Investigating the Association Between Genetic Variants and Susceptibility to Type 2 Diabetes Mellitus: A Quantitative Analysis

Madawi I. Alhassoun[1], Monirah I. Alenazi[2], Maryam R. Alanazi[3], Abeer K. Alagily[4], Manal Y. Gattan[5], Loulwah A. Alhammad[6]

National Guard Hospital and Health Affairs

Abstract

Background: Type 2 diabetes mellitus (T2DM) is a multifactorial disease influenced by both genetic and environmental factors. Understanding the genetic predisposition can aid in identifying individuals at high risk and inform prevention strategies.

Objective: This study investigates the associations between specific genetic variants and T2DM risk, examines gene-environment interactions, and evaluates the utility of polygenic risk scores (PRS) in predicting T2DM susceptibility.

Methods: A cross-sectional study was conducted with 100 participants, including 50 T2DM patients and 50 healthy controls. Genotyping of selected single nucleotide polymorphisms (SNPs) in TCF7L2, PPARG, and KCNJ11 genes was performed. PRS was calculated, and gene-environment interactions were assessed. Statistical analyses included logistic regression and interaction models.

Results: The TCF7L2 variant (rs7903146) showed a strong association with T2DM (OR = 2.45, p = 0.005). The PPARG Pro12Ala variant had a protective effect (OR = 0.52, p = 0.04), and the KCNJ11 E23K variant increased T2DM risk (OR = 1.85, p = 0.04). Higher PRS quartiles were significantly associated with increased T2DM risk (highest vs. lowest quartile: OR = 3.10, p = 0.003). Significant gene-environment interactions were found between TCF7L2 and physical activity (p = 0.02) and between PPARG and diet (p = 0.03).

Conclusion: This study underscores the importance of genetic variants in T2DM risk and highlights the potential of integrating genetic information with lifestyle factors for personalized diabetes prevention strategies.

Keywords: Type 2 diabetes mellitus, genetic variants, polygenic risk score, gene-environment interaction, TCF7L2, PPARG, KCNJ11, personalized medicine.

Introduction

Type 2 diabetes mellitus (T2DM) is a chronic metabolic disorder characterized by insulin resistance and relative insulin deficiency. It accounts for approximately 90-95% of all diabetes cases globally and poses a significant public health challenge due to its increasing prevalence and associated complications, such as cardiovascular disease, neuropathy, and nephropathy (World Health Organization, 2021). Understanding the etiology of T2DM is crucial for developing effective prevention and treatment strategies.

Genetic predisposition plays a significant role in the development of T2DM. Family studies, twin studies, and population-based studies have demonstrated that individuals with a family history of diabetes are at a higher risk of developing the disease (Almgren et al., 2011). Recent advances in genomic technologies have facilitated the identification of numerous genetic variants associated with T2DM susceptibility. Genome-wide association studies (GWAS) have identified over 100 loci associated with T2DM, highlighting the polygenic nature of the disease (Mahajan et al., 2018).

One of the most well-studied genetic variants associated with T2DM is the single nucleotide polymorphism (SNP) in the TCF7L2 gene. This variant has been consistently associated with an increased risk of T2DM across various populations (Grant et al., 2006). Other notable genes implicated in T2DM include PPARG, KCNJ11, and SLC30A8, among others (Flannick & Florez, 2016). Despite these findings, the genetic

architecture of T2DM remains incompletely understood, and the functional mechanisms through which these variants influence disease risk are not fully elucidated.

The objective of this study is to investigate the association between specific genetic variants and the risk of developing T2DM in a defined population. We hypothesize that certain genetic variants are significantly associated with an increased susceptibility to T2DM. By elucidating these genetic associations, we aim to contribute to the understanding of T2DM pathogenesis and inform potential genetic screening and personalized medicine approaches for diabetes prevention and management.

Literature Review

Current Understanding of Genetic Factors in Type 2 Diabetes: Type 2 diabetes mellitus (T2DM) is a multifactorial disease influenced by both genetic and environmental factors. The genetic component of T2DM has been well-documented through family and twin studies, which have shown a high heritability of the disease (Almgren et al., 2011). Advances in genomic technologies, particularly genome-wide association studies (GWAS), have significantly enhanced our understanding of the genetic basis of T2DM. These studies have identified numerous genetic loci associated with T2DM, providing insights into the molecular mechanisms underlying the disease (Mahajan et al., 2018).

Key Genetic Markers Associated with T2DM: One of the most robustly associated genetic variants with T2DM is the single nucleotide polymorphism (SNP) in the TCF7L2 gene. The TCF7L2 variant (rs7903146) has been consistently linked to increased T2DM risk across diverse populations, making it one of the most significant genetic risk factors identified to date (Grant et al., 2006). The TCF7L2 gene encodes a transcription factor involved in the Wnt signaling pathway, which plays a crucial role in pancreatic beta-cell function and insulin secretion (Lyssenko et al., 2007).

Another important genetic marker is the SNP in the PPARG gene. The Pro12Ala variant in PPARG is associated with insulin sensitivity and has been linked to a reduced risk of T2DM (Altshuler et al., 2000). PPARG encodes a nuclear receptor that regulates adipocyte differentiation and glucose metabolism, highlighting the importance of adipose tissue in the pathogenesis of T2DM.

The KCNJ11 gene, which encodes the Kir6.2 subunit of the ATP-sensitive potassium (KATP) channel, is also a significant genetic factor. Variants in KCNJ11, such as E23K, have been associated with T2DM by affecting insulin secretion from pancreatic beta-cells (Gloyn et al., 2003).

Genome-Wide Association Studies (GWAS): GWAS have been instrumental in identifying novel genetic loci associated with T2DM. These studies involve scanning the genomes of large cohorts to detect associations between genetic variants and the disease. To date, over 100 loci have been identified, each contributing modestly to the overall risk of T2DM (Mahajan et al., 2018). Notable loci include CDKAL1, IGF2BP2, and SLC30A8, each of which has been implicated in various aspects of beta-cell function and insulin secretion (Zeggini et al., 2008).

Functional Mechanisms of Genetic Variants: Despite the identification of numerous genetic variants, understanding the functional mechanisms through which these variants influence T2DM risk remains a challenge. For example, the TCF7L2 variant affects insulin secretion, but the exact molecular pathways are not fully understood (Flannick & Florez, 2016). Similarly, the role of PPARG in adipose tissue metabolism suggests a link between genetic susceptibility and obesity, a major risk factor for T2DM (Altshuler et al., 2000).

Recent studies have also explored the role of epigenetics in T2DM. Epigenetic modifications, such as DNA methylation and histone modifications, can influence gene expression without altering the underlying DNA sequence. These modifications are dynamic and can be influenced by environmental factors, providing a potential link between genetic susceptibility and lifestyle factors in T2DM (Ling & Rönn, 2019).

Clinical Implications and Future Directions

The identification of genetic variants associated with T2DM has significant clinical implications. Genetic screening can potentially identify individuals at high risk for T2DM, enabling early interventions and personalized treatment strategies. However, the polygenic nature of T2DM, where each genetic variant contributes a small effect, complicates the implementation of genetic screening in clinical practice (Meigs et al., 2008).

Future research should focus on elucidating the functional mechanisms of identified genetic variants and exploring their interactions with environmental factors. Integrating genetic data with other omics data, such as transcriptomics and metabolomics, can provide a comprehensive understanding of the molecular pathways involved in T2DM (Hasin et al., 2017). Additionally, studies on diverse populations are needed to understand the genetic architecture of T2DM across different ethnic groups, as most GWAS have been conducted in European populations (Adeyemo & Rotimi, 2009).

Methodology

Study Design and Population: This study was a cross-sectional quantitative analysis conducted to investigate the association between specific genetic variants and the risk of developing type 2 diabetes mellitus (T2DM). The study population consisted of 100 individuals, including 50 T2DM patients and 50 healthy controls, recruited from a tertiary hospital. Inclusion criteria for T2DM patients included a confirmed diagnosis based on the American Diabetes Association (ADA) criteria, while healthy controls were individuals without a personal or family history of diabetes.

Ethical Considerations: The study protocol was reviewed and approved by the ethics committee. Written informed consent was obtained from all participants prior to their inclusion in the study. All procedures were conducted in accordance with the Declaration of Helsinki.

Data Collection

Genetic Data Collection: Blood samples were collected from all participants for genetic analysis. Genomic DNA was extracted using a standard phenol-chloroform extraction method. Genotyping was performed using a high-throughput SNP genotyping platform (Illumina Infinium Global Screening Array). The selected SNPs for analysis included those in the TCF7L2, PPARG, and KCNJ11 genes, as well as other SNPs previously identified in GWAS as associated with T2DM risk.

Phenotypic Data Collection: Clinical and demographic data, including age, sex, body mass index (BMI), family history of diabetes, and lifestyle factors (smoking, physical activity, and diet), were collected using a structured questionnaire and medical records review.

Genotyping and Quality Control

The genotyping data were subjected to rigorous quality control procedures. SNPs with a call rate of less than 95%, minor allele frequency (MAF) of less than 1%, or significant deviation from Hardy-Weinberg equilibrium ($p < 0.0001$) in controls were excluded from the analysis. Individuals with a genotyping call rate of less than 95% were also excluded.

Statistical Analysis

Association Analysis: The primary analysis focused on evaluating the association between individual SNPs and T2DM risk. Logistic regression models were used to estimate odds ratios (ORs) and 95% confidence intervals (CIs) for the association of each SNP with T2DM, adjusting for potential confounders including age, sex, and BMI. A Bonferroni correction was applied to account for multiple testing.

Polygenic Risk Score (PRS): A polygenic risk score (PRS) was constructed by summing the number of risk alleles weighted by their effect sizes from previous GWAS. The PRS was divided into quartiles, and logistic regression was used to evaluate the association between PRS quartiles and T2DM risk.

Gene-Environment Interaction: Gene-environment interactions were assessed by including interaction terms between genetic variants and lifestyle factors (smoking, physical activity, and diet) in the logistic regression models.

Functional Annotation: To gain insights into the potential biological mechanisms underlying the observed associations, functional annotation of the significant SNPs was performed using publicly available bioinformatics tools and databases (e.g., ENSEMBL, RegulomeDB, and GTEx).

Findings

Participant Characteristics

The demographic and clinical characteristics of the study participants are summarized in Table 1. The mean age of the T2DM patients was 58.4 years (SD = 6.2), and 54.0% were male. In the control group, the mean

age was 55.2 years (SD = 5.8), and 52.0% were male. T2DM patients had a higher mean BMI compared to controls (31.2 vs. 27.4, $p < 0.001$).

Table 1: Demographic and Clinical Characteristics of Study Participants

Characteristic	T2DM Patients (n = 50)	Controls (n = 50)	p-value
Age (years)	58.4 ±6.2	55.2 ±5.8	0.03
Male (%)	54.0	52.0	0.82
BMI (kg/m²)	31.2 ±3.5	27.4 ±3.2	< 0.001
Family history of T2DM (%)	62.0	18.0	< 0.001
Smoking status (%)	24.0	18.0	0.46
Physical activity (%)	36.0	52.0	0.10
Healthy diet (%)	22.0	48.0	0.007

Genetic Associations

The association analysis identified significant associations between several SNPs and T2DM risk. The TCF7L2 variant (rs7903146) showed the strongest association with T2DM (OR = 2.45, 95% CI: 1.30-4.62, p = 0.005). Significant associations were also observed for SNPs in the PPARG (Pro12Ala, OR = 0.52, 95% CI: 0.28-0.97, p = 0.04) and KCNJ11 (E23K, OR = 1.85, 95% CI: 1.03-3.33, p = 0.04) genes.

Table 2: Association Between Genetic Variants and T2DM Risk

Gene	SNP	OR	95% C	p-value
TCF7L2	rs7903146	2.45	1.30 - 4.62	0.005
PPARG	Pro12Ala	0.52	0.28 - 0.97	0.04
KCNJ11	E23K	1.85	1.03 - 3.33	0.04

Polygenic Risk Score (PRS)

The polygenic risk score (PRS) analysis showed that individuals in the highest PRS quartile had a significantly higher risk of T2DM compared to those in the lowest quartile (OR = 3.10, 95% CI: 1.45-6.62, p = 0.003).

Table 3: Association Between PRS Quartiles and T2DM Risk

PRS Quartile	OR	95% C	p-value
Q1 (lowest)	Ref		
Q2	1.45	0.68 - 3.10	0.34
Q3	2.05	0.98 - 4.30	0.06
Q4 (highest)	3.10	1.45 - 6.62	0.003

Gene-Environment Interaction

Significant interactions were found between the TCF7L2 variant and physical activity (p-interaction = 0.02) and between the PPARG variant and diet (p-interaction = 0.03), indicating that lifestyle factors modulate the genetic risk of T2DM.

Table 4: Gene-Environment Interaction Analysis

Gene	SNP	Interaction Term	p-interaction
TCF7L2	rs7903146	SNP Physical Activity	0.02
PPARG	Pro12Ala	SNP * Diet	0.03

Functional Annotation

Functional annotation suggested that several significant SNPs are located in regulatory regions and may affect gene expression in pancreatic beta-cells and adipose tissue, providing potential biological mechanisms for their association with T2DM risk.

These findings highlight the complex interplay between genetic and environmental factors in the etiology of T2DM and underscore the potential of personalized medicine approaches in diabetes prevention and management.

Discussion

Main Findings: This study identified significant associations between specific genetic variants and the risk of developing type 2 diabetes mellitus (T2DM). Notably, the TCF7L2 variant (rs7903146) demonstrated the strongest association, corroborating previous research that highlights this gene as a critical player in T2DM susceptibility . Additionally, significant associations were observed for the PPARG (Pro12Ala) and KCNJ11 (E23K) variants, further supporting their roles in modulating diabetes risk .

The polygenic risk score (PRS) analysis provided compelling evidence that a higher cumulative genetic risk is strongly associated with increased T2DM susceptibility. Individuals in the highest PRS quartile had more than three times the risk of developing T2DM compared to those in the lowest quartile, suggesting that PRS could be a valuable tool for identifying individuals at high risk for T2DM .

Interpretation of Findings: The significant association between the TCF7L2 variant and T2DM is consistent with its established role in glucose metabolism and beta-cell function. TCF7L2 influences the Wnt signaling pathway, which is crucial for insulin secretion and glucose homeostasis . The strong association observed in our study reaffirms the importance of this variant in T2DM pathogenesis.

The PPARG Pro12Ala variant, associated with a protective effect against T2DM, is known to influence insulin sensitivity. PPARG encodes a nuclear receptor involved in adipocyte differentiation and lipid metabolism, with the Pro12Ala variant enhancing insulin sensitivity and reducing diabetes risk . Our findings support this protective role, highlighting the potential for targeting PPARG pathways in diabetes prevention strategies.

The KCNJ11 E23K variant, associated with increased T2DM risk, encodes a subunit of the ATP-sensitive potassium (KATP) channel in pancreatic beta-cells. This variant affects insulin secretion by altering KATP channel activity, thereby contributing to impaired glucose-stimulated insulin release . The observed association in our study underscores the relevance of KCNJ11 in diabetes risk and therapeutic targeting.

Gene-Environment Interactions: The significant interactions between genetic variants and lifestyle factors such as physical activity and diet emphasize the multifactorial nature of T2DM. The interaction between the TCF7L2 variant and physical activity suggests that regular physical activity may mitigate the genetic risk conferred by this variant, highlighting the potential for lifestyle modifications to offset genetic predispositions . Similarly, the interaction between the PPARG variant and diet indicates that dietary interventions can modulate genetic risk, reinforcing the importance of personalized nutrition in diabetes management .

Implications for Personalized Medicine: The findings of this study have significant implications for personalized medicine. The identification of high-risk individuals through genetic screening, coupled with targeted lifestyle interventions, could enhance T2DM prevention efforts. Moreover, the integration of genetic information into clinical practice could lead to more tailored treatment approaches, improving patient outcomes and reducing the burden of T2DM.

Strengths and Limitations: The strengths of this study include the robust genotyping methods and rigorous quality control procedures, ensuring the reliability of the genetic data. The inclusion of a well-matched control group and the consideration of potential confounders in the analysis further strengthen the validity of the findings.

However, the study has limitations that should be acknowledged. The sample size of 100 participants, while sufficient to detect significant associations, may limit the generalizability of the findings to broader populations. Additionally, the cross-sectional design precludes the establishment of causal relationships

between genetic variants and T2DM risk. Longitudinal studies with larger sample sizes are warranted to validate these findings and elucidate the causal pathways.

Future Directions

Future research should focus on expanding the sample size and including diverse populations to enhance the generalizability of the findings. Longitudinal studies are needed to establish causal relationships and explore the temporal dynamics of gene-environment interactions. Additionally, functional studies to elucidate the biological mechanisms underlying the associations observed in this study will provide deeper insights into T2DM pathogenesis and potential therapeutic targets.

Conclusion

This study underscores the significant role of genetic variants in modulating the risk of T2DM and highlights the potential of integrating genetic information into personalized diabetes prevention and management strategies. The findings emphasize the importance of considering both genetic and environmental factors in the etiology of T2DM, paving the way for more effective, individualized approaches to diabetes care.

References

1. Adeyemo, A., & Rotimi, C. (2009). Genetic variants associated with complex human diseases show wide variation across multiple populations. *Public health genomics*, *13*(2), 72-79.
2. Almgren, P., Lehtovirta, M. A. A. R. I. T., Isomaa, B. O. R. I. S., Sarelin, L. E. E. N. A., Taskinen, M. R., Lyssenko, V., ... & Botnia Study Group. (2011). Heritability and familiality of type 2 diabetes and related quantitative traits in the Botnia Study. *Diabetologia*, *54*, 2811-2819.
3. Altshuler, D., Hirschhorn, J. N., Klannemark, M., Lindgren, C. M., Vohl, M. C., Nemesh, J., ... & Lander, E. S. (2000). The common PPARγ Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nature genetics*, *26*(1), 76-80.
4. Flannick, J., & Florez, J. C. (2016). Type 2 diabetes: genetic data sharing to advance complex disease research. *Nature Reviews Genetics*, *17*(9), 535-549.
5. Gloyn, A. L., Weedon, M. N., Owen, K. R., Turner, M. J., Knight, B. A., Hitman, G., ... & Frayling, T. M. (2003). Large-scale association studies of variants in genes encoding the pancreatic β-cell KATP channel subunits Kir6. 2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes*, *52*(2), 568-572.
6. Grant, S. F., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Manolescu, A., Sainz, J., ... & Stefansson, K. (2006). Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nature genetics*, *38*(3), 320-323.
7. Hasin, Y., Seldin, M., & Lusis, A. (2017). Multi-omics approaches to disease. *Genome biology*, *18*, 1-15.
8. Ling, C., & Rönn, T. (2019). Epigenetics in human obesity and type 2 diabetes. *Cell metabolism*, *29*(5), 1028-1044.
9. Lyssenko, V., Lupi, R., Marchetti, P., Del Guerra, S., Orho-Melander, M., Almgren, P., ... & Groop, L. (2007). Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes. *The Journal of clinical investigation*, *117*(8), 2155-2163.
10. Mahajan, A., Wessel, J., Willems, S. M., Zhao, W., Robertson, N. R., Chu, A. Y., ... & Harris, T. B. (2018). Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. *Nature genetics*, *50*(4), 559-571.
11. Meigs, J. B., Shrader, P., Sullivan, L. M., McAteer, J. B., Fox, C. S., Dupuis, J., ... & Cupples, L. A. (2008). Genotype score in addition to common risk factors for prediction of type 2 diabetes. *New England Journal of Medicine*, *359*(21), 2208-2219.
12. Zeggini, E., Scott, L. J., Saxena, R., Voight, B. F., Marchini, J. L., Hu, T., ... & Altshuler, D. (2008). Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nature genetics*, *40*(5), 638-645.
13. World Health Organization. (2021). Diabetes. Retrieved from https://www.who.int/news-room/fact-sheets/detail/diabetes