Title: Sex-specific genetic loci linked to early and late onset type 2 diabetes

Short title: Sex-specific genetic loci in T2D

Jaime Berumen1*, Lorena Orozco2, Héctor Gallardo-Rincón3,4*, Rosa Elba Benuto5,
Espiridión Ramos-Martínez1, Fernando Rivas1, Humberto García-Ortiz2, Melissa Marín-
Medina5, Elizabeth Barrera1, Eligia Juárez-Torres5, Anabel Alvarado Silva5, Luis Alberto
Martínez-Juárez4, Julieta Lomelín-Gascón4, Alejandra Montoya4, Janinne Ortega-Montiel4,
Diego-Abelardo Alvarez-Hernández4, Roberto Tapia-Conyer6

1. Unidad de Investigación en Medicina Experimental, Facultad de Medicina, Universidad
Nacional Autónoma de México, Ciudad de México, México.
2. Instituto Nacional de Medicina Genómica, Ciudad de México, México.
3. Universidad of Guadalajara, Health Sciences University Center, Guadalajara, Jalisco,
México
4. Fundación Carlos Slim, Ciudad de México, México.
5. Laboratorio HuellaGénica, Ciudad de México, México.
6. Facultad de Medicina, Universidad Nacional Autónoma de México, Ciudad de México,
México

*Corresponding authors:
Jaime Berumen
Unidad de Investigación en Medicina Experimental, Hospital General de México. Dr. Balmis
148, Doctores, Cuauhtémoc, 06720 Ciudad de México, CDMX, México
Phone: 52 55 1850 4023
Email: jaimeberumen47@gmail.com
Héctor Gallardo-Rincón
Fundación Carlos Slim
Support for paper writing:

Writing support was funded by the Fundación Carlos Slim.

Disclosure statement:

H.G.R., L.A.M.J., J.L.G., A.M., J.O.M., and D.A.A.H. are employed by the Fundación Carlos Slim. All other authors have nothing to disclose.

Word count, abstract: 259

Word count, text: 5358

Number of tables/figures: 2/6

Supplementary materials: 14 tables
Abstract

Purpose
To investigate the effect of sex and age on the timing of a type 2 diabetes (T2D) diagnosis and the influence T2D-related genes, parental history of T2D, and obesity on T2D development.

Methods
In this case-control study, 1012 T2D cases and 1008 healthy subjects were selected from the Diabetes in Mexico Study database. Participants were stratified by sex and age at T2D diagnosis (early, ≤45 years; late, ≥46 years). Seventy T2D-associated SNPs were explored and the percentage contribution (R²) of T2D-related genes, parental history of T2D, and obesity (body mass index [BMI] and waist-hip ratio [WHR]) on T2D development was calculated using univariate and multivariate logistic regression models.

Results
T2D-related genes influenced T2D development most in males who were diagnosed early (R² = 23.5%; females diagnosed early, R² = 13.5%; males and females diagnosed late, R² = 11.9% and R² = 7.3%, respectively). With an early diagnosis, insulin production genes were more influential in males (76.0% of R²) whilst peripheral insulin resistance genes were more influential in females (52.3% of R²). With a late diagnosis, insulin production genes from chromosome region 11p15.5 notably influenced males while peripheral insulin resistance and inflammation genes notably influenced females. Influence of parental history was higher among those diagnosed early (males, 19.9%; females, 17.5%) versus late (males, 6.4%; females, 5.3%). Unilateral maternal T2D history was more influential than paternal T2D history. BMI influenced T2D development for all, while WHR exclusively influenced males.

Conclusions
The influence of T2D-related genes, maternal T2D history, and fat distribution on T2D development was greater in males than females.

Keywords: age, genetic loci, obesity, sex, type 2 diabetes, parental history
Introduction

The etiological parameters of type 2 diabetes (T2D) have not been fully elucidated; however, T2D is clearly associated with obesity, parental history, genes, ethnicity, poor eating habits, dyslipidemia, lack of physical activity, and aging, among others (1). Of these, obesity, parental history, and genes appear most important. Obesity, a modifiable factor, is clearly associated with T2D development and is considered the main risk factor; T2D risk increases linearly with increased body mass index (BMI) (2-4). However, the relationship between T2D and obesity may not be as direct as thought. For example, in countries such as China, India, and Japan, where the prevalence of T2D is high, the prevalence of obesity is relatively low. Currently, we know that approximately two-thirds of people with T2D are overweight or obese; however, only 2%–13% of obese people develop T2D. When using univariate logistic regression (ULR) models, we recently reported that the variability (R^2) of T2D attributed to obesity would be 12.6%–14.9% (5). However, these values vary substantially in multivariate logistic regression (MLR) models that include parental history of T2D and genes as variables (5).

The importance of parental history on T2D development has been well established in both family (6-8) and association studies (9). The heritability of T2D is variable, reported to be around 25% in family studies (9-11) and 20% in association studies (6, 7, 12). In a case–control study conducted in Mexico, the R^2 of T2D attributed to parental history of diabetes was 12.1%. There are currently over 400 genes with an identified association with T2D (13). However, based on their low individual odds ratios (OR), most genes have very little influence on the development of the disease, which can vary with different study populations. Most studies suggest that genetic involvement contributes to no more than 10% of T2D variability (13, 14). Recently, a study of 16 single nucleotide polymorphisms (SNPs) with the highest predictive power for T2D in a Latin American population reported that only nine were associated with T2D and that this contributed to only 6.5% of T2D variability (5). However, as previously reported (15), this study found that parental history and genes have a greater influence on T2D when it occurs before the age of 46 and that the influence of these factors...
decreases substantially when the disease is diagnosed later in life (≥46 years). Furthermore, this study also found an important difference in the association of genes, but not of parental history, with T2D between males and females (15). Here, the $R^2$ value was much higher in males than females (11.2% versus 4.1%), and given that the differences in gene association between males and females have not been previously reported or have only been reported for individual genes (16, 17), the differences are likely specific to the ancestry of populations in Mexico, which include a mixture of different Amerindian and European ancestries. In fact, in recent high-throughput studies, as part of the Slim Initiative in Genomic Medicine for the Americas (SIGMA), polymorphisms strongly associated with T2D have been discovered in Mexican and Latin populations that are either non-existent or very rare in European populations. These polymorphisms are located in \textit{SLC16A11} (18), \textit{INS-IGF2} (19), and \textit{HNF1A} (20). Here, the difference between the $R^2$ values of males and females may be partially explained by the difference in the effect size of these genes between sexes. We note that polymorphisms in well-established genes associated with T2D in multiple populations, such as \textit{KCNQ1} and \textit{TCF7L2}, are also important contributors to this difference (5). Because of the small number of genes analyzed in a previous study, we were unable to investigate which biological processes associated with T2D (e.g., insulin production, peripheral insulin resistance, inflammation) were most relevant for the genetic differences between men and women and between early- and late-onset T2D. Given the difference in the degree of association of genes between the two sexes and between the late and early presentation of T2D, it is likely that the biological processes associated with T2D in each group are markedly different.

In this work, a case–control study (n = 2020) was performed, including 69 SNPs associated with T2D located in genes involved in different biological processes. The aim was to solve the following research questions: 1) to confirm whether the influence of genes on T2D is different between men and women, 2) to investigate which biological processes have different weights of contribution in T2D between sexes and age at T2D presentation, and 3) to understand how the contribution of each biological process changes in each group when...
obesity and parental history of diabetes are included in the MLR models. To identify the factors associated with T2D in each group, univariate logistic regression (ULR) models were used. Then, MLR models were used to calculate the percentage contribution (R²) of each factor to the variability of T2D. In the latter, the genes (grouped by biological processes) and the remaining factors were entered in successive blocks to calculate the contribution of each factor (R²) in the global model. The variation in the order of entry of each factor allowed us to identify how obesity and parental history affect the importance of gene biological processes linked to insulin production and resistance in the variability of T2D between sexes and age of disease onset.

Material and Methods

Sample selection and study design

Study participants were part of the Diabetes in Mexico Study (DMS) (21). The DMS study design was previously described as part of the SIGMA Type 2 Diabetes Consortium. Briefly, participants were recruited from two tertiary-level hospitals in Mexico City (Centro Médico Nacional siglo XXI del Seguro Social Mexicano [IMSS], Centro Médico Nacional 20 de Noviembre del Seguro Social para los Trabajadores del Estado [ISSSTE]). T2D was diagnosed according to the criteria of the American Diabetes Association (22). A total of 1012 cases (unrelated individuals, >20 years old, with a previous diagnosis of T2D or fasting glucose levels >125 mg/dL) and 1008 controls (healthy subjects, >50 years old, fasting glucose levels <100 mg/dL) were selected from the DMS database, which consisted of 1500 controls and 1500 cases of the Latin American mestizo population. Participants and same-sex, same-hospital controls were recruited between November 2009 and August 2013. Clinical information including weight, waist and hip circumference, and parental history of T2D was collected during the initial interview. For fasting glucose measurements and DNA extraction, 10 mL of intravenous blood was collected. The association of the 69 selected SNPs, parental history of T2D, BMI, and WHR with T2D was investigated using ULR and MLR models.
This study was approved by the Ethics and Research Committees of the National Institute of Genomic Medicine and the Federal Commission for the Protection against Health Risks (COFEPRIS; approval number CAS/OR/CMN/113300410D0027-0577/2012). The study protocols were carried out in accordance with the ethical principles outlined in the 1964 Declaration of Helsinki. Written informed consent was obtained from all participants prior to their inclusion in the study. This study was not registered in a clinical trial registry.

**SNP selection and genotyping**

Sixty-nine SNPs were selected; nine associated with T2D in a previous study (5) and 60 from different GWAS databases and published articles (Table S1). SNPs with a minor allele frequency (MAF) ≥10% and OR ≥ 1.2 or ≤ 0.83 in the Latin American mestizo population and those localized in genes with a role in cellular processes involved in T2D development were prioritized. All DNA samples were genotyped for the 69 SNPs listed in Table S1. Assays were designed using Applied Biosystems TaqMan SNP assay design technology (Foster City, CA, USA). Genotyping was performed by the allelic discrimination assay-by-design TaqMan® method on OpenArray® plates. The plates were analyzed on the QuantStudio™ 12 K Flex Real-Time PCR System (ThermoFisher Scientific, Waltham, MA, USA). The genotypes were analyzed by Genotyper™ Software v1.3 (ThermoFisher Scientific, Waltham, MA, USA).

A total of 23 SNPs that had a difference of distribution between cases and controls of \( P < 0.2 \) in the ULR models were ultimately included in the MLR models. The list of SNPs and gene symbols are presented in Table 1, and the names of genes and their function in Table S2. The SNPs were classified into four groups according to the function and location of the gene in which they occur: Group 1, insulin production located in chromosome region 11p15.5 (n = 5); Group 2, insulin production located in other regions (n = 7); Group 3, peripheral insulin resistance (n = 6); and Group 4, inflammation and other functions (n = 5) (Table 1, Table S2).
**Statistical analyses**

When determining the sample size, we considered MLR models to have good performance when there was a baseline of 50 cases (as in the ULR models) and 10–15 additional cases for each variable introduced in the model (23). Given that we planned to introduce 15–20 variables in the MLR models, we calculated a minimum number of 200–350 cases in the comparison groups.

Participant age, age at T2D diagnosis, BMI, and WHR were expressed as mean ± standard deviation (SD). Distribution of the frequency of genotypes according to the Hardy–Weinberg law based on the allelic frequency and the formula: $(a + b)^2 = a^2 + 2ab + b^2$, where $a$ and $b$ are the allelic frequencies in the control group, was investigated. Differences in the distribution of genotypes between the observed and expected results were calculated using the chi-square test. The BMI was adjusted (BMIadj) for participants who had been diagnosed with T2D for ≥3 years using data from patients who had been diagnosed with T2D for ≤2 years (5). In the ULR and MLR models, case (diagnosis of T2D) or control was considered a dependent variable; values of the alleles and genotypes of the SNPs, parental history, BMI, and WHR were considered explanatory variables. Interactions between each SNP and both sex and age of T2D presentation were assessed in the ULR models. All variables (BMIadj, WHR, parental history of T2D, and SNPs) were analyzed with stratification by sex (male and female) and age at T2D diagnosis (≤45 years and ≥46 years). The risk conferred by each factor (explanatory variables) was calculated by comparing cases and controls using ULR models. Association was expressed as an OR with 95% confidence intervals (CIs); the contribution to the variability of T2D was expressed as Nagelkerke’s $R^2$, which represents the percentage of T2D variability that can be explained by a named factor (24). Confounders were identified using a theoretical strategy based on a backstep, stepwise MLR model and the change-in-estimate criterion. Variables with $P < 0.20$ in the ULR analysis were considered for entry in the MLR model. Confounders were defined as those variables for which the percentage difference between the values of the regression $\beta$ between the
adjusted and non-adjusted variables in the stepwise MLR model was larger than 10% \((P > 0.1)\). Therefore, the total variability and the contribution of each factor on T2D was calculated using this MLR model. These factors were included successively in the model in different blocks, and the contribution of each factor to the model was assessed by the increase of \(R^2\) and the decrease in the \(-2\) log likelihood ratio value from one block to the next; the Omnibus test was used to determine whether the differences between the successive blocks were statistically significant. A post-hoc power analysis was performed for each logistic regression model using the software G * Power 3.1.9.2, considering the sample size, the OR, the probability of the event in the control group, and an \(\alpha = 0.05\) (25). In addition, for MLR models, the value of the total \(R^2\) obtained at the end of the model was introduced for power calculation.

All statistical tests were two-sided, and differences were considered significant when \(P < 0.05\). Statistical analyses were conducted using SPSS version 20 software (IBM Corp., Armonk, NY, USA).
Results

Participant characteristics
The 2020 participants (cases, 1012; controls, 1008) were included in the SNP and BMI analyses; given missing data for some participants, 1690 were included in the parental history analysis of T2D and 1335 in the WHR analysis. WHR was not included in some MLR models; a total of 1690 participants were included in models where parental history was introduced. The demographic characteristics of patients with T2D and non-diabetic controls are presented in Table S3.

Among the participants, 53.9% were female and 46.1% were male. The mean ± SD age of non-diabetic controls (59.1 ± 11.3 years) was higher than that of cases (55.5 ± 11.7 years) at the time of enrollment. The mean ± SD age at T2D diagnosis was 45.8 ± 10.7 years, and the number of years with T2D varied widely (range, 0–46 years; mean ± SD, 9.5 ± 8.7 years). At the time of study recruitment, most patients (89.5%) were taking medication for their T2D, and few were smokers (cases, 20.2%; controls, 19.8%).

Identification of genes associated with T2D using ULR models
The allelic frequencies of only 23 (Table 1) of the 69 SNPs explored (Table S1) were significantly different ($P < 0.1$) between cases and controls in all participants (Table S4), stratified by age at T2D diagnosis ($\leq$45 years and $\geq$46 years) (Table S4), sex (male and female) (Table S5), or both (Table S6).

In the allelic ULR analysis, 12 of 23 SNPs showed a significant association with T2D for the entire study population (Table 1). The $R^2$ sum of these 12 SNPs only explains 5.3% of the variability of T2D etiology.

When analyzed according to age of T2D diagnosis ($\leq$45 years versus $\geq$46 years), 8 of the 12 SNPs associated with T2D in the whole sample were associated with T2D in both age groups (Table S7). SNPs in SLC30A8 (rs3802177), NOS3 (rs2070744), and KHDRBS3 (rs6577691) were associated only with early T2D diagnosis and SNPs in CACNA1H (rs4984636) were only associated with late diagnosis. Two additional SNPs, not associated
with T2D in the analysis of the entire study population, were also identified: **WFS1** (rs4458523) (associated with early T2D diagnosis) and **HMG20A** (rs1005752) (associated with late T2D diagnosis) (**Table S7**). No major differences were observed in the association of gene groups between early or late T2D diagnosis. However, all SNPs except **INS-IGF2** (rs149483638) had a greater association with early than late T2D diagnosis. Compared with the overall analysis, the $R^2$ sum increased in SNPs associated with early T2D diagnosis ($R^2$ sum = 7.37%) and decreased in those associated with late T2D diagnosis ($R^2$ sum = 3.32%).

There was a significant interaction ($P < 0.1$) between sex and some SNPs, both in the entire study population ($n = 8$) and when stratified by early ($n = 9$) and late ($n = 6$) T2D diagnosis (**Table 1, Table S7**). **Table S5** shows allelic frequency and **Table S7** shows the ULR results stratified by sex. Twelve and eight of 23 SNPs were associated with T2D exclusively in males and females, respectively; five were associated with both sexes. A large difference was observed in the association of different groups of genes between males and females. For example, all five SNPs on the chromosome region 11p15.5 were associated with T2D in males but only three SNPs were in females. The two SNPs that were not associated with T2D in females were located at the upstream end of the region where **INS** and **IGF2** are located. Those shared with males, however, were located in the downstream region (**KCNQ1** and **SLC22A18**) (**Figure 1**). The two SNPs located in **KCNQ1** had a stronger association with T2D in males than females (OR >1.6 versus OR = 1.2; $P < 0.0001$ versus $P < 0.05$), while the SNP in **SLC22A18** had a slightly higher association in females versus males (OR >1.43 versus OR = 1.32; $P < 0.0002$ versus $P < 0.007$). None of the seven SNPs in genes involved with insulin production located in other genomic regions (Group 2) were associated with T2D in females; however, three SNPs (**IGF2BP2** [rs4402960], **TCF7L2** [rs7903146], **SLC30A8** [rs3802177]) were associated in males. Peripheral resistance appears more important than insulin production for T2D development in females. The SNP in **SLC16A11** had a greater effect in females than males (OR >1.37 versus OR = 1.24; $P < 0.0004$ versus $P < 0.021$) and the SNP in **PPP1R3A** was only associated with T2D in females.
Overall, ORs were higher and \( P \) values lower in genes associated with T2D in males versus females. In females, the highest ORs were reported for SNPs in \textit{SLC16A11} (OR, 1.37) and \textit{SLC22A18} (OR, 1.43), while there were five SNPs with an OR > 1.5 (located in \textit{IGF2BP2}, \textit{INS}, \textit{KCNQ1}, and \textit{TCF7L2}) for males. The \( R^2 \) sum was 3.5% for females and 10.7% for males. When the analysis was stratified by sex and age of T2D detection some differences in the associations of groups of genes were seen between the two groups of males and females (Table 2 and see below in the MLR analysis).

The observed genotypic frequency was distributed according to the Hardy–Weinberg law in all 23 SNPs studied (Table S8). Genotypic frequency is shown in Table S9 (the total sample stratified by age at T2D diagnosis), Table S10 (stratified by sex), and Table S11 (stratified by age at T2D diagnosis and sex). Overall, SNPs associated with T2D in the allelic ULR analysis also showed a significant association in the ULR genotype analysis (Table S12). However, in all ULR genotypic models, \( R^2 \) sum values were much higher than those observed in the univariate allelic models, clearly indicating that genotypes rather than alleles should be used in MLR models.

**Degree of participation of each group of genes in the variability of T2D using MLR models**

The MLR analyses were performed with the SNP groups (Group 1, Group 2, Group 3, and Group 4) introduced in successive blocks. Most SNPs identified by ULR, except the SNP rs450208 (\textit{SLC22A18}) in males and SNP rs2237897 (\textit{KCNQ1}) in females, which could be in partial linkage disequilibrium with some of the other three SNPs located in chromosome region 11p15.5 (Table S13), remained in the MLR models (Table S14).

The influence of genes on T2D was greatest in males diagnosed early (23.5%), followed by females diagnosed early (13.5%), males diagnosed late (11.9%), and females diagnosed late (7.3%). Genes on the chromosome region 11p15.5 (40.3% of the \( R^2 \); Figure 2A) and others involved in insulin production (35.7% of the \( R^2 \); Figure 2A) had the greatest effect in males diagnosed early, which contrasts with the major influence of genes involved
in peripheral insulin resistance in females diagnosed early (52.3% of the $R^2$; Figure 2A).

Also notable, males diagnosed late had a predominance of associated SNPs in genes of the chromosome 11p15.5 region, which contrasts with the lower $R^2$ values of the same SNPs in females diagnosed late. Although an increase in the influence of genes involved with inflammation was observed in males and females diagnosed late, it accounted for nearly one quarter of the genes’ influence in females.

Influence of parental history of T2D

Parental history had the strongest influence in those cases diagnosed with T2D at ≤45 years of age (males, 19.9%; females, 17.5%) and the least influence in those cases diagnosed with T2D later in life (≤6.4%) (Figure 2B). History of T2D in both parents had the most influence in both males and females diagnosed early. In contrast, having only a father with T2D had little or no influence on T2D development whereas having a mother with T2D influenced T2D diagnosis, especially among males diagnosed early (25%) and females diagnosed late (50%). The risk of developing T2D early was approximately 10 times higher when both parents and three times higher when one parent had a history of T2D versus neither. In contrast, a history of T2D in both parents conferred a lower risk of T2D in those diagnosed late compared with an early diagnosis (males, 3.6 times higher; females, 2.7 times higher versus neither parent with T2D) (data not shown).

In the four groups, the total value of $R^2$ was higher when parental history was the first block and genes was the second block, particularly for males diagnosed early ($R^2$, 0.403) (Figure 2C). In this group, gene effects decreased from 0.251 to 0.201 (Figure 2C) as compared with genes in the first block. In contrast, the effect of parental history decreased when analyzing in the opposite order ($R^2$ decreased from 0.202 to 0.119), suggesting a common effect on 5%–8% of T2D variance linked to genes and parental history of T2D, and that much of the effect of parental history is not directly due to the assessed genetic polymorphisms. Interestingly, when the effect of genes is reduced, there is not a uniform reduction for the four groups of genes. Rather, the influence ($R^3$) of genes related to insulin
production is decreased (chromosome region 11p15.5 [Group 1], 32%; other genomic regions [Group 2], 34%) and in genes related to peripheral insulin resistance (Group 3), the influence is increased by 25%. In the reverse model (first block, genes; second block, parental history) the decreased influence of parental history was not uniform for a history in the father (decreased from a small effect to no effect), mother (60% decrease), or both (32% decrease) in males diagnosed early (Figure 2C). These data may suggest a relationship between SNPs on genes in the chromosome region 11p15.5 and a maternal history of T2D, especially given that genes in this region have genetic imprinting on the maternal side (26).

When stratified by the type of parental inheritance, the distribution of risk alleles compared with alternative alleles for several 11p15.5 genes and TCF7L2 only differs between cases and controls with unilateral maternal inheritance in males diagnosed early (Figure 3).

Influence of obesity and fat distribution

ULR revealed that BMI was important in all four groups of T2D (males diagnosed early, $R^2 = 0.198$; males diagnosed late, $R^2 = 0.147$; females diagnosed early, $R^2 = 0.132$; females diagnosed late, $R^2 = 0.106$) (Figure 4A). WHR was most important for both male groups (Figure 4A). BMI results were similar in MLR models when BMI was the first block (Figure 5A and Figure 5B). However, if BMI was the third block (following parental history and genetic influence), the effect decreases substantially (males diagnosed early, 54.6%; males diagnosed late, 34.3%; women diagnosed early, 40.7%; females diagnosed late, 19.9%) (Figure 4B).

Using MLR, the effects of genes and parental history were analyzed separately to determine how much each individually absorbs the effect of BMI (first block, genes or parental history, second block, BMI; first block, BMI, second block, genes or parental history). In males diagnosed early, parental history or genes decreased the effect of BMI by 43.6% or 23.3%, respectively. With parental history as the first block, the effect was mainly absorbed by having a mother with T2D (87.2% increase) or both parents with T2D (72% increase); the effect of a father with T2D was minor (Figure 5B). With genes as the first
block, the effect of BMI was mainly absorbed by genes located in the chromosome region 11p15.5 (58.5% increase) (Figure 5A). Similar increases with parental history or genes as the first block were observed in males diagnosed late and females diagnosed early (Figure 5A and Figure 5B). A similar analysis was performed using WHR rather than BMI. The findings in males diagnosed early were similar to those found for BMI (Figure 5C and Figure 5D) and together these results suggest that some components of parental history, and to a lesser extent genes on the chromosome 11p15.5 region, are linked to obesity.

Using the Pearson correlation, we found a positive correlation between the number of risk alleles in KCNQ1 (chromosome 11p15.5 region) and both BMI and WHR in males diagnosed early ($r = 0.125$, $P = 0.001$ and $r = 0.130$, $P < 0.009$, respectively) and late ($r = 0.103$, $P < 0.007$ and $r = 0.140$, $P < 0.003$, respectively) (Figure 6). A positive correlation was also observed between SNP in INS and BMI in males diagnosed early ($r = 0.09$, $P = 0.019$; data not shown).

When BMI was introduced as the last block, the influence of BMI ($R^2$) was similar among the four groups of T2D ($R^2$ range: 0.078–0.094). However, the strength of the BMI effect related to parental history and genes appeared stronger for a late diagnosis compared with an early diagnosis (Figure 4B), indicating that obesity is more important than genetic influence and parental history for males and females diagnosed with late-onset T2D.
Discussion

This study identified notable differences in the association of T2D-related genes, parental history of T2D, BMI, and WHR with T2D development between males and females of Latin American mestizo origin. When participants were stratified by age of T2D presentation, the differences observed between males and females remained regardless of whether they were diagnosed with T2D early ($\leq 45$ years) or late ($\geq 46$ years). However, in males and females diagnosed late, the influence of genes and family history decreased drastically, while the influence of obesity increased. The reduced effect of T2D-related genes and parental history in late-onset T2D has previously been reported in other studies on White populations (27, 28). The three factors (genetic loci, parental history, and BMI) appear related to each other, given that they each compete for a percentage of $R^2$ (percent of T2D variability explained by each factor). In fact, the relationship of these three factors seems to be much more important in this studied population than in previously analyzed European populations (29).

The differences between males and females were most evident in those diagnosed with T2D early: 1) the genetic contribution involved in insulin production was predominant in males while genes involved in peripheral resistance were more important in females, 2) the influence of obesity, especially WHR, was much greater in males than females, and 3) the influence of a mother (unilateral) with a history of T2D greatly influenced males. Given genes that are involved in insulin production are located in autosomes, differences in T2D association should not be related to a differential distribution of risk alleles between males and females. Rather, this is most likely associated with external factors that differentially influence these genes in males and females. Our findings suggest that obesity, hormones, and maternal inheritance may be involved in the selection of males, but not females, who have risk alleles involved in insulin production for the early development of T2D. The association of BMI and especially WHR was much higher in males than females in the group that was diagnosed with T2D early (WHR $16\times$ higher). Also, in males there was a linear correlation between BMI and WHR and the number of risk alleles in several genes located in the chromosome region 11p15.5 ($\text{INS}$ and $\text{KCNQ1}$) that are associated with insulin.
production. There is clear evidence that sex hormones play an important role in fat
distribution, measured as WHR (30). WHR is a measurement of visceral fat (waist
circumference) and gluteal fat (hip circumference). Gluteal-femoral fat predominates in
females prior to menopause whereas abdominal fat predominates in postmenopausal
females and males of all ages. Abdominal fat has been associated with an increased risk of
developing T2D and cardiovascular disease (31). This may be contributing to the lower risks
of developing T2D in premenopausal females compared with males of the same age (32).

The overall prevalence of T2D is slightly higher in males than females (33-35); in the
2021 International Diabetes Federation report, the prevalence of T2D in males and females
20–79 years of age was 10.8% and 10.2%, respectively (17.7 million more males than
females with T2D). However, this difference is not uniform across age ranges; the difference
is greatest among those ≤55 years of age (approximately 16.5% versus 13.5%) compared
with those who are older, for whom the difference disappears with increasing age (17, 34-
36). In Mexico, this difference appears even greater. For example, in 2019 the incidence of
T2D was 17.2% higher in males than females aged 15–49 years (524 versus 447
cases/100,000 persons) (37). This lower prevalence in younger females coincides with the
premenopausal period, suggesting a protective effect for estrogens. This protective role has
been demonstrated in studies of estradiol (38). Hormone replacement therapy is reported to
reduce the incidence of T2D by 35% versus placebo in postmenopausal females (39) and
early menopause is associated with an increased risk of T2D (40). Fasting glucose levels in
premenopausal females are lower than those in age- and weight-matched males (41), while
rat models have shown that estrogens protect against T2D by increasing insulin sensitivity
(42, 43) and insulin secretion (43), and decreasing oxidative stress (42). This may suggest
that a hormonal effect may protect or delay the impact of genes on the development of T2D
in females. Furthermore, overweight or obese males may have low concentrations of serum
testosterone, which is associated with an increased risk of T2D (44). A study of males aged
50-74 years, with a waist circumference ≥95 cm, serum testosterone concentration ≤14.0
nmol/L, and impaired glucose tolerance or newly diagnosed T2D investigated the effects of
participation in a lifestyle program with or without testosterone treatment on T2D (45). The study found that those who received testosterone treatment for 2 years had a greater reduction in 2-hour glucose versus baseline than those who received placebo (mean difference −0.75 mmol/L; p<0.0001). Pancreatic islets have also been shown to be more susceptible to oxidative stress in males than females (46).

It has previously been shown that the influence of maternal history of diabetes is more important than paternal influence, at least for the early development of T2D (47). In the present study, we observed that the parental history of T2D through the mother (unilateral) may play an important role in differences in the association of risk alleles between the sexes in early-diagnosed T2D for two reasons: 1) MLR analysis revealed a substantial reduction in the influence of a mother with T2D when the groups of genes involved in insulin secretion were included in the first block and parental inheritance in the second, and 2) the distribution of risk alleles in \textit{INS} and \textit{KCNQ1} (chromosome 11p15.5 region) and \textit{TCF7L2} was significantly different between cases and controls only when the mother had a history of T2D (not with the father or both parents). At least in the case of the chromosome 11p15.5 region, there is evidence of genetic imprinting on the maternal side; a previous study reported that risk alleles in \textit{KCNQ1} conferred a risk for T2D only when are inherited by the mother (26) and that risk alleles influenced methylation levels of a regulatory sequence in fetal human pancreas, suggesting that diabetes risk effects may be mediated in early development (48). Interestingly, the effect of maternal inheritance on genes from that region was only observed in males with an early T2D diagnosis. Such an association has not been previously reported.

There were also important differences in the T2D models of males and females with a late T2D diagnosis. The effect size of T2D-related genes was larger in males than females, primarily at the expense of the genes involved with insulin production from the chromosome 11p15.5 region. It is also noteworthy that in both groups with a late diagnosis, the proportion of the participation of genes involved with inflammation and other processes, relative to the total variance due to genes, was higher than in those with an early diagnosis. This could be related to increased inflammatory processes with aging (49).
This study had some limitations. As noted, the number of participants with missing data for WHR was notable. However, the important difference in the median WHR and the degree of association with T2D between men and women, as well as the significant correlation between WHR and the number of risk alleles of the KCNQ1 gene, are very notable. These data could be considered valid because they were explored in ULR analyses. However, it is clear that the interpretation of the effect of WHR in MLR models should be confirmed in studies with a larger sample of cases and controls. Furthermore, this study is also limited by the inclusion of only Mexican participants and therefore the study findings are not necessarily generalizable to other populations.

The previous results allow us to conclude that the influence of T2D-related genes, maternal T2D history, and fat distribution on T2D development was greater in males than females.
Funding

This work was funded by the Fundación Carlos Slim, the Laboratory Huella Génica, and the Faculty of Medicine of the Universidad Nacional Autónoma de México (UNAM).

Acknowledgements

The authors thank Sarah Bubeck, PhD, of Edanz (www.edanz.com) for providing medical writing support, which was funded by Fundación Carlos Slim, in accordance with Good Publication Practice (GPP3) guidelines (http://www.ismpp.org/gpp3).

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Author contributions

References


23

W, Ma RCW, Pollin TI, Sandhu M, Tandon N, Froguel P, Barroso I, Teo YY, Zeggini
E, Loos RJF, Small KS, Ried JS, DeFronzo RA, Grallert H, Glaser B, Metspalu A,
Wareham NJ, Walker M, Banks E, Gieger C, Ingelsson E, Im HK, Illig T, Franks PW,
Buck G, Trakalo J, Buck D, Prokopenko I, Mägi R, Lind L, Farjoun Y, Owen KR,
Gloyn AL, Strauch K, Tuomi T, Kooner JS, Lee JY, Park T, Donnelly P, Morris AD,
Hattersley AT, Bowden DW, Collins FS, Atzmon G, Chambers JC, Spector TD,
Laakso M, Strom TM, Bell GI, Blangero J, Duggirala R, Tai ES, McVean G, Hanis
Gabriel S, Burtt NP, Mohlke KL, Meitinger T, Groop L, Abecasis G, Florez JC, Scott
LJ, Morris AP, Kang HM, Boehnke M, Altshuler D, McCarthy MI. The genetic

10. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI,
Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L,
Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG,
Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding

based rare variant association analysis: A fast and efficient method of multivariate

identification of genetic loci jointly influencing diabetes-related traits and sleep traits

13. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, Payne AJ,
Steinthorsdottir V, Scott RA, Grarup N, Cook JP, Schmidt EM, Wuttke M, Sarnowski
C, Mägi R, Nano J, Gieger C, Trompet S, Lecoeur C, Preuss MH, Prins BP, Guo X,
Bielak LF, Below JE, Bowden DW, Chambers JC, Kim YJ, Ng MCY, Petty LE, Sim X,
Zhang W, Bennett AJ, Bork-Jensen J, Brummett CM, Canouil M, Ec Kardt KU,
Fischer K, Kardia SLR, Kronenberg F, Läll K, Liu CT, Locke AE, Luan J, Ntalla I,
D, Martin NG, Marroni F, Mangino M, Magnusson PK, Lucas G, Luben R, Loos RJ,
Lokki ML, Lettre G, Langenberg C, Launer LJ, Lakatta EG, Laaksonen R, Kyyvik KO,
Kronenberg F, König IR, Khaw KT, Kaprio J, Kaplan LM, Johansson A, Jarvelin MR,
Janssens AC, Ingelsson E, Igl W, Kees Hovingh G, Hottenga JJ, Hofman A, Hicks
AA, Hengstenberg C, Heid IM, Hayward C, Havulinna AS, Hastie ND, Harris TB,
Haritunians T, Hall AS, Gyllensten U, Guiducci C, Groop LC, Gonzalez E, Gieger C,
Freimer NB, Ferrucci L, Erdmann J, Elliott P, Ejebe KG, Döring A, Dominiczak AF,
Demissie S, Deloukas P, de Geus EJ, de Faire U, Crawford G, Collins FS, Chen YD,
Caulfield MJ, Campbell H, Burtt NP, Bonnycastle LL, Boomsma DI, Boekholdt SM,
Bergman RN, Barroso I, Bandinelli S, Ballantyne CM, Assimes TL, Quertermous T,
Altshuler D, Seielstad M, Wong TY, Tai ES, Feranil AB, Kuzawa CW, Adair LS,
Taylor HA Jr, Borecki IB, Gabriel SB, Wilson JG, Holm H, Thorsteinsdottir U,
Gudnason V, Krauss RM, Mohlke KL, Ordovas JM, Munroe PB, Kouner JS, Tall AR,
Hegele RA, Kastelein JJ, Schadt EE, Rotter JL, Boerwinkle E, Strachan DP, Mooser
V, Stefansson K, Reilly MP, Samani NJ, Schunkert H, Cupples LA, Sandhu MS,
Ridker PM, Rader DJ, van Duijn CM, Peltonen L, Abecasis GR, Boehnke M,
Kathiresan S. Biological, clinical and population relevance of 95 loci for blood lipids.
Nature. 2010;466(7307):707-713.

age of onset for type 2 diabetes mellitus in addition to the conventional risk factors. J

16. Humphries SE, Gable D, Cooper JA, Ireland H, Stephens JW, Hurel SJ, Li KW,
Palmen J, Miller MA, Cappuccio FP, Elkeles R, Godsland I, Miller GJ, Talmud PJ.
Common variants in the TCF7L2 gene and predisposition to type 2 diabetes in UK
European Whites, Indian Asians and Afro-Caribbean men and women. J Mol Med
(Berl). 2006;84(12)1005-1014.


**Figure legends**

**Figure 1.** Graphic representation of the chromosome 11p15.5 region.

**Figure 2.** Influence ($R^2$) of genes and parental history of T2D on T2D development.

MLR analysis stratified by age at T2D diagnosis and sex was conducted to demonstrate (A) the effect of groups of genes, (B) parental history, and (C) both factors on T2D development (early diagnosis was ≤45 years whereas late diagnosis was ≥46 years). Gene groups and parental history as factors were introduced into the models as successive blocks. The gene groups were as follows: Group 1, five SNPs located in genes on chromosome region 11p15.5 that are in involved in insulin production (INS, IGF2, and KCNQ1); Group 2, seven SNPs located in genes outside of chromosome 11 and involved in insulin production; Group 3, six SNPs located in genes involved in peripheral insulin resistance; Group 4, five SNPs located in genes involved in inflammation and other functions. The influence of parental history of T2D was investigated according to having a mother with T2D, a father with T2D, or both parents with T2D. In panel C, the effects of genes and parental history were investigated with genes as the first block and parental history as the second block and vice versa.

Abbreviations: MLR, multivariate logistic regression; SNP, single nucleotide polymorphism; T2D, type 2 diabetes.

**Figure 3.** Analysis of allelic frequency stratified by type of parental history of T2D in males diagnosed early. A comparison of frequencies of risk and alternative alleles for SNPs in genes located in the chromosome region 11p15.5 (INS and KCNQ1) and TCF7L2 between cases and controls stratified by the type of parental history of T2D (mother only, father only, both parents). $P$ values were calculated using a chi-square test.

Abbreviations: SNP, single nuclear polymorphism; T2D, type 2 diabetes.
**Figure 4.** Influence (R²) of genes, parental history of T2D, BMI, and WHR on T2D development. (A) MLR analysis was conducted to determine the involvement of BMI and WHR in the four models of T2D (stratified by age of T2D diagnosis and sex). (B) MLR analysis of the participation of parental history, genes, and adjusted BMI in the four models. Abbreviations: BMI, body mass index; MLR, multivariate logistic regression; T2D, type 2 diabetes; WHR, waist-hip ratio.

**Figure 5.** Influence (R²) of obesity and fat distribution and genes and parental history on T2D development overlap. Overlapping analysis was performed with stratification by age of T2D diagnosis and sex. Overlapping analysis of (A) BMI with genes, (B) BMI with parental history, (C) WHR with genes, and (D) WHR with parental history are shown. The overlapping effect of R² between genes and BMI and between parental history and BMI was observed by introducing BMI as the first block and either genes or parental history as the second block and then conducting the analysis with the blocks reversed. Abbreviations: BMI, body mass index; PH, parental history; T2D, type 2 diabetes; WHR, weight distribution.

**Figure 6.** Correlation analysis between the number of risk alleles in KCNQ1 in relation to obesity and fat distribution. (A) Correlation between the number of risk alleles in KCNQ1 and BMI in males and females with a T2D diagnosis ≤45 years and ≥46 years of age. (B) Correlation between the number of risk alleles in KCNQ1 and WHR in males and females with a T2D diagnosis ≤45 years and ≥46 years of age. Correlations and P values were calculated using the Pearson correlation test. Abbreviations: BMI, body mass index; T2D, type 2 diabetes; WHR, weight distribution.
Table 1. Association of 23 SNPs with type 2 diabetes among the whole sample (N = 4040 chromosomes)

<table>
<thead>
<tr>
<th>Locus (SNP)</th>
<th>Chr</th>
<th>Position</th>
<th>RA</th>
<th>AA</th>
<th>MAF</th>
<th>ULR</th>
<th>OR (95% CI)</th>
<th>p-Wald</th>
<th>R²</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>INS-IGF2</em> (rs149483638)</td>
<td>11</td>
<td>2140300</td>
<td>C</td>
<td>T</td>
<td>0.281</td>
<td>1.25 (1.1–1.4)</td>
<td>0.0022</td>
<td>0.003</td>
<td>&gt;0.05</td>
<td></td>
</tr>
<tr>
<td><em>INS</em> (rs689)</td>
<td>11</td>
<td>2160994</td>
<td>A</td>
<td>T</td>
<td>0.188</td>
<td>1.32 (1.1–1.5)</td>
<td>0.0005</td>
<td>0.004</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td><em>KCNQ1</em> (rs2237897)</td>
<td>11</td>
<td>2837316</td>
<td>C</td>
<td>T</td>
<td>0.392</td>
<td>1.44 (1.3–1.6)</td>
<td>&lt; 0.0001</td>
<td>0.010</td>
<td>&gt;0.05</td>
<td></td>
</tr>
<tr>
<td><em>KCNQ1</em> (rs163168)</td>
<td>11</td>
<td>2803115</td>
<td>C</td>
<td>T</td>
<td>0.436</td>
<td>1.39 (1.2–1.6)</td>
<td>&lt; 0.0001</td>
<td>0.008</td>
<td>&gt;0.05</td>
<td></td>
</tr>
<tr>
<td><em>SLC22A18</em> (rs450208)</td>
<td>11</td>
<td>2910751</td>
<td>T</td>
<td>G</td>
<td>0.311</td>
<td>1.38 (1.2–1.6)</td>
<td>&lt; 0.0001</td>
<td>0.007</td>
<td>&gt;0.05</td>
<td></td>
</tr>
<tr>
<td><strong>Group 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>IGF2BP2</em> (rs4402960)</td>
<td>3</td>
<td>185793899</td>
<td>T</td>
<td>G</td>
<td>0.163</td>
<td>1.28 (1.1–1.5)</td>
<td>0.0031</td>
<td>0.003</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td><em>TCF7L2</em> (rs7903146)</td>
<td>10</td>
<td>112998590</td>
<td>T</td>
<td>C</td>
<td>0.121</td>
<td>1.5 (1.3–1.8)</td>
<td>&lt; 0.0001</td>
<td>0.007</td>
<td>&gt;0.05</td>
<td></td>
</tr>
<tr>
<td><em>CDKN2A</em> (rs10811661)</td>
<td>9</td>
<td>22134095</td>
<td>T</td>
<td>C</td>
<td>0.096</td>
<td>1.12 (0.9–1.4)</td>
<td>0.029</td>
<td>0</td>
<td>&gt;0.05</td>
<td></td>
</tr>
<tr>
<td><em>SLC30A8</em> (rs3802177)</td>
<td>8</td>
<td>117172786</td>
<td>A</td>
<td>G</td>
<td>0.272</td>
<td>1.15 (1–1.3)</td>
<td>0.044</td>
<td>0.001</td>
<td>&gt;0.05</td>
<td></td>
</tr>
<tr>
<td><em>HNF1A</em> (rs483353044)</td>
<td>12</td>
<td>120999288</td>
<td>A</td>
<td>G</td>
<td>0.003</td>
<td>2 (0.8–5)</td>
<td>0.13</td>
<td>0.001</td>
<td>&gt;0.05</td>
<td></td>
</tr>
<tr>
<td><em>WFS1</em> (rs4458523)</td>
<td>4</td>
<td>6288259</td>
<td>G</td>
<td>T</td>
<td>0.233</td>
<td>1.07 (0.9–1.2)</td>
<td>0.35</td>
<td>0</td>
<td>&gt;0.05</td>
<td></td>
</tr>
<tr>
<td><em>HMG20A</em> (rs1005752)</td>
<td>15</td>
<td>77525786</td>
<td>C</td>
<td>A</td>
<td>0.419</td>
<td>0.94 (0.8–1.1)</td>
<td>0.34</td>
<td>0</td>
<td>&gt;0.05</td>
<td></td>
</tr>
</tbody>
</table>
### Group 3

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>Position</th>
<th>RA</th>
<th>AA</th>
<th>CI</th>
<th>OR</th>
<th>MAF</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLC16A11 (rs75493593)</td>
<td>17</td>
<td>7041768</td>
<td>T^a</td>
<td>G</td>
<td>0.374</td>
<td>1.31</td>
<td>1.2–1.5</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>IRS1 (rs1801278)</td>
<td>2</td>
<td>226795828</td>
<td>T^a</td>
<td>C</td>
<td>0.027</td>
<td>0.98</td>
<td>0.7–1.4</td>
<td>0.91</td>
</tr>
<tr>
<td>FABP2 (rs1799883)</td>
<td>4</td>
<td>119320747</td>
<td>T^a</td>
<td>C</td>
<td>0.241</td>
<td>0.99</td>
<td>0.9–1.1</td>
<td>0.89</td>
</tr>
<tr>
<td>CAPN10 (rs7607759)</td>
<td>2</td>
<td>240596709</td>
<td>G^a</td>
<td>A</td>
<td>0.053</td>
<td>1.04</td>
<td>0.8–1.4</td>
<td>0.76</td>
</tr>
<tr>
<td>PPP1R3A (rs1799999)</td>
<td>7</td>
<td>113878379</td>
<td>A^a</td>
<td>C</td>
<td>0.291</td>
<td>1.09</td>
<td>0.9–1.2</td>
<td>0.23</td>
</tr>
<tr>
<td>PTPRD (rs10511567)</td>
<td>9</td>
<td>11606348</td>
<td>C^a</td>
<td>T</td>
<td>0.278</td>
<td>1.02</td>
<td>0.9–1.2</td>
<td>0.76</td>
</tr>
</tbody>
</table>

### Group 4

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>Position</th>
<th>RA</th>
<th>AA</th>
<th>CI</th>
<th>OR</th>
<th>MAF</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL6 (rs1800795)</td>
<td>7</td>
<td>22727026</td>
<td>C^a</td>
<td>G</td>
<td>0.079</td>
<td>1.01</td>
<td>0.8–1.3</td>
<td>0.93</td>
</tr>
<tr>
<td>NOS3 (rs2070744)</td>
<td>7</td>
<td>150992991</td>
<td>C^a</td>
<td>T</td>
<td>0.145</td>
<td>0.84</td>
<td>0.7–1</td>
<td>0.054</td>
</tr>
<tr>
<td>CPED1 (rs10261386)</td>
<td>7</td>
<td>120711855</td>
<td>C^a</td>
<td>T</td>
<td>0.347</td>
<td>1.04</td>
<td>0.9–1.2</td>
<td>0.55</td>
</tr>
<tr>
<td>KHDRBS3 (rs6577691)</td>
<td>8</td>
<td>135961229</td>
<td>G^a</td>
<td>T</td>
<td>0.123</td>
<td>0.84</td>
<td>0.7–1</td>
<td>0.076</td>
</tr>
<tr>
<td>CACNA1H (rs4984636)</td>
<td>16</td>
<td>1202441</td>
<td>C^a</td>
<td>T</td>
<td>0.113</td>
<td>0.83</td>
<td>0.7–1</td>
<td>0.073</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. genes</th>
<th>12</th>
</tr>
</thead>
</table>

| Sum of univariate R² | 0.053 |

Abbreviations: AA, alternative allele; Chr, chromosome; CI, confidence interval; MAF, major allele frequency; OR, odds ratio; R², variability of T2D explained by the SNP; RA, risk allele; SI, sex interaction; SNP, single nucleotide polymorphism; ULR, univariate logistic regression.

^aMinor allele.
Group 1, insulin production located in the chromosome region 11p15.5; Group 2, insulin production located in other chromosomes; Group 3, peripheral insulin resistance; Group 4, inflammation and other functions.
Table 2. Allele association of 23 SNPs with type 2 diabetes stratified by age of type 2 diabetes diagnosis and sex among the total population (N = 4040 chromosomes).

<table>
<thead>
<tr>
<th>Locus (SNP)</th>
<th>T2D diagnosis ≤45 years</th>
<th>T2D diagnosis ≥46 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Females</td>
<td>Males</td>
</tr>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>p-Wald</td>
</tr>
<tr>
<td><strong>Group 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>INS-IGF2</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(rs149483638)</td>
<td>1.06 (0.8–1.3)</td>
<td>0.61</td>
</tr>
<tr>
<td><em>INS</em> (rs689)</td>
<td>1.09 (0.8–1.4)</td>
<td>0.52</td>
</tr>
<tr>
<td><em>KCNQ1</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(rs2237897)</td>
<td>1.31 (1.1–1.6)</td>
<td>0.016</td>
</tr>
<tr>
<td><em>CDKN1C</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(rs163168)</td>
<td>1.36 (1.1–1.7)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.0043</td>
</tr>
<tr>
<td><em>SLC22A18</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(rs450208)</td>
<td>1.46 (1.2–1.9)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.0014</td>
</tr>
<tr>
<td><strong>Group 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>IGF2BP2</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.16 (0.9–1.5)</td>
<td>0.26</td>
</tr>
<tr>
<td>Locus</td>
<td>Group</td>
<td>MAF</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>TCF7L2</td>
<td></td>
<td>0.15</td>
</tr>
<tr>
<td>CDKN2A</td>
<td></td>
<td>0.049</td>
</tr>
<tr>
<td>SLC30A8</td>
<td></td>
<td>0.018</td>
</tr>
<tr>
<td>HNF1A</td>
<td></td>
<td>0.43</td>
</tr>
<tr>
<td>WFS1</td>
<td></td>
<td>0.52</td>
</tr>
<tr>
<td>HMG20A</td>
<td></td>
<td>0.81</td>
</tr>
<tr>
<td>SLC16A11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IRS1</td>
<td></td>
<td>0.14</td>
</tr>
<tr>
<td>FABP2</td>
<td></td>
<td>0.34</td>
</tr>
<tr>
<td>Gene</td>
<td>rs</td>
<td>Group 1</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>CAPN10</td>
<td>7607759</td>
<td>0.76 (0.5–1.2)</td>
</tr>
<tr>
<td>PPP1R3A</td>
<td>17999999</td>
<td>1.36 (1.1–1.7)</td>
</tr>
<tr>
<td>PTPRD</td>
<td>10511567</td>
<td>1.28 (1–1.6)</td>
</tr>
<tr>
<td>Group 4</td>
<td></td>
<td>IL6 (rs1800795)</td>
</tr>
<tr>
<td></td>
<td>NOS3</td>
<td>0.71 (0.5–1)</td>
</tr>
<tr>
<td></td>
<td>CPED1</td>
<td>0.95 (0.8–1.2)</td>
</tr>
<tr>
<td></td>
<td>KHDRBS3</td>
<td>0.9 (0.7–1.2)</td>
</tr>
<tr>
<td></td>
<td>CACNA1H</td>
<td>0.89 (0.6–1.2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>No. SNPs</th>
<th>9</th>
<th>12</th>
<th>4</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of univariate $R^2$</td>
<td>0.056</td>
<td>0.14</td>
<td>0.02</td>
<td>0.079</td>
<td></td>
</tr>
</tbody>
</table>
Abbreviations: CI, confidence interval; OR, odds ratio; R², variability of T2D explained by these genes; SNP, single nucleotide polymorphism; ULR, univariate logistic regression.

*Indicates SNPs that remain in the multivariate model.
GRCh38 chromosome 11p15.5
(*IGF2-INS-TH-KCNQ1-SLC22A18*)

Figure 1

---

CC-BY-NC-ND 4.0 International license
It is made available under a
is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
(which was not certified by peer review)
The copyright holder for this preprint this version posted October 27, 2022; doi: medRxiv preprint
Figure 2A.

Sex and age of T2D presentation

R²

- Females (early T2D diagnosis)
- Males (early T2D diagnosis)
- Females (late T2D diagnosis)
- Males (late T2D diagnosis)

Group 1
Group 2
Group 3
Group 4

CC-BY-NC-ND 4.0 International license
It is made available under a
is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

The copyright holder for this preprint this version posted October 27, 2022. ; https://doi.org/10.1101/2022.10.27.22281587
doi: medRxiv preprint
Figure 2B

Sex and age of T2D presentation

- Females (early T2D diagnosis)
- Males (early T2D diagnosis)
- Females (late T2D diagnosis)
- Males (late T2D diagnosis)

- Father with T2D
- Mother with T2D
- Both parents with T2D
Figure 2C

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Father with T2D</td>
<td>Mother with T2D</td>
<td>Both parents with T2D</td>
<td></td>
</tr>
</tbody>
</table>

Female (early T2D diagnosis) | Male (late T2D diagnosis) | Female (early T2D diagnosis) | Male (late T2D diagnosis)
Figure 3

Parental history of diabetes

**INS (rs689)**

- Frequency (%)
- Controls: 80, Cases: 80
- Controls: 60, Cases: 60
- Controls: 40, Cases: 40
- Controls: 20, Cases: 20
- Controls: 0, Cases: 0

**KCNQ1 (rs2237897)**

- Frequency (%)
- Controls: 80, Cases: 80
- Controls: 60, Cases: 60
- Controls: 40, Cases: 40
- Controls: 20, Cases: 20
- Controls: 0, Cases: 0

**KCNQ1 (rs163168)**

- Frequency (%)
- Controls: 80, Cases: 80
- Controls: 60, Cases: 60
- Controls: 40, Cases: 40
- Controls: 20, Cases: 20
- Controls: 0, Cases: 0

**TCF7L2 (rs7903146)**

- Frequency (%)
- Controls: 80, Cases: 80
- Controls: 60, Cases: 60
- Controls: 40, Cases: 40
- Controls: 20, Cases: 20
- Controls: 0, Cases: 0

Significance levels:
- P = 0.018
- P = 0.025
- P = 0.01
- P = 0.012
Figure 4
Figure 5B

- Father with T2D
- Mother with T2D
- Both parents with T2D
- Adjusted BMI

B

R²

0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35

PH + BMI  BMI + PH  PH + BMI  BMI + PH  PH + BMI  BMI + PH  PH + BMI  BMI + PH

Female (early T2D diagnosis)  Male (early T2D diagnosis)  Female (late T2D diagnosis)  Male (late T2D diagnosis)
Figure 5C
Figure 5D

- **Female (early T2D diagnosis)**
- **Male (early T2D diagnosis)**
- **Female (late T2D diagnosis)**
- **Male (late T2D diagnosis)**

- Father with T2D
- Mother with T2D
- Both parents with T2D
- WHR

<table>
<thead>
<tr>
<th>PH + WHR</th>
<th>WHR + PH</th>
<th>PH + WHR</th>
<th>WHR + PH</th>
<th>PH + WHR</th>
<th>WHR + PH</th>
<th>PH + WHR</th>
<th>WHR + PH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female (early T2D diagnosis)</td>
<td>Male (early T2D diagnosis)</td>
<td>Female (late T2D diagnosis)</td>
<td>Male (late T2D diagnosis)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
BMI

No. of risk alleles

Figure 6