panel. The inclusion criteria for LD SNPs were set as a distance limit of 500 KB from the query SNP with $r^2 > 0.8$ from pairwise LD calculations. A total of 1,086 potential functional T2D SNPs was identified.

Histone modification datasets

We selected T2D-relevant cell/tissues based on their crucial role in pathogenesis of T2D through comprehensive, systematic literature searches in PubMed. In total, we identified 27 T2D-relevant cell/tissue types, such as adipose tissue [10–13], aorta [14–16], bladder [17–19], liver [20–22], T cell [13, 23–25], pancreas [26–29], thymus [30], brain [31–34], osteoblasts [35], skeletal muscle and myoblast [36–38]. The human histone modification WIG/BIGWIG files (H3K27ac) in 27 T2D-relevant cell/tissue types were downloaded from Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>). A detailed list of all T2D-relevant cell/tissue types and sample information are included in [S1 Table](#).

Identification of T2D super enhancer and T2D super enhancer SNPs

Histone modification WIG/BIGWIG files were mapped to the human reference genome (hg19) using Bowtie v1.1.2 (<http://bowtie-bio.sourceforge.net/index.shtml>). MACS v1.4.1 was

used to identify enhancer regions [39], it applies a dynamic parameter, λ_{local} , to calculate the p-value of each candidate peak and removes potential false positives due to local biases. The significant threshold was defined as Bonferroni adjust p-value $< 1 \times 10^{-9}$. Super-enhancers were separated from enhancers using the ROSE algorithm [40]. A detailed description of the super enhancer identification process was illustrated in the previously published pipeline [9]. We identified a total of 18,422 super enhancers in 27 T2D-relevant cell/tissue types (S2 Table) and mapped 1,086 potential functional T2D SNPs to these super enhancers. Finally, we achieved 286 T2D super enhancer SNPs (S3 Table).

Functional prediction of T2D super enhancer SNPs

Characterization of the regulatory features of T2D super enhancer SNPs was carried out through the bioinformatics tool rVarBase [41]. rVarBase utilizes experimentally supported regulatory elements from ENCODE and other data resources to make relevant annotation, it provides reliable regulatory feature for human variants. rVarBase annotates regulatory feature of variants from chromatin state of the queried variant surrounding region, regulatory elements that overlapped with the queried variant, and potential target genes of queried variant. To further validate the potential functional consequence of T2D super enhancer SNPs, we applied HaploReg v4.1 [42], a comprehensive resource which utilizes data from Roadmap Epigenomics, ENCODE and Genotype-Tissue Expression Project (GTEx) *etc.*, to explore the effect of T2D super enhancer SNPs on gene expression (eQTLs) and regulatory motif alterations within sets of genetically linked T2D super enhancer SNPs. Furthermore, we performed transcription factor enrichment analysis using SNP2TFBS (<http://ccg.vital-it.ch/snp2tfbs/>) tool which selects and visualizes user defined variants that affect single or multiple transcription factors.

Long-range interaction analysis of T2D super enhancer SNPs

GWAS3D (<http://jjwanglab.org/gwas3d>) systematically calculates the probability of GWAS-associated SNPs which affect regulatory pathways and underlie disease associations by integrating functional genomics, chromatin state, sequence motif, and conservation information [43]. Moreover, the results can be interpreted by comprehensive visualization. In this study, we used GWAS3D to identify the significant T2D super enhancer SNPs which have a long-range interaction signal with their distal regulatory elements using the system default parameters: Fisher's combined p-value $< 1 \times 10^{-5}$, $r^2 = 0.8$, CEU population panel (HapMap phase I+II+III), binding site p-value = 0.01, size of 30 variants, and 3 interaction size.

Gene ontology (GO) analysis

To find out the functional enrichment (*e.g.* biological processes) of T2D super enhancer SNPs, we assigned each T2D super enhancer SNPs to their nearest genes based on the human genome assembly GRCH37 (hg19) from Genome Reference Consortium. We carried out GO analyses using these T2D super enhancer SNPs target genes through GOEAST toolkit (<http://omicslab.genetics.ac.cn/GOEAST/>). The p-values were calculated by hypergeometric tests and adjusted for multiple comparisons using stringent Yekutieli (FDR under dependency) method [44]. The statistical significant threshold was defined as Yekutieli adjust p-value < 0.05 .

Protein-protein interaction networks

STRING v10.0 (<http://string-db.org/>) is a database of known and predicted protein interaction, including physical and function association. It quantitatively integrates interaction data

derived from genomic context, high-throughput experiments, co-expression and previous knowledge [45]. The database currently contains 9.6 million proteins from 2031 organisms constructed more than 184 million interactions. In order to explore the functional partnership and interaction of the identified T2D super enhancer SNP target genes (S3 Table), we construct gene (protein) interaction networks via the STRING v10.0 database with default settings (observed interaction, 70; expected interaction, 8.83; Benjamini adjust p-value < 1×E-10; proteins, 76).

Results

Identification of T2D super enhancer SNPs

In this study, we first retrieved 792 potential T2D SNPs at GWAS significance threshold (p-value < 5×E-8) from the largest T2D meta-analysis study DIAGRAM Consortium. By using LD information from the 1000 Genomes Project with CEU population panel, we inferred 1,086 potential functional T2D SNPs. Then we took advantage of the public available human histone modification ChIP-seq datasets from GEO database to construct a comprehensive list of super enhancer in T2D-relevant cell/tissue types. In total, we examined and confirmed 18,422 super enhancer regions in 27 human cell/tissue types (S2 Table). Then 1,086 potential functional T2D SNPs were mapped to the identified T2D cell/tissue-specific super enhancer. Finally, we obtained a total of 286 potential functional T2D super enhancer SNPs (S3 Table).

To identify whether T2D super enhancer SNPs enriched in specific cell/tissue types, we characterized the distribution of the T2D super enhancer SNPs in different cell/tissue types and identified a significant SNP enrichment in adipose tissue (161/286), aorta (115/286), pancreas (113/286) and brain (44/286) (S4 Table). Notably, 160 SNPs were mapped to super enhancer in a wide variety of cell/tissue types (S4 Table). Interestingly, some of these SNPs enriched in the same or neighboring super enhancers, the strong enrichment of T2D super enhancer SNPs on the chromosome 2p21 (*THADA*), 3q21.1 (*ADCY5*), 4p16.1 (*WFS1*) and 10q25.2 (*TCF7L2*) are rather striking (Fig 1). In addition, we checked the LD status among each of the SNPs pairs that mapped to the same or neighboring super enhancers. To our surprise, we identified several SNP pairs with low LD which were mapped to the same super enhancers (S4 Table & S5 Table), e.g. the SNP pair rs11196182 and rs12245680 with low LD ($r^2 = 0.016$) were mapped to the same super enhancer in adipose tissue (10q25.2). The full list of SNPs that map to a specific super enhancer and LD between those SNPs is included in S5 Table.

Functional prediction of T2D super enhancer SNPs

To investigate the potential functional impact of the identified T2D super enhancer SNPs, we annotated 286 T2D super enhancer SNPs to various regulatory elements through rVarBase [41]. rVarBase is a database which utilizes experimentally supported regulatory elements from multiple data resources to make relevant functional prediction, it provides reliable regulatory feature for human variants [41]. Finally, we identified 57 regulatory SNPs which were supported by experimental evidences. Of the 57 SNPs identified, 20 SNPs are involved in chromatin interactive regulation, 4 SNPs are involved in regulation of lncRNA expression, and 4 SNPs overlapped with CpG islands (S6 Table). In addition, to validate the potential functional consequence of the T2D super enhancer SNPs, we investigated the effect of T2D super enhancer SNPs on regulatory motifs and gene expression (eQTLs) by using data from ENCODE and GTEx projects through HaploReg v4.1. Finally, we observed that 242 T2D super enhancer SNPs changed transcription factor binding motif, this result highlighted the strong regulatory potential of these T2D super enhancer SNPs (S7 Table). Moreover, we identified 186 T2D

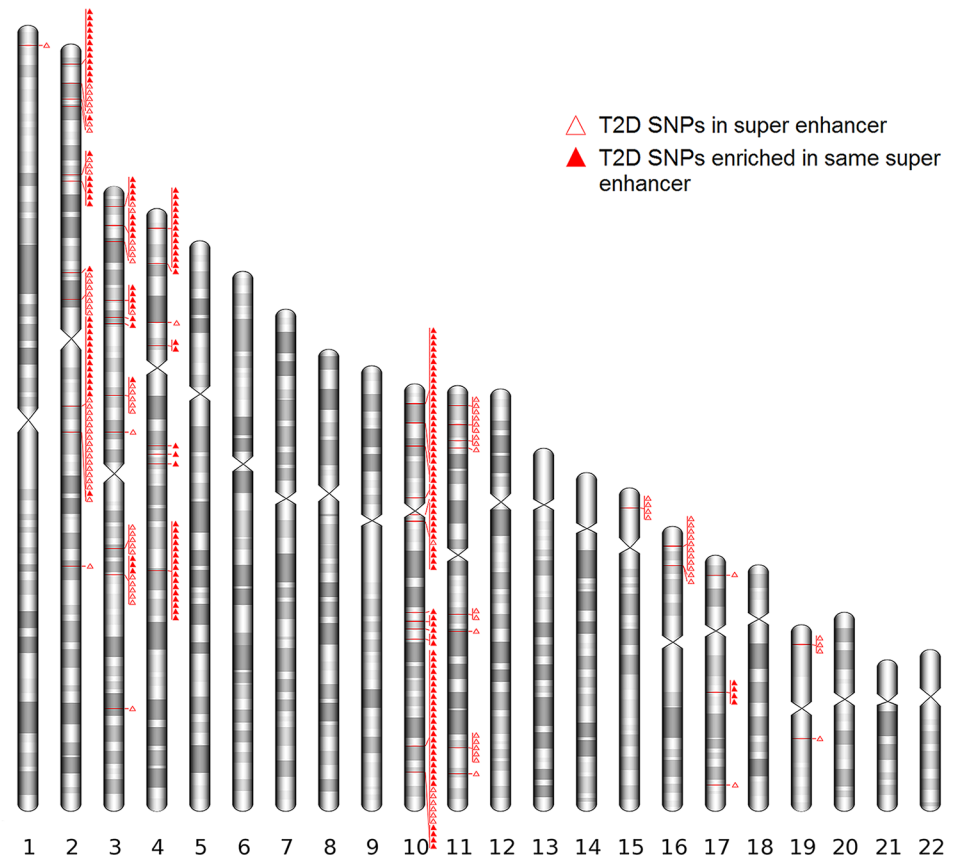


Fig 1. Distribution of T2D super enhancer SNPs in human genome. T2D super enhancer SNPs were annotated to the human genome assembly GRCH37 (hg19). The enrichment of T2D super enhancer SNPs was observed on chromosome 2p21 (*THADA*), 3q21.1 (*ADCY5*), 4p16.1 (*WFS1*) and 10q25.2 (*TCF7L2*).

<https://doi.org/10.1371/journal.pone.0192105.g001>

super enhancer SNPs which were reported eQTL evidences, including 120 T2D super enhancer SNPs showed eQTL evidences in a wide variety of tissues and cell types (S7 Table). Importantly, we found that all 57 regulatory SNPs identified by rVarBase were replicated in HaploReg regulatory prediction results (S6 Table & S7 Table).

The effect of T2D super enhancer SNPs on the transcription factor binding affinity was further analyzed by SNP2TFBS. The top five enrichment transcription factors are Tcf12, Sox6, JUNB, Myog, and Gata4 (Fig 2). Transcription factors Tcf12 is an important component member of WNT signaling and ERK signaling pathway. Diabetogenic factors have been found to influence insulin signaling through activation of the ERK signaling pathway [46]. Sox6 and JUNB are involved in T2D metabolism, T2D-related cell development and differentiation [47, 48]. Gata4 is important transcription factor which is involved in the regulation of T2D development and metabolism [49].

Long-range interaction of T2D super enhancer SNPs

GWAS3D was used to identify T2D super enhancer SNPs which have a long-range interaction signal with their distal regulatory elements. Finally, we detected 82 SNPs (Fisher’s combined p-value < 1×E-5) affecting the long range interaction based on the CEU population and all cell types listed in GWAS3D (Fig 3, S8 Table). For example, SNP rs13411629, located in the

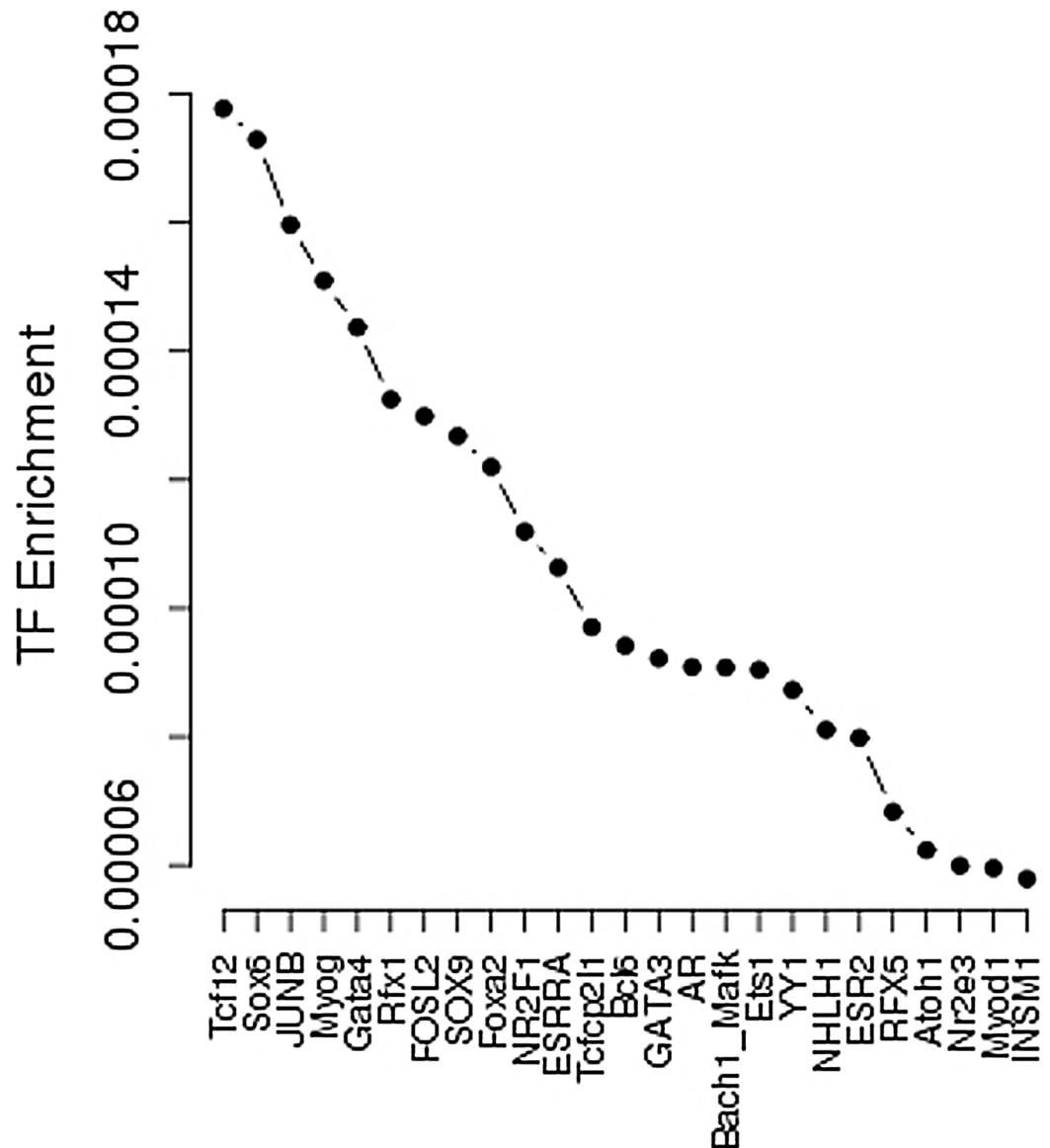


Fig 2. Magnified transcription factors enrichment plot for T2D super enhancer SNPs. Transcription factors are sorted based on their enrichment.

<https://doi.org/10.1371/journal.pone.0192105.g002>

intronic region of *THADA* on chromosome 2, has a long-range interaction signal with locus 11p15.4. SNP rs2124500, located in the intronic region of *ADCY5* on chromosome 3, has two long-range interaction signals with locus 7q36.1 and 18q12.2. SNP rs1552224, located in the intronic region of *ARAP1* on chromosome 11q13.4, showed two long-range interaction signals with locus 10p23.2 and locus near *POTEG* respectively. Furthermore, GWAS3D also revealed 105 SNPs which may affect promoter activity by changing transcription factor binding site affinity (**S8 Table**), including 76 SNPs have direct effect by GWAS leading SNPs and 29 variants have indirect effect by high LD of GWAS leading SNPs. Interestingly, among the 105 functional SNPs identified by GWAS3D, 96 SNPs were also confirmed by HaploReg regulatory prediction results (**S7 Table & S8 Table**).

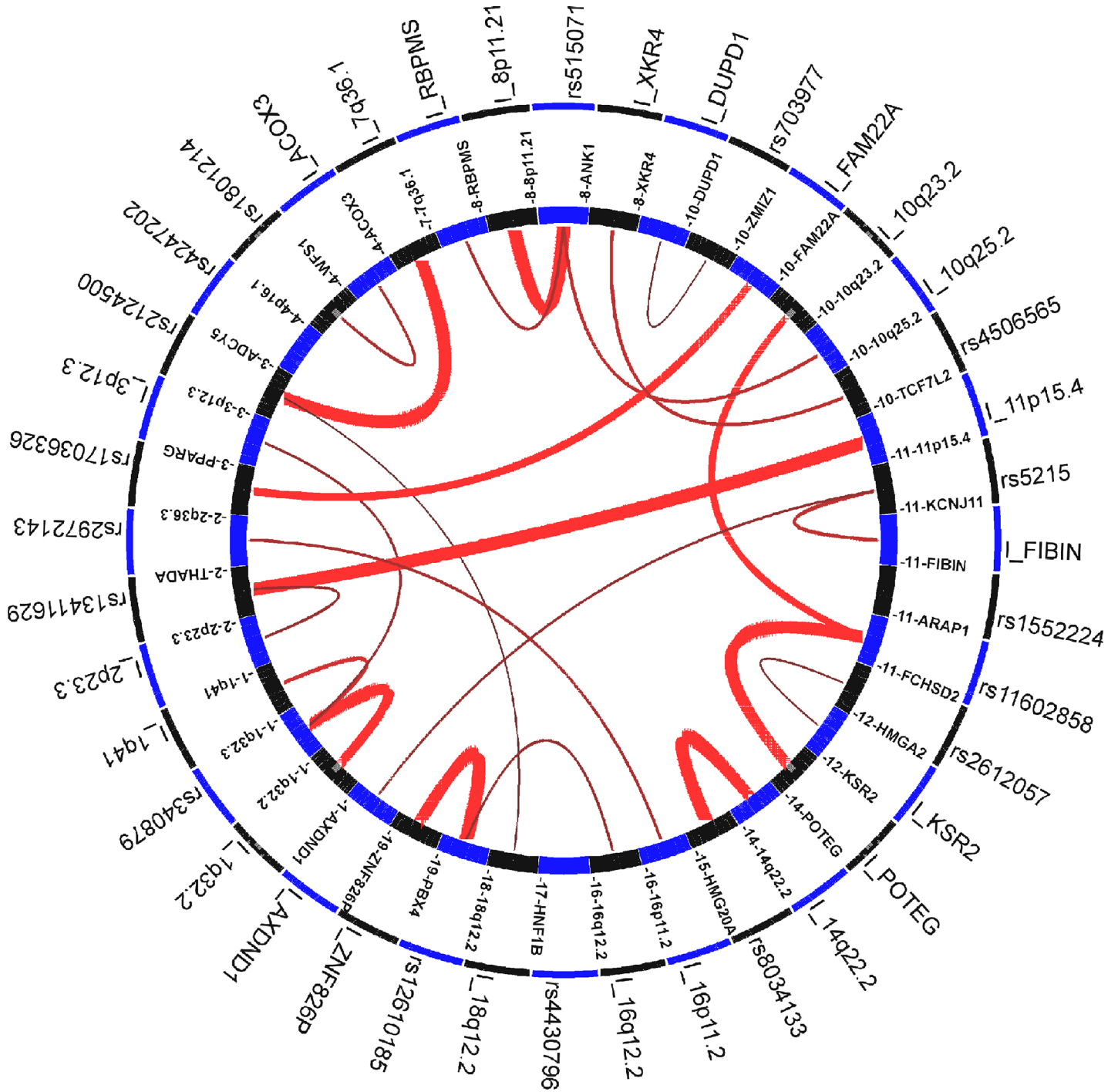


Fig 3. The circle plot of GWAS3D for T2D super enhancer SNPs. The annotation based on all cell line and CEU population. The red line indicated long-range interaction signals, and the intensity of interaction was represented by the width of the line. Interactive elements with significant SNP will start with 'L'.

<https://doi.org/10.1371/journal.pone.0192105.g003>

GO analysis of T2D super enhancer SNPs target genes

To systematically investigate whether the identified T2D super enhancer SNPs were specifically associated with T2D, we conduct GO analysis using 26 T2D super enhancer SNPs target genes ([S3 Table](#)) through GOEAST. We identified 151 significant GO terms with Yekutieli

Table 1. The most significant GO terms for T2D super enhancer SNPs target genes.

GOID	Term	log_odds_ratio	p-value
GO:0048008	platelet-derived growth factor receptor signaling pathway	6.72	6.97E-13
GO:0010172	embryonic body morphogenesis	7.70	7.26E-10
GO:0050794	regulation of cellular process	1.14	4.98E-08
GO:0023052	Signaling	1.64	1.10E-07
GO:0044700	single organism signaling	1.64	1.10E-07
GO:0050789	regulation of biological process	1.08	1.90E-07
GO:0007154	cell communication	1.59	2.04E-07
GO:0065007	biological regulation	1.02	3.75E-07
GO:0007166	cell surface receptor signaling pathway	2.27	6.59E-07
GO:0051974	negative regulation of telomerase activity	7.88	9.20E-07
GO:0016055	WNT signaling pathway	3.99	3.00E-06
GO:0198738	cell-cell signaling by WNT	3.99	3.00E-06
GO:0007165	signal transduction	1.57	3.07E-06
GO:0007186	G-protein coupled receptor signaling pathway	3.21	3.19E-06
GO:0051972	regulation of telomerase activity	7.39	3.53E-06

GO enrichment analysis was performed using the GOEAST tool. The p-values were calculated by hypergeometric tests and adjusted for multiple comparisons using stringent Yekutieli (FDR under dependency) adjustment.

<https://doi.org/10.1371/journal.pone.0192105.t001>

adjust p-value < 0.05 (S9 Table). Interestingly, we observed a significant enrichment in signaling pathways and regulation process (Table 1, S9 Table), e.g. WNT signaling pathway and G-protein coupled receptor signaling pathway, which play a key role in the pathogenesis of T2D [50, 51]. Other interesting GO terms, such as positive regulation of lipid metabolic process and negative regulation of inflammatory response, may also have a potential function in the metabolism of T2D.

Protein–protein interaction networks

To partially characterize the functional partnership and their interaction networks among the identified T2D super enhancer SNPs target genes, we uploaded the 26 T2D super enhancer SNPs target genes (S3 Table) into STRING v10.0 database with default settings (observed interaction, 70; expected interaction, 8.83; Benjamini adjust p-value < 1×E-10; proteins, 76). Fig 4 revealed a strong association between the topological properties and biological function of T2D target genes. The hub genes with the strong connections were *KCNJ11*, *PPARG*, *ABCC8*, *THADA*, *WFS1*, *TCF7L2*, and *ADCY5*. Interestingly, these genes are involved in the energy reserve metabolic process, glucose homeostasis, and negative regulation of pancreatic β cell apoptotic process, which revealed their potential functional impacts on the pathogenesis of T2D.

Discussion

In the human genome, there are much more genetic variations located in the noncoding regions when compared with protein-coding regions. Some important functional regulatory elements, such as the super enhancer, have a great impact on cell-type specific gene expression. SNPs located in super enhancer may play an essential role in disease metabolism. In this study, we integrated GWASs data from DIAGRAM and T2D cell/tissue-specific histone modification ChIP-seq data from GEO to identify T2D super enhancer SNPs. To explore the potential functional consequence of these SNPs and their target genes, we characterized their regulatory

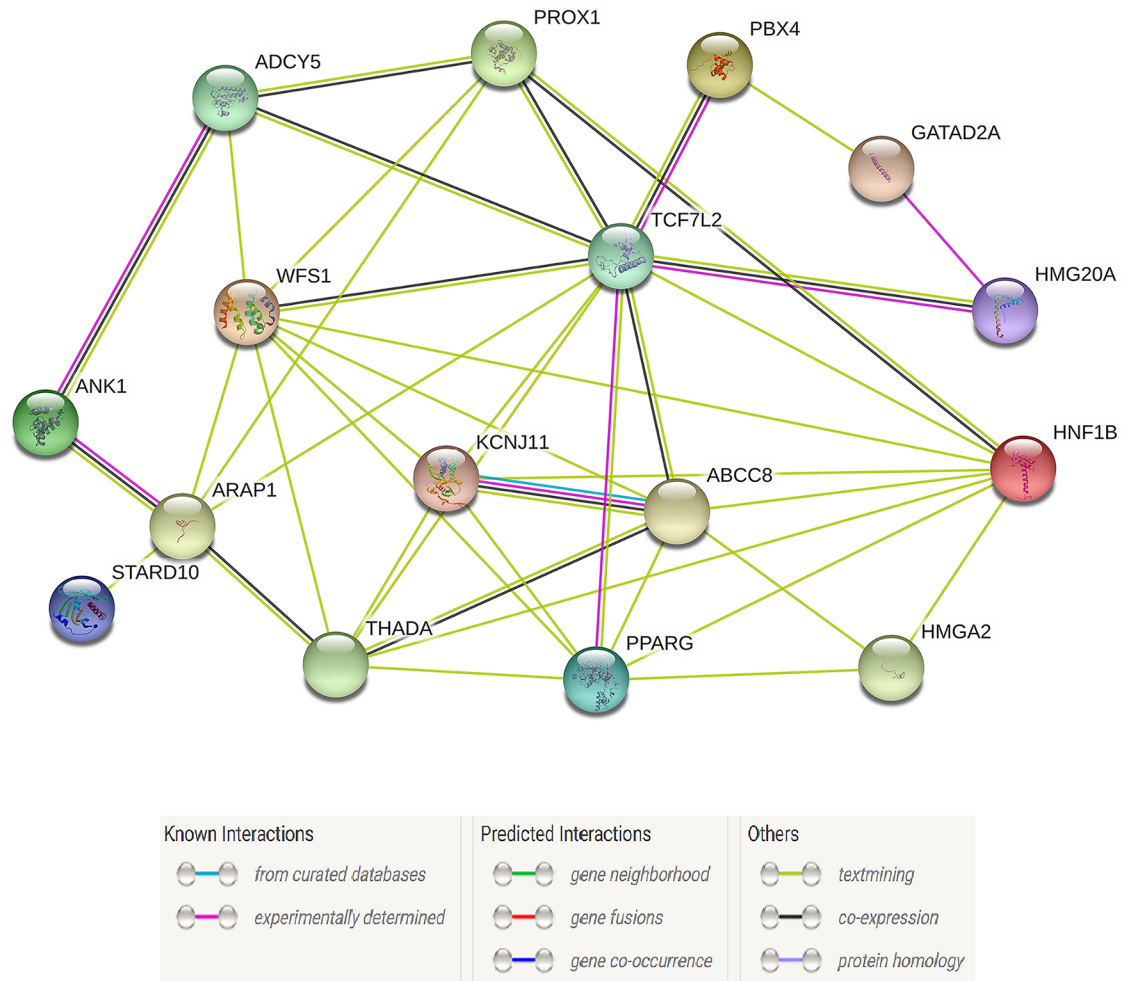


Fig 4. A functional protein-protein interaction network for T2D super enhancer SNPs target genes. Connections are based on co-expression and experimental evidence with a STRING v10.0 summary score above 0.4. Each filled node denotes a gene; edges between nodes indicate protein-protein interactions between protein products of the corresponding genes. Different edge colors represent the types of evidence for the association.

<https://doi.org/10.1371/journal.pone.0192105.g004>

feature using several different database, performed transcription factor enrichment analysis, long-range interaction analysis, and conduct GO and protein-protein interaction network analysis.

Our results indicate a total of 286 potential functional T2D super enhancer SNPs. Interestingly, some of these SNPs showed a strong enrichment in the T2D genes or novel candidate genes, *e.g.* *ADCY5*, *ANK1*, *PPARG*, *TCF7L2*, *THADA*, *WFS1*, and *ZMIZ1*. We identified 18 super enhancer SNPs enriched in T2D-associated gene *ADCY5*, this gene encodes a member of the membrane-bound adenylyl cyclase enzymes which are expressed in heart, brain, and pancreatic β cells, *etc.* Adenylyl cyclases mediate G-protein coupled receptor signaling through the synthesis of the second messenger cAMP, which can modulate glucose-stimulated insulin secretion by several possible mechanisms [52]. Another interesting gene *ANK1*, located on chromosome 8p11.21, contains 6 super enhancer SNPs, this gene encodes a member of the ankyrin family which links the integral membrane proteins to the spectrin-actin cytoskeleton [53]. It is expressed in several human tissues, *e.g.* endocrine pancreas, which is relevant to the glucose metabolism [54]. A recent study reported that variants in *ANK1* were significantly

associated with the susceptibility to T2D and there exists tissue-specific expression profile of *ANK1* in the human islet, pancreas, skeletal muscle, adipose and liver tissues which are important organs for glucose metabolism [55]. There are 16 super enhancer SNPs enriched in *PPARG* gene. This gene encodes a member of the nuclear receptor family of ligand-activated transcription factors and is predominantly expressed in adipose tissue [56]. Werman *et al.* suggested that *PPARG* is a major regulator of adipogenesis via stimulation by fatty acids, *e.g.* insulin enhances the transcriptional activity of the *PPARG* [57]. Importantly, we identified that super enhancer SNPs were highly enriched (88/286) in *TCF7L2* gene. Variants in *TCF7L2* have been associated with T2D in multiple ethnic groups [58–60]. As a transcription factor, *TCF7L2* plays a critical role in WNT signaling pathway and is involved in the development of various cell lineages and tissues [61]. Potential mechanisms of *TCF7L2* SNPs influencing on T2D include its role in adipogenesis, pancreatic islet development, and insulin secretory granule function.[62]. It is also involved in impaired glucose production and tolerance via direct effects on pancreatic β cells [63].

Functional annotation of T2D super enhancer SNPs through rVarBase showed 57 regulatory SNPs, including 20 SNPs which are involved in chromatin interactive regulation, 4 SNPs which are involved in regulation on lncRNA, and 4 SNPs overlapped with CpG islands. Additional functional prediction analysis showed that all these 57 regulatory SNPs were replicated in HaploReg results, which highlighted the strong and reliable regulatory potential of these T2D super enhancer SNPs. Interestingly, 22 out of 54 regulatory SNPs targeted to gene *WFS1*, this gene encodes a transmembrane protein which maintains calcium homeostasis of the endoplasmic reticulum and is ubiquitously expressed in brain, pancreas, heart, and pancreatic β cell lines [64]. *WFS1* is critical for the endoplasmic reticulum stress response in insulin-producing pancreatic β cells which contributes to the risk of T2D [64]. A recent study shows that the SNPs associated with prostate cancer can modulate lncRNA expression. This result indicates that lncRNA may play an essential role in human complex disease. Notably, we identified 4 regulatory SNPs, rs508419 in gene *ANK1*, rs4689393, rs3821941 and rs4688987 in gene *WFS1* which are involved in regulation on lncRNA, SNP rs508419 also targeted to a CpG island. In addition, we indicated that T2D super enhancer SNPs may alter the transcription factor binding affinity of Tcf12, Sox6, JUNB, Myog, and Gata4, *etc.* For example, transcription factors Gata4 is critical for the regulation of T2D development and metabolism [49]. Transcription factors Sox6 and JUNB are involved in T2D metabolism, T2D-related cell development and differentiation [47, 48]. Furthermore, we examined the effect of T2D super enhancer SNPs on expression of their target genes (eQTLs) by using data from GTE projects, 186 T2D super enhancer SNPs showed eQTL evidences, including 120 T2D super enhancer SNPs have numerous reported eQTL evidences in a wide variety of tissues and cell types.

GWAS3D revealed 82 T2D super enhancer SNPs which have long-range interaction signals. As long-range interaction can also be cell/tissue type specific, we performed additional analysis to examine whether long-range interaction of T2D super enhancer SNPs and their distal regulatory elements occurred in the same cell/tissue type. Due to the limited cell/tissue types listed in GWAS3D, we did not find the direct long-range interactions in the same cell/tissue type. However, we discovered 21 T2D super enhancer SNPs in adipose tissue which showed long-range interactions in K562 cell line. This cell line was the first human immortalised myelogenous leukemia line to be established [65]. Interestingly, among the 21 T2D super enhancer SNPs in adipose tissue, 13 SNPs were mapped to *TCF7L2* gene. This gene encodes a WNT transcription factor which has been linked to a higher risk of developing T2D [58–60, 66]. Previous studies have also indicated that WNT signaling pathway is the most commonly dysregulated process in myelogenous leukemia. Notably, *TCF7L2* plays a crucial role in maintaining the proliferation and viability of myeloid leukaemia cells, and overexpression of

TCF7L2 is associated with poor clinical outcome [67]. *TCF7L2* maybe a valid target in myelogenous leukemia therapy. In addition, according to a recently published study, the myelogenous leukemia treatment drug Gleevec has been stumbled on possible use for curing T2D [68]. The researchers discovered that Gleevec blocks CDK5-mediated PPAR γ phosphorylation, and thus lowers the level insulin resistance and reduces the risk of hyperglycemia and obesity. They also observed that Gleevec reduce lipogenic and gluconeogenic gene expression in liver and improved inflammation in adipose tissues. Although Gleevec is a new hope for both myelogenous leukemia and T2D cure, the exact mechanisms and their functional relationships remain unclear.

GO analysis of T2D super enhancer SNPs target genes not only confirmed well known WNT signaling pathway and G-protein coupled receptor signaling pathway, but also identified many other mechanisms which may likely contribute to T2D pathogenesis, such as positive regulation of lipid metabolic process and negative regulation of inflammatory response. Protein–protein interaction network analysis identified several interacted hub genes (protein), *e.g.* *KCNJ11*, *PPARG*, *ABCC8*, *THADA*, *WFS1*, *TCF7L2*, and *ADCY5*. Most of them have been reported to be associated with the pathogenesis of T2D or T2D-related metabolic pathways, *e.g.* gene *WFS1* is involved in the energy reserve metabolic process, glucose homeostasis and negative regulation of pancreatic β cells apoptotic process [69].

In the current study, we performed comprehensive bioinformatics analyses to demonstrate the functional importance of the T2D super enhancer SNPs. However, the databases used in these bioinformatics analyses are specialized, *e.g.* we only focus on SNPs from the 1000 genomes project with MAF > 0.001. We admit that a number of known variants satisfying the MAF criterion were missed in the functional prediction analysis. In additional, the results of functional annotation exclusively depend on computationally predicted regulation features, therefore the model and algorithm selection is critical for this kind of analysis. For example, the reliability of the transcription factor binding site predicted by SNP2TFBS is a function of the accuracy of the PWM model which is less accurate than recently proposed approaches [70]. Further, genomic sequence with higher affinity to a transcription factor is often insufficient for *in vivo* binding. A number of other factors may impact on binding such as chromatin accessibility [71]. Therefore, the functional importance of these candidate T2D super enhancer SNPs may be overshadowed, further experiment validation should be conduct to confirm the functional mechanism of these potential T2D super enhancer SNPs.

In conclusion, by integrating T2D GWAS and histone modification ChIP-seq data, we identified hundreds of T2D super enhancer SNPs. Comprehensive bioinformatics analyses of T2D super enhancer SNPs highlighted their potential functional importance in the regulation of T2D pathogenesis, which may yield novel insights into the genetic basis of T2D.

Supporting information

S1 Table. A detailed list of T2D-relevant cell & tissue types.

(XLSX)

S2 Table. Super enhancers in T2D-relevant cell & tissue types.

(XLSX)

S3 Table. The potential functional T2D super enhancer SNPs.

(XLSX)

S4 Table. The distribution of T2D super enhancer SNPs in different cell & tissue types.

(XLSX)

S5 Table. List of SNPs that map to the same super enhancer and LD between those SNPs.
(XLSX)

S6 Table. Functional prediction of T2D super enhancer SNPs by rVarBase.
(XLSX)

S7 Table. Functional prediction of T2D super enhancer SNPs by HaploReg.
(XLSX)

S8 Table. Long-range interaction analysis of T2D super enhancer SNPs.
(XLSX)

S9 Table. GO analysis of T2D super enhancer SNPs target genes.
(XLSX)

Acknowledgments

We thank the DIAGRAM Consortium and GEO project for generating and sharing the data used in this paper.

Author Contributions

Conceptualization: Shuai Liu.

Data curation: Jielong Tang.

Funding acquisition: Chunping Zeng.

Project administration: Chunping Zeng.

Resources: Juan Chen.

Software: Shuai Liu.

Supervision: Chunping Zeng.

Validation: Juan Chen.

Visualization: Daqing Deng.

Writing – original draft: Weiping Sun, Sihong Yao.

Writing – review & editing: Weiping Sun, Sihong Yao, Chunping Zeng.

References

1. American Diabetes A. Diagnosis and classification of diabetes mellitus. *Diabetes care*. 2009 Jan; 32 Suppl 1:S62–7. <https://doi.org/10.2337/dc09-S062> PMID: 19118289. Pubmed Central PMCID: 2613584.
2. Saddik B, Al-Dulajjan N. Diabetic patients' willingness to use tele-technology to manage their disease—A descriptive study. *Online journal of public health informatics*. 2015; 7(2):e214. <https://doi.org/10.5210/ojphi.v7i2.6011> PMID: 26284148. Pubmed Central PMCID: 4537853.
3. Santos FR, Bernardo V, Gabbay MA, Dib SA, Sigulem D. The impact of knowledge about diabetes, resilience and depression on glycemic control: a cross-sectional study among adolescents and young adults with type 1 diabetes. *Diabetology & metabolic syndrome*. 2013 Sep 28; 5(1):55. <https://doi.org/10.1186/1758-5996-5-55> PMID: 24289093. Pubmed Central PMCID: 3849685.
4. Willemsen G, Ward KJ, Bell CG, Christensen K, Bowden J, Dalgard C, et al. The Concordance and Heritability of Type 2 Diabetes in 34,166 Twin Pairs From International Twin Registers: The Discordant Twin (DISCOTWIN) Consortium. *Twin research and human genetics: the official journal of the International Society for Twin Studies*. 2015 Dec; 18(6):762–71. <https://doi.org/10.1017/thg.2015.83> PMID: 26678054.

5. Almgren P, Lehtovirta M, Isomaa B, Sarelin L, Taskinen MR, Lyssenko V, et al. Heritability and familiality of type 2 diabetes and related quantitative traits in the Botnia Study. *Diabetologia*. 2011 Nov; 54(11):2811–9. <https://doi.org/10.1007/s00125-011-2267-5> PMID: 21826484.
6. Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proceedings of the National Academy of Sciences of the United States of America*. 2009 Jun 09; 106(23):9362–7. <https://doi.org/10.1073/pnas.0903103106> PMID: 19474294. Pubmed Central PMCID: 2687147.
7. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, et al. Systematic localization of common disease-associated variation in regulatory DNA. *Science*. 2012 Sep 07; 337(6099):1190–5. <https://doi.org/10.1126/science.1222794> PMID: 22955828. Pubmed Central PMCID: 3771521.
8. Bulger M, Groudine M. Functional and mechanistic diversity of distal transcription enhancers. *Cell*. 2011 Feb 04; 144(3):327–39. <https://doi.org/10.1016/j.cell.2011.01.024> PMID: 21295696. Pubmed Central PMCID: 3742076.
9. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, et al. Super-enhancers in the control of cell identity and disease. *Cell*. 2013 Nov 07; 155(4):934–47. <https://doi.org/10.1016/j.cell.2013.09.053> PMID: 24119843. Pubmed Central PMCID: 3841062.
10. Guilherme A, Virbasius JV, Puri V, Czech MP. Adipocyte dysfunctions linking obesity to insulin resistance and type 2 diabetes. *Nat Rev Mol Cell Biol*. 2008 May; 9(5):367–77. <https://doi.org/10.1038/nrm2391> PMID: 18401346. Pubmed Central PMCID: 2886982.
11. Cusi K. The role of adipose tissue and lipotoxicity in the pathogenesis of type 2 diabetes. *Current diabetes reports*. 2010 Aug; 10(4):306–15. <https://doi.org/10.1007/s11892-010-0122-6> PMID: 20556549.
12. Kohlgruber A, Lynch L. Adipose tissue inflammation in the pathogenesis of type 2 diabetes. *Current diabetes reports*. 2015 Nov; 15(11):92. <https://doi.org/10.1007/s11892-015-0670-x> PMID: 26374569.
13. McLaughlin T, Liu LF, Lamendola C, Shen L, Morton J, Rivas H, et al. T-cell profile in adipose tissue is associated with insulin resistance and systemic inflammation in humans. *Arterioscler Thromb Vasc Biol*. 2014 Dec; 34(12):2637–43. <https://doi.org/10.1161/ATVBAHA.114.304636> PMID: 25341798. Pubmed Central PMCID: 4445971.
14. Cardoso CR, Ferreira MT, Leite NC, Salles GF. Prognostic impact of aortic stiffness in high-risk type 2 diabetic patients: the Rio de Janeiro Type 2 Diabetes Cohort Study. *Diabetes care*. 2013 Nov; 36(11):3772–8. <https://doi.org/10.2337/dc13-0506> PMID: 23877987. Pubmed Central PMCID: 3816863.
15. Lekva T, Bollerslev J, Norwitz ER, Aukrust P, Henriksen T, Ueland T. Aortic Stiffness and Cardiovascular Risk in Women with Previous Gestational Diabetes Mellitus. *PloS one*. 2015; 10(8):e0136892. <https://doi.org/10.1371/journal.pone.0136892> PMID: 26309121. Pubmed Central PMCID: 4550237.
16. Mansour AS, Yannoutsos A, Majahalmeh N, Agnoletti D, Safar ME, Ouerdane S, et al. Aortic stiffness and cardiovascular risk in type 2 diabetes. *Journal of hypertension*. 2013 Aug; 31(8):1584–92. <https://doi.org/10.1097/HJH.0b013e3283613074> PMID: 23822928.
17. Golbidi S, Laher I. Bladder dysfunction in diabetes mellitus. *Frontiers in pharmacology*. 2010; 1:136. <https://doi.org/10.3389/fphar.2010.00136> PMID: 21833175. Pubmed Central PMCID: 3153010.
18. Kebapci N, Yenilmez A, Efe B, Entok E, Demirustu C. Bladder dysfunction in type 2 diabetic patients. *Neurourology and urodynamics*. 2007; 26(6):814–9. <https://doi.org/10.1002/nau.20422> PMID: 17455274.
19. Karoli R, Bhat S, Fatima J, Priya S. A study of bladder dysfunction in women with type 2 diabetes mellitus. *Indian journal of endocrinology and metabolism*. 2014 Jul; 18(4):552–7. <https://doi.org/10.4103/2230-8210.137518> PMID: 25143916. Pubmed Central PMCID: 4138915.
20. Targher G, Bertolini L, Poli F, Rodella S, Scala L, Tessari R, et al. Nonalcoholic fatty liver disease and risk of future cardiovascular events among type 2 diabetic patients. *Diabetes*. 2005 Dec; 54(12):3541–6. PMID: 16306373.
21. Targher G, Bertolini L, Padovani R, Rodella S, Tessari R, Zenari L, et al. Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. *Diabetes care*. 2007 May; 30(5):1212–8. <https://doi.org/10.2337/dc06-2247> PMID: 17277038.
22. Leite NC, Salles GF, Araujo AL, Villela-Nogueira CA, Cardoso CR. Prevalence and associated factors of non-alcoholic fatty liver disease in patients with type-2 diabetes mellitus. *Liver international: official journal of the International Association for the Study of the Liver*. 2009 Jan; 29(1):113–9. <https://doi.org/10.1111/j.1478-3231.2008.01718.x> PMID: 18384521.
23. Wang Q, Zhai X, Chen X, Lu J, Zhang Y, Huang Q. Dysregulation of circulating CD4+CXCR5+ T cells in type 2 diabetes mellitus. *APMIS: acta pathologica, microbiologica, et immunologica Scandinavica*. 2015 Feb; 123(2):146–51. <https://doi.org/10.1111/apm.12330> PMID: 25351463.
24. Xia C, Rao X, Zhong J. Role of T Lymphocytes in Type 2 Diabetes and Diabetes-Associated Inflammation. *Journal of diabetes research*. 2017; 2017:6494795. <https://doi.org/10.1155/2017/6494795> PMID: 28251163. Pubmed Central PMCID: 5307004 publication of this paper.

25. Kumar NP, Sridhar R, Nair D, Banurekha VV, Nutman TB, Babu S. Type 2 diabetes mellitus is associated with altered CD8(+) T and natural killer cell function in pulmonary tuberculosis. *Immunology*. 2015 Apr; 144(4):677–86. <https://doi.org/10.1111/imm.12421> PMID: 25363329. Pubmed Central PMCID: 4368174.
26. Cnop M, Welsh N, Jonas JC, Jorns A, Lenzen S, Eizirik DL. Mechanisms of pancreatic beta-cell death in type 1 and type 2 diabetes: many differences, few similarities. *Diabetes*. 2005 Dec; 54 Suppl 2:S97–107. PMID: 16306347.
27. Butler AE, Janson J, Bonner-Weir S, Ritzel R, Rizza RA, Butler PC. Beta-cell deficit and increased beta-cell apoptosis in humans with type 2 diabetes. *Diabetes*. 2003 Jan; 52(1):102–10. PMID: 12502499.
28. Lorenzo A, Razzaboni B, Weir GC, Yankner BA. Pancreatic islet cell toxicity of amylin associated with type-2 diabetes mellitus. *Nature*. 1994 Apr 21; 368(6473):756–60. <https://doi.org/10.1038/368756a0> PMID: 8152488.
29. Del Guerra S, Lupi R, Marselli L, Masini M, Bugliani M, Sbrana S, et al. Functional and molecular defects of pancreatic islets in human type 2 diabetes. *Diabetes*. 2005 Mar; 54(3):727–35. PMID: 15734849.
30. Dworacki G, Urazayev O, Bekmukhambetov Y, Iskakova S, Frycz BA, Jagodzinski PP, et al. Thymic emigration patterns in patients with type 2 diabetes treated with metformin. *Immunology*. 2015 Nov; 146(3):456–69. <https://doi.org/10.1111/imm.12522> PMID: 26271466. Pubmed Central PMCID: 4610634.
31. Seeley RJ, Sandoval DA. Targeting the brain as a cure for type 2 diabetes. *Nat Med*. 2016 Jul 07; 22(7):709–11. <https://doi.org/10.1038/nm.4137> PMID: 27387884.
32. Moran C, Phan TG, Chen J, Blizzard L, Beare R, Venn A, et al. Brain atrophy in type 2 diabetes: regional distribution and influence on cognition. *Diabetes care*. 2013 Dec; 36(12):4036–42. <https://doi.org/10.2337/dc13-0143> PMID: 23939539. Pubmed Central PMCID: 3836136.
33. Korf ES, White LR, Scheltens P, Launer LJ. Brain aging in very old men with type 2 diabetes: the Honolulu-Asia Aging Study. *Diabetes care*. 2006 Oct; 29(10):2268–74. <https://doi.org/10.2337/dc06-0243> PMID: 17003305.
34. Krabbe KS, Nielsen AR, Krogh-Madsen R, Plomgaard P, Rasmussen P, Erikstrup C, et al. Brain-derived neurotrophic factor (BDNF) and type 2 diabetes. *Diabetologia*. 2007 Feb; 50(2):431–8. <https://doi.org/10.1007/s00125-006-0537-4> PMID: 17151862.
35. Ehnert S, Freude T, Ihle C, Mayer L, Braun B, Graeser J, et al. Factors circulating in the blood of type 2 diabetes mellitus patients affect osteoblast maturation—description of a novel in vitro model. *Exp Cell Res*. 2015 Mar 15; 332(2):247–58. <https://doi.org/10.1016/j.yexcr.2014.12.011> PMID: 25557875.
36. Scott LJ, Erdos MR, Huyghe JR, Welch RP, Beck AT, Wolford BN, et al. The genetic regulatory signature of type 2 diabetes in human skeletal muscle. *Nature communications*. 2016 Jun 29; 7:11764. <https://doi.org/10.1038/ncomms11764> PMID: 27353450. Pubmed Central PMCID: 4931250.
37. Jackson S, Bagstaff SM, Lynn S, Yeaman SJ, Turnbull DM, Walker M. Decreased insulin responsiveness of glucose uptake in cultured human skeletal muscle cells from insulin-resistant nondiabetic relatives of type 2 diabetic families. *Diabetes*. 2000 Jul; 49(7):1169–77. PMID: 10909975.
38. Dietze D, Koenen M, Rohrig K, Horikoshi H, Hauner H, Eckel J. Impairment of insulin signaling in human skeletal muscle cells by co-culture with human adipocytes. *Diabetes*. 2002 Aug; 51(8):2369–76. PMID: 12145147.
39. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008; 9(9):R137. <https://doi.org/10.1186/gb-2008-9-9-r137> PMID: 18798982. Pubmed Central PMCID: 2592715.
40. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*. 2013 Apr 11; 153(2):307–19. <https://doi.org/10.1016/j.cell.2013.03.035> PMID: 23582322. Pubmed Central PMCID: 3653129.
41. Guo L, Du Y, Qu S, Wang J. rVarBase: an updated database for regulatory features of human variants. *Nucleic Acids Res*. 2016 Jan 04; 44(D1):D888–93. <https://doi.org/10.1093/nar/gkv1107> PMID: 26503253. Pubmed Central PMCID: 4702808.
42. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res*. 2012 Jan; 40(Database issue):D930–4. <https://doi.org/10.1093/nar/gkr917> PMID: 22064851. Pubmed Central PMCID: 3245002.
43. Li MJ, Wang LY, Xia Z, Sham PC, Wang J. GWAS3D: Detecting human regulatory variants by integrative analysis of genome-wide associations, chromosome interactions and histone modifications. *Nucleic Acids Res*. 2013 Jul; 41(Web Server issue):W150–8. <https://doi.org/10.1093/nar/gkt456> PMID: 23723249. Pubmed Central PMCID: 3692118.

44. Daniel Yekutieli YB. Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *Journal of Statistical Planning and Inference*. 1999; 82(1–2):171–96.
45. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguéz P, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res*. 2011 Jan; 39(Database issue):D561–8. <https://doi.org/10.1093/nar/gkq973> PMID: 21045058. Pubmed Central PMCID: 3013807.
46. Jager J, Gremeaux T, Cormont M, Le Marchand-Brustel Y, Tanti JF. Interleukin-1beta-induced insulin resistance in adipocytes through down-regulation of insulin receptor substrate-1 expression. *Endocrinology*. 2007 Jan; 148(1):241–51. <https://doi.org/10.1210/en.2006-0692> PMID: 17038556. Pubmed Central PMCID: 1971114.
47. Iguchi H, Urashima Y, Inagaki Y, Ikeda Y, Okamura M, Tanaka T, et al. SOX6 suppresses cyclin D1 promoter activity by interacting with beta-catenin and histone deacetylase 1, and its down-regulation induces pancreatic beta-cell proliferation. *The Journal of biological chemistry*. 2007 Jun 29; 282(26):19052–61. <https://doi.org/10.1074/jbc.M700460200> PMID: 17412698.
48. Cunha DA, Ladriere L, Ortis F, Igoillo-Esteve M, Gurzov EN, Lupi R, et al. Glucagon-like peptide-1 agonists protect pancreatic beta-cells from lipotoxic endoplasmic reticulum stress through upregulation of BiP and JunB. *Diabetes*. 2009 Dec; 58(12):2851–62. <https://doi.org/10.2337/db09-0685> PMID: 19720788. Pubmed Central PMCID: 2780890.
49. Shaw-Smith C, De Franco E, Lango Allen H, Battle M, Flanagan SE, Borowiec M, et al. GATA4 mutations are a cause of neonatal and childhood-onset diabetes. *Diabetes*. 2014 Aug; 63(8):2888–94. <https://doi.org/10.2337/db14-0061> PMID: 24696446.
50. Zhou JB, Yang JK, Zhang BH, Lu J. Interaction of Wnt pathway related variants with type 2 diabetes in a Chinese Han population. *PeerJ*. 2015; 3:e1304. <https://doi.org/10.7717/peerj.1304> PMID: 26509107. Pubmed Central PMCID: 4621788.
51. Reimann F, Gribble FM. G protein-coupled receptors as new therapeutic targets for type 2 diabetes. *Diabetologia*. 2016 Feb; 59(2):229–33. <https://doi.org/10.1007/s00125-015-3825-z> PMID: 26661410. Pubmed Central PMCID: 4705132.
52. Prentki M, Matschinsky FM. Ca²⁺, cAMP, and phospholipid-derived messengers in coupling mechanisms of insulin secretion. *Physiological reviews*. 1987 Oct; 67(4):1185–248. <https://doi.org/10.1152/physrev.1987.67.4.1185> PMID: 2825225.
53. Bennett V, Baines AJ. Spectrin and ankyrin-based pathways: metazoan inventions for integrating cells into tissues. *Physiological reviews*. 2001 Jul; 81(3):1353–92. <https://doi.org/10.1152/physrev.2001.81.3.1353> PMID: 11427698.
54. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature genetics*. 2012 Sep; 44(9):981–90. <https://doi.org/10.1038/ng.2383> PMID: 22885922. Pubmed Central PMCID: 3442244.
55. Imamura M, Maeda S, Yamauchi T, Hara K, Yasuda K, Morizono T, et al. A single-nucleotide polymorphism in ANK1 is associated with susceptibility to type 2 diabetes in Japanese populations. *Human molecular genetics*. 2012 Jul 01; 21(13):3042–9. <https://doi.org/10.1093/hmg/dds113> PMID: 22456796.
56. Huang P, Chandra V, Rastinejad F. Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics. *Annual review of physiology*. 2010; 72:247–72. <https://doi.org/10.1146/annurev-physiol-021909-135917> PMID: 20148675. Pubmed Central PMCID: 3677810.
57. Werman A, Hollenberg A, Solanes G, Bjorbaek C, Vidal-Puig AJ, Flier JS. Ligand-independent activation domain in the N terminus of peroxisome proliferator-activated receptor gamma (PPARgamma). Differential activity of PPARgamma1 and -2 isoforms and influence of insulin. *The Journal of biological chemistry*. 1997 Aug 08; 272(32):20230–5. PMID: 9242701.
58. Cauchi S, Meyre D, Dina C, Choquet H, Samson C, Gallina S, et al. Transcription factor TCF7L2 genetic study in the French population: expression in human beta-cells and adipose tissue and strong association with type 2 diabetes. *Diabetes*. 2006 Oct; 55(10):2903–8. <https://doi.org/10.2337/db06-0474> PMID: 17003360.
59. Miyake K, Horikawa Y, Hara K, Yasuda K, Osawa H, Furuta H, et al. Association of TCF7L2 polymorphisms with susceptibility to type 2 diabetes in 4,087 Japanese subjects. *J Hum Genet*. 2008; 53(2):174–80. <https://doi.org/10.1007/s10038-007-0231-5> PMID: 18097733.
60. Yang Y, Xu JR, Wang YJ, Liu XM. Association of TCF7L2 gene polymorphisms with susceptibility to type 2 diabetes mellitus in a Chinese Hui population. *Genetics and molecular research: GMR*. 2015 Aug 21; 14(3):10064–71. <https://doi.org/10.4238/2015.August.21.13> PMID: 26345943.
61. Klaus A, Birchmeier W. Wnt signalling and its impact on development and cancer. *Nature reviews Cancer*. 2008 May; 8(5):387–98. <https://doi.org/10.1038/nrc2389> PMID: 18432252.

62. Shu L, Sauter NS, Schulthess FT, Matveyenko AV, Oberholzer J, Maedler K. Transcription factor 7-like 2 regulates beta-cell survival and function in human pancreatic islets. *Diabetes*. 2008 Mar; 57(3):645–53. <https://doi.org/10.2337/db07-0847> PMID: 18071026.
63. Khan IU, Beck-Sickinger AG. Targeted tumor diagnosis and therapy with peptide hormones as radio-pharmaceuticals. *Anti-cancer agents in medicinal chemistry*. 2008 Feb; 8(2):186–99. PMID: 18288921.
64. Sandhu MS, Weedon MN, Fawcett KA, Wasson J, Debenham SL, Daly A, et al. Common variants in WFS1 confer risk of type 2 diabetes. *Nature genetics*. 2007 Aug; 39(8):951–3. <https://doi.org/10.1038/ng2067> PMID: 17603484. Pubmed Central PMCID: 2672152.
65. Lozzio CB, Lozzio BB. Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood*. 1975 Mar; 45(3):321–34. PMID: 163658.
66. Jin T, Liu L. The Wnt signaling pathway effector TCF7L2 and type 2 diabetes mellitus. *Mol Endocrinol*. 2008 Nov; 22(11):2383–92. <https://doi.org/10.1210/me.2008-0135> PMID: 18599616.
67. Siti SD AF, Marie NG, Alan KB, Robert H, Alex T, Richard LD. Identification of the Wnt Signalling Protein, TCF7L2 As a Significantly Overexpressed Transcription Factor in AML. *Blood*. 2012; 120:1281.
68. Choi SS, Kim ES, Jung JE, Marciano DP, Jo A, Koo JY, et al. PPARgamma Antagonist Gleevec Improves Insulin Sensitivity and Promotes the Browning of White Adipose Tissue. *Diabetes*. 2016 Apr; 65(4):829–39. <https://doi.org/10.2337/db15-1382> PMID: 26740599. Pubmed Central PMCID: 5314706.
69. Fu Z, Gilbert ER, Liu D. Regulation of insulin synthesis and secretion and pancreatic Beta-cell dysfunction in diabetes. *Current diabetes reviews*. 2013 Jan 01; 9(1):25–53. PMID: 22974359. Pubmed Central PMCID: 3934755.
70. Kumar S, Ambrosini G, Bucher P. SNP2TFBS—a database of regulatory SNPs affecting predicted transcription factor binding site affinity. *Nucleic Acids Res*. 2017 Jan 04; 45(D1):D139–D44. <https://doi.org/10.1093/nar/gkw1064> PMID: 27899579. Pubmed Central PMCID: 5210548.
71. Slattery M, Zhou T, Yang L, Dantas Machado AC, Gordan R, Rohs R. Absence of a simple code: how transcription factors read the genome. *Trends Biochem Sci*. 2014 Sep; 39(9):381–99. <https://doi.org/10.1016/j.tibs.2014.07.002> PMID: 25129887. Pubmed Central PMCID: 4149858.