



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Programa de Maestría y Doctorado en Ciencias Bioquímicas

“Caracterización de variantes polimórficas y mutaciones en los genes *HNF4A* y *GCK* en población indígena mexicana mazahua y teenek con diabetes mellitus tipo 2”

TESIS

QUE PARA OPTAR POR EL GRADO DE:

Doctor en Ciencias

PRESENTA:

M en C María de los Ángeles Granados Silvestre

TUTOR PRINCIPAL

Dra. Marta Alicia Menjívar Iraheta (FQ, UNAM)

MIEMBROS DEL COMITÉ TUTOR

Dra. Amelia Farrés González Saravia (FQ, UNAM)

Dr. Julio Granados Arriola (INCMNSZ)

Ciudad de México, mayo de 2018



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

Este trabajo fue apoyado por el proyecto DGAPA IN208507

Beca Conacyt 201106

A la Secretaria de Salud de los Estados de: Campeche, Nayarit, Oaxaca, San Luis Potosí y Tabasco

Al laboratorio de Endocrinología Molecular del Hospital Juárez de México

A todas las personas de los grupos indígenas que voluntariamente participaron en este proyecto

A la Dra. Marta Alicia Menjívar Iraheta por su tutoría, apoyo y confianza en el desarrollo de esta tesis

A los doctores miembros del jurado: Marcia Hiriart Urdanivia, Andrea Díaz Villaseñor, Samuel Canizales Quinteros y Julián Valdés por sus generosas observaciones a esta tesis

DEDICATORIA

“Después de todo, ¿qué es un científico entonces? Es un Hombre curioso que mira a través del ojo de una cerradura, la cerradura de la naturaleza, tratando de saber qué es lo que sucede”

Jacques Yves Cousteau

A Dios, por haberme permitido concluir esta etapa

A mis padres porque fueron ejemplo para conducirme por la vida

A la Universidad Nacional Autónoma de México por brindarme todo el apoyo para mi formación académica

A la Dra. Marta Menjivar por su confianza y entusiasmo, que nos transmite en cada proyecto que emprende. Por su paciencia durante el desarrollo de esta tesis. Por su amistad

A mis hermanos: Carmen, Juanita, Jero, Martín y Jesús, por su apoyo en cada tarea que emprendo

A mis compañeras de laboratorio: Lupita, Katy, Caro y Barbara, por su ayuda, acompañamiento y ánimo durante este camino

A todas aquellas personas que hicieron posible llevar a cabo este trabajo

ÍNDICE

INDICE DE FIGURAS	4
INDICE DE TABLAS	5
ABREVIATURAS	6
RESUMEN	8
I. ANTECEDENTES	10
1. INTRODUCCIÓN	12
1.1 COSTOS DE LA DIABETES EN MÉXICO	13
1.2 PREVALENCIA DE DT2 EN MÉXICO	14
1.3. FISIOPATOLOGÍA DE LA DT2	15
1.3.1 Regulación de la glucosa por el hígado	17
1.3.2 Regulación de la glucosa en músculo esquelético.....	17
1.3.3 Regulación de la glucosa en tejido adiposo	17
1.3.4 Regulación de la glucosa en intestino	18
1.3.5 Regulación de la glucosa en riñón	18
2. PARTICIPACIÓN DE LOS GENES EN LA PATOGÉNESIS DE LA DT2	19
2.1 Heredabilidad en la DT2	19
2.2. HNF4A como gen de riesgo para DT2	23
2.3 Glucocinasa (GCK) como gen de riesgo para DT2	27
3. POBLACIÓN INDÍGENA MEXICANA	30
3.1 Participación de los genes en la susceptibilidad a la DT2 en población Mexicana.....	32
II. JUSTIFICACIÓN	34
III. HIPÓTESIS	34
IV. OBJETIVOS	35
V. MATERIAL Y MÉTODOS	36
VI. RESULTADOS	42
ANCESTRÍA	42
EVALUACIÓN EXÓMICA DE HNF4A Y GCK EN LOS GRUPOS MAZAHUA Y TEENEK.....	43
IDENTIFICACIÓN DE VARIANTES.....	43
FRECUENCIA DE LA MUTACIÓN T130I DEL GEN HNF4A	53
ASOCIACIÓN DE LA VARIANTE T130I (RS1800961) CON RASGOS METABÓLICOS.	58
VII. DISCUSIÓN	63
VIII. CONCLUSIONES	72
IX. PERSPECTIVAS	73
X. BIBLIOGRAFÍA	74
ANEXOS	81

ÍNDICE DE FIGURAS

- Figura 1 Análisis comparativo de las tendencias de los costos sanitarios en la DT2 para países de América Latina
- Figura 2 Antecedentes heredofamiliares de diabetes en pacientes reportados por la DGE. México 2016
- Figura 3 Polimorfismos de un Solo Nucleótido (SNPs), como factores de riesgo de DT2
- Figura 4 Loci asociados a DT2 en población mexicana
- Figura 5 Estructura de *HNF4A*, gen y proteína
- Figura 6 Isoformas de *HNF4A*
- Figura 7 Homología de la proteína de *HNF4A* con y sin la variante rs1800961
- Figura 8 Representación lineal de los dominios de la proteína *HNF4A* y localización tridimensional de la variante T130I.
- Figura 9 Estructura de *GCK* unión a ATP y glucosa
- Figura 10 Estructura del gen *GCK* de la célula β
- Figura 11 Localización geográfica de los 11 grupos indígenas evaluados
- Figura 12 Tamaño de los exones 2, 4 y 5 de *HNF4A*
- Figura 13 Corrimiento electroforético del PCR preparativo del exón 4 del gen *HNF4A*
- Figura 14 Autoradiografía de la SSCP del exón 2 del gen *HNF4A*
- Figura 15 Autoradiografía de la SSCP del exón 4 del gen *HNF4A*
- Figura 16 Electroferograma del exón 2 de *HNF4A* de la muestra MZ 4
- Figura 17 Electroferograma del exón 2 de *HNF4A*, MZ 1
- Figura 18 Electroferograma del exón 2 de *HNF4A*, MZ 5
- Figura 19 Electroferograma de la muestra MZ20 del exón 4 de *HNF4A*.
- Figura 20 Corrimiento electroforético del PCR preparativo del exón 1a de *GCK*
- Figura 21 Autoradiografía de la SSCP del exón 1a de *GCK*
- Figura 22 Electroferograma de la secuencia del exón 1a de *GCK* del individuo MZ50
- Figura 23 Autoradiografía de la SSCP del exón 2 de *GCK*
- Figura 24 Electroferograma de la secuencia del exón 2 de *GCK* del individuo MZ42
- Figura 25 SSCP del exón 6 de *GCK*. Tinción de plata
- Figura 26 Electroferograma de la secuencia del exón 6 de *GCK* del individuo TK17
- Figura 27 Extensión de la homocigocidad haplotípica para la variante T130I (rs1800961)
- Figura 28 Extensión de la homocigocidad haplotípica para la variante T130I (rs1800961) en población europea y asiática.

ÍNDICE DE TABLAS

- Tabla 1 Frecuencia de la variante rs1800961 (T130I) en diferentes poblaciones en personas con y sin diabetes.
- Tabla 2 Oligonucleótidos empleados para amplificación de *HNF4A*.
- Tabla 3 Oligonucleótidos empleados para amplificación de *GCK*
- Tabla 4 Estimación de la ancestría de 11 grupos indígenas mexicanos
- Tabla 5 Variantes de secuencia de *HNF4A* en población mazahua y teenek
- Tabla 6 Variantes de secuencia de *GCK* en población mazahua y teenek
- Tabla 7 Frecuencias alélicas y genotípicas del rs1800961 (T130I) del gen *HNF4A* en población indígena mexicana
- Tabla 8 Comparación de frecuencias alélicas del polimorfismo T301 (rs1800961) entre etnias
- Tabla 9 Comparación de las frecuencias alélicas y genotípicas del rs1800961 entre los indígenas mexicanos : cora, chontal, chatino, maya, mazahua, teenek, mixteco, purépecha, otomí, zapoteco y yaqui, con las poblaciones del Proyecto Internacional HapMap3
- Tabla 10 Características bioquímicas y antropométricas de 7 grupos indígenas mexicanos.
- Tabla 11 Asociación de la variante T130I (rs1800961) del gen *HNF4A*, con factores de riesgo para DT2 en el total de Asociación de la variante T130I (rs1800961), con factores de riesgo para DT2 por etnia la población indígena mexicana evaluada
- Tabla 12 Asociación de la variante T130I (rs1800961) del gen *HNF4A*, con factores de riesgo para DT2 por grupo indígena

ABREVIATURAS

ABCA1	Cassette transportador de unión a ATP tipo A1
ADA	Asociación Americana de Diabetes
AGA	Alteración de glucosa en ayuno
AGL	Ácidos grasos libres
ATG	Alteración de la tolerancia a la glucosa
CAMK1D	Proteína cinasa 1D dependiente de calcio/calmodulina
CAPN10	Calpaína 10
CDC123	Proteína de división del ciclo celular 123
CDKN2A	Inhibidor de cinasas dependiente de ciclina 2A
CDKN2B	Inhibidor de cinasas dependiente de ciclina 2B
DT1	Diabetes tipo 1
DT2	Diabetes tipo 2
ENSANUT	Encuesta Nacional de Salud y Nutrición
GCK	Glucocinasa
GIP	Péptido insulínico
GLP1	Péptido 1 similar a glucagon
GLUT2	Transportador 2 de glucosa
GLUT4	Transportador 4 de glucosa
GWAS	Estudio de Asociación del Genoma Completo
HLA	Antígeno leucocitario humano
Hb1Ac	Hemoglobina Glicosilada
HDL	Lipoproteínas de Alta Densidad
HHEX	Human Hematopoietically expressed homeobox
HNF4a	Factor Nuclear del Hepatocito 4 α
HOMA-IR	Modelo homeostático de la evaluación de la resistencia a la insulina
HTA	Hipertensión arterial
ICC	Índice Cadera-Cintura
IDF	Federación Internacional de Diabetes
IMC	Índice de Masa Corporal

INEGI	Instituto Nacional de Estadística, Geografía e informática
IRS1	Receptor de sustrato de insulina 1
IRS2	Receptor de sustrato de insulina 2
KCNJ11	Canal de potasio, subfamilia J, miembro 11
LDL	Lipoproteínas de Baja Densidad
MODY	Diabetes de la edad madura que se presenta en el joven
MAF	Frecuencia del alelo mayor
OCDE	Organización para la cooperación y el desarrollo económico
OMS	Organización Mundial de la Salud
OR	Razón de momios
PAD	Presión Arterial Diastólica
PAS	Presión Arterial Sistólica
PPARG	Receptor Activado por Proliferador de Peroxisoma gamma
RI	Resistencia a la Insulina
SGLT2	Cotransportador sodio-glucosa 2
SLC16A11	Transportador de solutos familia 16 miembro 11
SLC30A8	Transportador de solutos integrante de la familia 30 miembro 8
SNP	Polimorfismo de un solo nucleótido
SSCP	Polimorfismo conformacional de cadena simple
TCF7L2	Factor 7 de células T similar al 2
TNG	Tolerancia normal a la glucosa
VLDL	Lipoproteínas de Muy Baja Densidad
VR	Valores de referencia

Resumen

La DT2 es el resultado de la pérdida de la homeostasis de la glucosa en la cual intervienen factores ambientales y un componente genético. Para nuestro país, la Federación Internacional de Diabetes (2017) reporta que hay 12 millones de personas con DT2 entre los 20 y 79 años, 4.5 millones de individuos no diagnosticados y no menos importante son los 12.1 millones de sujetos que presentan una alteración en la tolerancia a la glucosa, lo cual nos habla de la importancia de la enfermedad.

Uno de los abordajes para el estudio de la genética de la DT2 es la evaluación de genes que participan en el metabolismo de la glucosa. En este contexto, el Factor Hepático Nuclear (HNF4A) y la enzima Glucocinasa (GCK) son elementos claves que participan en la homeostasis de la glucosa en la célula β , además se ha informado que variantes en los genes de *HNF4A* y *GCK* están asociados en la aparición de DT2 en diversas poblaciones incluyendo la mestiza mexicana.

Aunque epidemiológicamente se ha mostrado la importancia de la DT2 en México, las variantes genéticas implicadas en la aparición de DT2 aún no son del todo conocidas en la población indígena mexicana.

El objetivo de este trabajo fue identificar en grupos indígenas mexicanos variantes de secuencia de los genes *HNF4A* y *GCK* y su probable relación con factores de riesgo para desarrollar DT2.

En este estudio, se incluyeron los grupos indígenas: cora, chatin, chontal, maya, mazahua, mixteco, otomí, purépecha, teenek, yaqui y zapoteco. El total de la población indígena evaluada fue de 879 personas, de las cuales 675 no tenían DT2. En una primera etapa, en los grupos mazahua y teenek se realizó la búsqueda de las variantes de secuencia de *HNF4A* y *GCK* a través de la técnica de polimorfismo conformacional de cadena simple, y posteriormente las variantes encontradas se secuenciaron por el método de Sanger. En una segunda etapa, con los resultados obtenidos en la primera, se buscó la frecuencia de las variantes no sinónimas de los genes en estudio en los 9 grupos restantes. Se llevó a cabo la estimación de ancestría genotipando 10 marcadores genéticos. Se realizó una correlación estadística entre la presencia de las variantes no sinónimas encontradas con factores de riesgo metabólico para el desarrollo de DT2.

En la estimación de ancestría encontramos en promedio 97.6% de componente amerindio. En la evaluación del gen *HNF4A* en la primera etapa, se encontraron tres cambios: rs736824 y rs745975, dentro de la región intrónica y el tercer cambio fue el rs1800961 que corresponde al polimorfismo T130I, estas tres variantes se encontraron en ambas poblaciones. Respecto al gen de *GCK*, encontramos tres variantes: D3D (rs144723656) y Q26H (rs193922329) en población mazahua y Y215Y (rs142553382) en población teenek, estos cambios se encontraron en tres individuos no relacionados. La predicción *in silico* de la variante Q26H, no predijo cambios que impacten en la función de la proteína.

Por la frecuencia en que fue encontrada y además porque la variante T130I (rs1800961) de *HNF4A*, en un estudio previo por nuestro grupo de trabajo se asocio con DT2 de aparición temprana en mestizos mexicanos, buscamos su frecuencia en el resto de las poblaciones en una segunda etapa. La frecuencia en la población total indígena evaluada fue en promedio de 16.6%, siendo los grupos maya y cora los que tuvieron la frecuencia más alta (28.2 % y 25.7% respectivamente) y no fue encontrada en el grupo yaqui. La comparación de las frecuencias obtenidas de las poblaciones del Proyecto Internacional Hap Map, muestra una diferencia estadística significativa con respecto a la frecuencia de la variante T130I (rs1800961) en los pueblos mexicanos.

Se encontró una asociación entre la variante T130I (rs1800961) y concentraciones altas de triglicéridos ($P < 0.05$) en los grupos chatino, chontal y cora, resultado que se conservo tomando en cuenta a toda la población indígena sin diabetes ($P < 0.05$).

Los resultados obtenidos evidencian la presencia de variantes genéticas en población indígena mexicana que se asocian a factores de riesgo para el desarrollo de DT2. De manera particular la variante T130I (rs1800961) de *HNF4A* se asocia con concentraciones altas de triglicéridos. Posteriores estudios serán necesarios para evaluar el impacto de las variantes intrónicas del gen *HNF4A* en la susceptibilidad al desarrollo de DT2. La inclusión de los pueblos originales de México en los estudios genéticos permitirá que estas comunidades participen de los beneficios que puedan aportar este tipo de resultados en el pronóstico y tratamiento de la enfermedad.

Summary

Type 2 Diabetes (T2D) results from the lost glucose homeostasis in which environmental factors interfere and a genetic component. In Mexico, in 2017, the International Diabetes Federation reported 12 million people with T2D within 20 and 79 years old, 4.5 millions of them do not have a diagnostic, and 12.1 million present impaired tolerance glucose. Hence, T2D has acquired a remarkable relevance.

The genes that participate in glucose metabolism are one of the critical approaches for the study of the genetic background of T2D. In this light, the Hepatic Nuclear Factor 4 alpha (HNF4A) and Glucokinase enzyme (GCK) are essential elements that participate in the glucose homeostasis of β cell.

Furthermore, allelic variants in *HNF4A* and *GCK* genes are associated to the appearance of T2D in diverse populations including Mexican Mestizo. In our country, T2D has notable epidemiologic importance. Nonetheless, the genetic variants involved in the appearance of T2D are not entirely known in Mexican indigenous population.

This work aimed to identify sequence variants in *HNF4A* and *GCK* genes and their relation as risk factors to develop T2D in Mexican indigenous groups. Specifically, this study was focused eleven populations: Cora, Chatin, Chontal, Maya, Mazahua, Mixteco, Otomí, Purépecha, Teenek, Yaqui, and Zapoteco. The total of the indigenous population evaluated was 879 individuals, where 675 of them did not have T2D. On the first stage, we searched sequence variants of *HNF4A* and *GCK* genes in Mazahua and Teenek groups through the simple chain conformational polymorphism technique and, then, the found variants were sequenced by the Sanger method. Based on these results, we determined the frequency of the nonsynonymous variants of the stated genes in the remaining groups.

Ten genetic markers were employed for the ancestry estimation. A statistical correlation between the presence of the found non-synonymous variants and metabolic risk factors for the development of T2D was done.

Respecting the ancestry estimation, we found, on average, a 97.6% of an Amerindian component. Concerning the evaluation of the *HNF4A* gene during the first stage, three changes were discovered: rs736824 and rs745975 (located in the intronic region), and the third one was the rs1800961 that corresponds to the T130I polymorphism. These three variants are present in both the Mazahua and Teenek populations. Regarding *GCK* gene we

detected three variants: D3D (rs144723656) and Q26H (rs193922329) in the Mazahua group, and Y215Y (rs142553382) in the Teenek group; these changes were found in three nonrelated individuals. It is worth noting that the prediction *in silico* of the Q26H variant did not refer changes that impact in the protein function.

On a previous study from our research group, the T130I (rs1800961) variant of *HNF4A* gene was associated with early onset T2D in Mexican Mestizo population besides to its high frequency. For this reason, we determined the frequency in the rest of population in the second stage. In the overall indigenous population evaluated, the T130I frequency was on average 16.6%, being the Maya and Cora groups who had the highest frequency (28.2% and 25.7%, respectively) and it was not found in the Yaqui group. The frequency comparison between the populations of the International Hap Map Project and the ones of this study showed a statistically significant difference, concerning the T130I (rs1800961) variant. In Chatino, Chontal and Cora groups, an association between T130I (rs1800961) variant and high triglycerides concentrations ($P < 0.05$) was found. This association was kept considering all nondiabetic indigenous population ($P < 0.05$).

The obtained results evidence the presence of genetic variants in Mexican indigenous population that are related to risk factors for the development of T2D. In particular, the T130I (rs1800961) variant of *HNF4A* gene is associated with high triglyceride concentrations. Further studies will be required in pursuance of evaluating the impact of intronic variants of *HNF4A* gene and its susceptibility to the development of T2D. The inclusion of the indigenous populations from Mexico in genetic studies will allow these communities to participate in the benefits that this type of results can bring in the prognosis and treatment of the disease.

I. Antecedentes

1. Introducción

La diabetes se define como una enfermedad crónico-degenerativa, ocasionada por diversos factores entre los que se encuentran la predisposición genética y los factores ambientales. Se caracteriza clínicamente por hiperglucemia crónica debido a la deficiencia en la producción o acción de la insulina, lo que afecta al metabolismo de los carbohidratos, proteínas y grasas [1].

La gran mayoría de los casos de diabetes se dividen en dos amplias categorías etiopatogénicas. La diabetes tipo 1 (DT1), en donde existe una deficiencia absoluta de secreción de insulina. Las personas con mayor riesgo de desarrollar este tipo de diabetes a menudo pueden ser identificadas por evidencia serológica de un proceso patológico autoinmune que ocurre en los islotes pancreáticos y por marcadores genéticos [2]. En la otra categoría, que es mucho más prevalente, la diabetes tipo 2 (DT2), la combinación de resistencia a la acción de la insulina y una respuesta secretora de la insulina compensatoria inadecuada parece ser la causa [3]. En la DT2, una hiperglucemia sin síntomas clínicos causa cambios patológicos y funcionales en varios tejidos diana y puede estar presente durante un largo período de tiempo antes de que se detecte. Durante este período asintomático, es posible demostrar una anormalidad en el metabolismo de los carbohidratos mediante la medición de la glucosa en plasma en ayunas, por hemoglobina glucosilada (HbA1c), o después de un desafío con una carga de glucosa oral [4].

Los síntomas de hiperglucemia, que son las características clínicas clásicas de la diabetes incluyen: poliuria, polidipsia, pérdida de peso, a veces polifagia y visión borrosa. La susceptibilidad a ciertas infecciones también pueden acompañar a la hiperglucemia crónica.

Además de DT1 y DT2, la diabetes se ha clasificado en otros tipos, de acuerdo al fenotipo clínico, los antecedentes familiares, el genotipo y/o las causas ambientales específicas, entre ellas se encuentran: la diabetes gestacional, la diabetes tipo MODY (Maturity Onset Diabetes of the Young), diabetes por endocrinopatías, entre otras [4].

1.1 Costos de la diabetes en México

Actualmente más de 415 millones de personas en el mundo tienen DT2. Para el 2040 se predice que el número de individuos afectados puede incrementar a 642 millones, principalmente en los países de bajo y mediano ingreso [5]. No obstante los avances en el tratamiento, la diabetes es causante de enfermedad renal crónica, ceguera, amputación de miembros y es el principal factor de riesgo cardiovascular, infarto y defectos en el nacimiento [6]. El costo económico de la DT2 en México muestra tendencias crecientes al igual que otros países en Latinoamérica (figura 1) [7]. El reporte final de la ENSANUT 2016 [8] no da cifras del gasto de la diabetes en México, sin embargo, de la ENSANUT 2012 se sabe que el costo anual de diabetes en 2011 fue de cerca de \$8,000 mil pesos por persona al año, por lo que para 2012 se requirieron aproximadamente \$48,000 mil millones de pesos para el manejo de este padecimiento [9]. Por otra parte, el estudio del Instituto Mexicano para la Competitividad (IMCO) señala que el costo del tratamiento médico de personas mayores a 20 años, con DT2 ocasionada por obesidad, fue de más de \$62,000 mil millones en el año 2015 [10] a este costo, se suman los gastos originados por las complicaciones, estimados en \$26,000 mil millones anuales. Se calcula también, que a causa de la diabetes cada año se pierden un promedio de 312 millones de horas laborales, lo que equivale al 25 % de los empleos que se generaron en el 2014 [10]. Por su parte, la Fundación Mexicana para la Salud en 2015 publicó el estudio Carga Económica de la Diabetes Mellitus en México, en el cual se desglosan los costos económicos de la atención de este padecimiento en nuestro país en ese año, de tal modo que considerando los costos directos e indirectos se estimó el gasto en \$362,859 millones, cifra mayor al crecimiento real anual de la economía mexicana [11]. De acuerdo a lo anterior, se puede observar que existen diversas estimaciones económicas a lo largo del tiempo y que coinciden con la importancia de la magnitud del problema al cual nos estamos enfrentando y que representa además de toda la carga económica una disminución muy importante en la calidad de vida de las personas que padecen DT2.

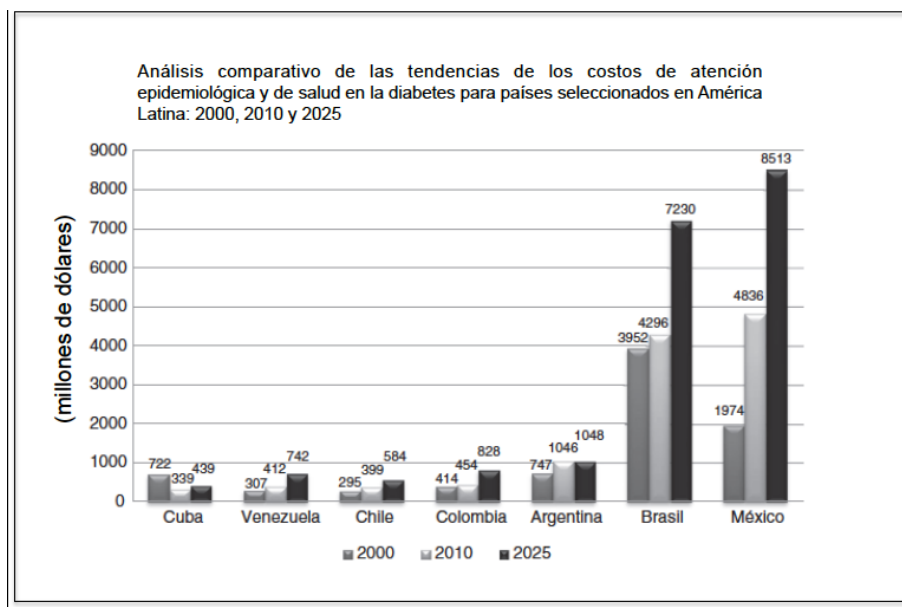


Figura 1. Análisis comparativo de las tendencias de los costos sanitarios en la DT2 para países de América Latina: 2000, 2010 y 2025 (millones de US \$) [7]

1. 2 Prevalencia de DT2 en México

La prevalencia de diabetes en México ha ido en aumento durante las últimas décadas. Los reportes de la Organización para la Cooperación y el Desarrollo Económico (OCDE) desde el 2015 [12] refieren que nuestro país tuvo la prevalencia más alta (15.9%) entre los miembros de esa organización, así mismo que el inicio temprano de la diabetes en México es más frecuente que en ningún otro país 23.8% en personas de 40 a 59 años y 5.9% en personas de 20 a 39 años, las cifras más altas de la OCDE, cuyos promedios son de 8.9% y 1.7%, respectivamente [12]. El inicio temprano de la DT2 tiene importantes implicaciones para el estado de salud de un paciente, así como para su condición social y económica.

Por su parte, los resultados de la ENSANUT de Medio Camino 2016 [8] muestran una prevalencia de 9.4% de personas con diabetes, lo que representa un ligero aumento con respecto a la ENSANUT 2012 (9.2%)[9] y un aumento mayor con respecto a la ENSANUT 2006 (7.2%) [13] en la prevalencia por diagnóstico médico previo. Como puede observarse las cifras de la OCDE son mayores que las reportadas por la ENSANUT, y de lo cual se desprende que la realización de estudios epidemiológicos futuros serán muy importantes

para evitar que exista una subestimación de la prevalencia de diabetes en México y no se alcance la cobertura médica a todas las personas afectadas.

Por otra parte, la Federación Internacional de la diabetes (IDF) coloca a México dentro de los diez primeros países con mayor número de personas con diabetes en el mundo (12 millones de personas afectadas) ocupando el quinto lugar, sin embargo la proyección de la IDF para el 2040 nos coloca en el cuarto lugar con 21.8 millones de afectados, lo cual representará una carga económica muy grande para el sector salud [5].

En la ENSANUT 2016, las complicaciones que los diabéticos reportaron en mayor proporción fueron: visión disminuida (54.5%), daño en la retina (11.2%), pérdida de la vista (9.95%), úlceras(9.1%) y amputaciones (5.5%) [8]. Todas éstas, excepto daño en la retina, se reportaron en mayor proporción en comparación con lo reportado en la ENSANUT 2012 [9].

La situación actual y el panorama estimado para los próximos años en nuestro país hace imperativo impulsar a todos los niveles medidas que ayuden a disminuir la aparición de nuevos casos, la mortalidad y mejorar la calidad de vida de las personas afectadas. Así los resultados de las investigaciones de corte clínico y genético serán de suma importancia en el campo de la diabetes.

1. 3 Fisiopatología de la DT2

La DT2 surge de una alteración de la capacidad del músculo, la grasa y el hígado para responder a la insulina, combinada con una incapacidad de la célula β para secretar apropiadamente esta hormona. Aunque la contribución relativa de estos dos defectos a la patogénesis de la DT2 continúa debatiéndose, los estudios longitudinales en individuos con alto riesgo de DT2 sugieren que la resistencia a la insulina (RI) es un fenómeno temprano, ocurriendo años antes de cualquier evidencia de intolerancia a la glucosa, mientras que la falla en la célula β se desarrolla más tarde durante la enfermedad [3, 14].

El avance de la tolerancia normal a la glucosa (TNG) hacia una DT2 tiene dos tipos distintos de hiperglucemia intermedia, aunque en cierta medida superpuesta: alteración de la glucosa en ayunas (AGA) y alteración de la tolerancia a la glucosa (ATG). AGA es principalmente el resultado de la resistencia hepática a la insulina y un defecto de fase temprana en la función de las células β , y es diagnosticada con valores de glucosa

plasmática en ayunas entre 5,6 mmol/l y <7,0 mmol/l. La ATG es principalmente el resultado de la resistencia a la insulina periférica y la disfunción progresiva de las células β , y se diagnostica cuando a las dos horas después de una prueba de tolerancia oral a la glucosa con una carga de 75 g, la persona presenta concentraciones de glucosa >7,8 mmol/l y <11,1 mmol/l [3, 14]. Hay evidencia de que la prevalencia de AGA y de ATG es diferente entre los diferentes grupos étnicos, así mismo también es influenciada por el sexo y la edad. Por ejemplo, la ATG es más común en las mujeres, mientras que la AGA lo es en los hombres [15].

La alteración de la acción de la insulina y de la función de las células β ocurre muy temprano en el desarrollo de DT2. La resistencia a la insulina (RI) puede detectarse en personas con tolerancia normal a la glucosa que tienen un mayor riesgo de desarrollar DT2 de 10-20 años antes de que se diagnostique la enfermedad. Además, las personas que están pasando de una tolerancia alterada de la glucosa a la DT2 pueden haber perdido hasta el 80% de la función de sus células β [16]. Uno de los factores que se cree que desempeñan un papel importante en este sentido es la edad, numerosos estudios han demostrado una disminución relacionada con la edad en la función de las células β y la secreción de insulina. Esto es consistente con la mayor prevalencia de DT2 con el envejecimiento [17].

La dieta hipercalórica y el sedentarismo; las dietas ricas en carbohidratos de rápida absorción dan como resultado niveles elevados de insulina y glucosa en la sangre, adicionalmente el aumento de grasa en el hígado y el músculo incrementa la RI en estos tejidos. Estos factores aumentan la demanda de insulina y mayor actividad de las células β , lo que a largo plazo, puede llevar a una falla progresiva de éstas [18].

Las elevadas concentraciones plasmáticas de ácidos grasos libres (AGL) también alteran la secreción de insulina en pacientes con riesgo de desarrollar DT2. Los AGL elevados en las células β conducen a un aumento del estrés oxidativo y la apoptosis debido a su lipotoxicidad [19].

Por otra parte, el advenimiento de las técnicas genómicas ha permitido la identificación de varios genes asociados con la disfunción de la insulina y las células β en pacientes con DT2, incluidas variantes genéticas asociadas con el desarrollo pancreático y el almacenamiento y secreción de insulina [20]. Se han propuesto que diversos polimorfismos genéticos en

individuos predispuestos a DT2 producen una falla de la célula β lo que la imposibilita para adaptarse a la mayor demanda de insulina [20, 21].

1.3.1 Regulación de la glucosa por el hígado

Diversos órganos y tejidos participan en la homeostasis de la glucosa, por lo cual alteraciones en cualquiera de ellos puede contribuir a incrementar las concentraciones de glucosa en un individuo. El hígado es el principal órgano responsable de la producción de glucosa. La síntesis y liberación de glucosa hepática a la circulación proviene de la gluconeogénesis y la glucogenólisis. El aumento de las tasas de producción de glucosa hepática en el diabético son el resultado de un desequilibrio en varios factores afectando los diferentes aspectos implicados en la regulación hepática de glucosa, entre estos, se encuentran: 1) incremento de los sustratos gluconeogénicos, 2) resistencia del hígado a la acción de la insulina (RI) que conduce a la inhibición de la supresión de la producción de glucosa hepática, (la RI se adquiere temprano en la progresión de DT2, inicialmente la tolerancia a la glucosa es normal debido a un aumento compensatorio en la secreción de insulina) y 3) la falta de supresión de la secreción de glucagon posprandial de las células pancreáticas α en pacientes con DT2 [22, 23].

1.3.2 Regulación de la glucosa en músculo esquelético

En el músculo esquelético el transporte de glucosa estimulado por insulina es el mecanismo principal para la eliminación de una carga de glucosa exógena. El principal transportador involucrado en la captación de glucosa en el músculo esquelético, es el transportador de glucosa 4 (GLUT4). La insulina y el ejercicio estimulan de forma aguda la translocación de GLUT4 a la membrana de las células musculares, lo que aumenta la absorción de glucosa. En pacientes con DT2, el músculo esquelético es resistente a las acciones de la insulina debido a defectos en su señalización y a menudo un bajo nivel de actividad física ocasiona menor translocación de GLUT4 a la membrana [24]. Esto conduce a una disminución en la absorción de glucosa que contribuye al desarrollo de la hiperglucemia [24, 25].

1.3.3 Regulación de la glucosa en tejido adiposo

La insulina es un potente inhibidor de la lipólisis y frena la liberación de AGL del adipocito inhibiendo la enzima lipasa sensible a hormonas. En personas con DT2 la capacidad de la insulina para inhibir la lipólisis y para reducir la concentración plasmática de AGL está

notablemente alterada. Los incrementos crónicos en AGL estimulan la gluconeogénesis, inducen RI en el hígado y el músculo y alteran la secreción de insulina en el páncreas. El tejido adiposo disfuncional produce cantidades excesivas de citocinas inflamatorias y aterogénicas que pueden contribuir a la inducción de la RI y al mismo tiempo no secreta adecuadamente adipocitocinas que sensibilizan a la insulina [26].

1.3.4 Regulación de la glucosa en intestino

El péptido 1 similar al glucagon (GLP-1) y el polipéptido insulínico dependiente de glucosa (GIP) (incretinas) son hormonas liberadas por el tracto gastrointestinal en respuesta a la ingestión de nutrientes. GLP-1 y GIP actúan sobre las células β para estimular la liberación de insulina y son responsables de hasta 60% de la secreción de insulina después de una comida. GLP-1 también promueve la saciedad, ralentiza el vaciado gástrico e inhibe la secreción de glucagón, reduciendo así la producción de glucosa hepática. Los pacientes con DT2 tienen secreción alterada de GLP-1 y una respuesta reducida a GIP. Esto da como resultado un aumento de la motilidad gastrointestinal, disminución de la secreción de insulina dependiente de glucosa, aumento de la secreción de glucagón y aumento de la liberación de glucosa en el hígado, todo lo cual afecta negativamente el control glucémico [27, 28].

1.3.5 Regulación de la glucosa en riñón

En condiciones normales, más del 99% de la glucosa filtrada por los riñones se reabsorbe en el túbulo proximal. La mayor parte de la glucosa es reabsorbida por el SGLT2, junto con el transportador de glucosa GLUT2. Una vez que las concentraciones de glucosa en plasma superan el umbral renal para la reabsorción (aproximadamente 180 mg/dL en individuos sanos), la glucosa comienza a aparecer en la orina. Un estudio reciente demostró que la capacidad del riñón para reabsorber glucosa aumenta en pacientes con DT2 en comparación con individuos sanos. Por lo tanto, en pacientes con DT2, los riñones reabsorben la glucosa en exceso y la devuelven a la circulación, lo que puede empeorar la hiperglucemia [29]. El hígado y los riñones son los únicos órganos que poseen las enzimas necesarias para la gluconeogénesis y que posteriormente liberan glucosa recién formada en la circulación. En el estado preprandial en humanos sanos, la gluconeogénesis renal representa aproximadamente el 20% de la glucosa total liberada en la circulación, y el

hígado contribuye con el resto. Se ha sugerido que la síntesis renal de glucosa aumenta en pacientes con DT2 en relación con individuos sanos. Así, en pacientes con DT2, los riñones pueden exacerbar aún más la hiperglucemia al continuar la reabsorción de glucosa y aumentar la producción de glucosa [29-31].

De esta manera, en la homeostasis de la glucosa participan mayormente los órganos mencionados, los cuales son capaces de mantener una estabilidad interna que les permite llevar a cabo sus funciones. El contenido de glucosa, tanto en las células como en el organismo completo, depende de la entrada y de su consumo, flujo que se encuentra regulado por señales que mantienen el estado energético. La regulación metabólica celular preserva los niveles de energía en equilibrio, lo cual es un proceso finamente regulado que se refleja en el organismo completo. Romper entonces este equilibrio desemboca en la aparición de enfermedades como la DT2.

2. Participación de los genes en la patogénesis de la DT2

El desarrollo de DT2 es el resultado de la interacción entre factores ambientales y un fuerte componente genético incluso con las mismas exposiciones ambientales, algunas personas son más susceptibles que otras a desarrollar diabetes, y este riesgo incrementado parece ser heredado. Pero si bien los factores genéticos desempeñan claramente un papel en el desarrollo de la diabetes, las variantes genéticas implicadas en este riesgo eran completamente desconocidas antes del advenimiento de las tecnologías genéticas modernas. El avance de los estudios genéticos humanos en la década de 1980 finalmente permitió intentar identificar los *loci* que subyacen a este componente genético [32].

A diferencia de DT1, donde el riesgo genético se concentra principalmente en la región HLA, el componente genético del riesgo de DT2 no se concentra en una región específica y parece ser el resultado de la interacción de múltiples genes a lo largo de todo el genoma. Es posible y muy probable que el componente genético de la DT2 se deba a múltiples variantes comunes de efecto pequeño (hipótesis variante común de enfermedad común), sin embargo puede ser que también contribuyan variantes raras de mayor efecto [33-35].

2. 1 Heredabilidad en la DT2

Un gran número de estudios a nivel mundial apoyan el concepto de la heredabilidad en la DT2, ya que ésta se presenta en familiares consanguíneos de generación en generación. Así,

el riesgo de tener DT2 se incrementa cuando hay una historia familiar positiva de la enfermedad. Aproximadamente del 15-25% de los parientes en primer grado de pacientes con DT2 desarrollan ya sea ATG o diabetes. El estudio Framingham revela que los niños que tienen un padre con DT2 tienen 3.5 más de riesgo de desarrollar la enfermedad comparado con un individuo de la población general y 6.1 veces más si ambos padres tienen DT2 [36]. Resultados de prevalencia de diabetes e intolerancia a la glucosa en una cohorte de 199 descendientes de padres diabéticos muestran que la ocurrencia de diabetes en ellos se acercaba al 60% a la edad de 60 años [37]. Estudios con gemelos han resultado de mucha utilidad para el conocimiento de la herencia en la aparición de la DT2. Se ha demostrado que la tasa de concordancia para la diabetes difiere entre el tipo de parejas de gemelos y ésta se eleva cuando se adiciona un factor de riesgo. Así, los rangos, entre 35% y 58% en los gemelos monocigotos se elevan al 88% cuando la tolerancia a la glucosa se altera y es de 17% a 20% en los gemelos dicigóticos [37-39]. A este respecto, la Dirección General de Epidemiología en México, reportó que entre las características epidemiológicas de los casos con DT2 reportados en el 2016, el 84.3 % de ellos tienen antecedentes heredofamiliares y el 54.1% el familiar es de primer grado (figura 2) [40].

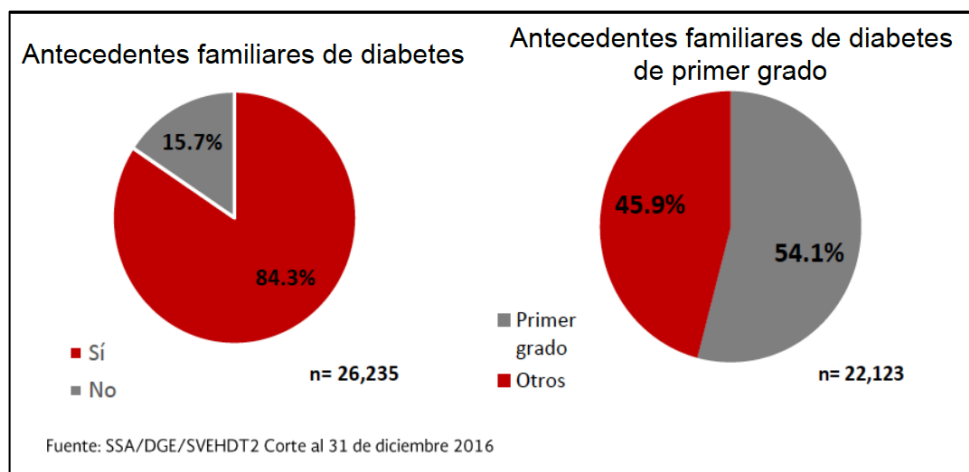


Figura 2. Antecedentes heredofamiliares de diabetes en pacientes reportados por la DGE. México 2016 [40]

Lo descrito anteriormente apoya la evidencia del papel del componente genético en el riesgo a desarrollar DT2. Además, el conocimiento de que la prevalencia de la enfermedad varía sustancialmente entre grupos étnicos que comparten el mismo ambiente, refuerza la teoría que los factores genéticos contribuyen a la predisposición de la enfermedad [41, 42].

Así, las prevalencias varían desde el 1% en algunas poblaciones indígenas, como en los indios mapuches o Chinos que viven en zonas rurales, a valores extremadamente altos como los encontrados en los Nauru (25%)[43] y los Pima en Arizona (50%) [44]. Sin embargo, el estudio del componente genético en la DT2 es complejo por la participación de múltiples genes en el origen de la enfermedad, así dependiendo del factor de riesgo o vía metabólica que se afecta actualmente pueden agruparse aquellos genes que se asocian a RI, disfunción de la célula β , obesidad y sedentarismo entre otros (figura 3) [45].

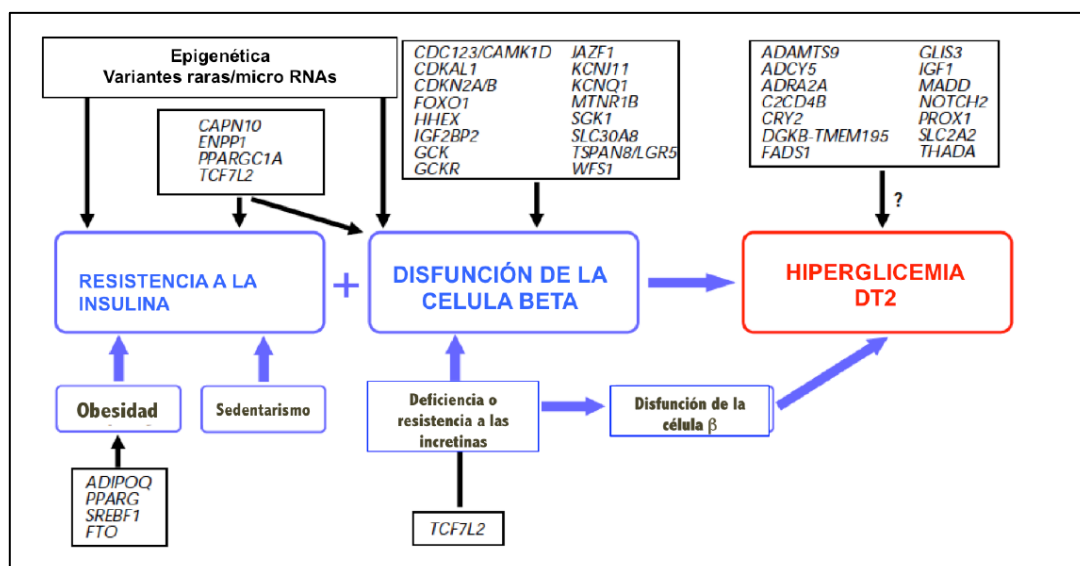


Figura 3. Polimorfismos de un Solo Nucleótido (SNPs), como factores de riesgo para DT2 [45].

Con el enfoque de genes candidatos, PPARG fue el primer gen identificado de riesgo para DT2. Posteriormente, con el uso de estudios de asociación del genoma completo, más de 120 *loci* se han relacionado con la DT2 [46]. Además, 53 *loci* se han relacionado con las concentraciones de insulina y glucosa, de los cuales 33 están asociados con la DT2, aunque algunos de ellos tienen relación con la obesidad y la resistencia a la insulina, la mayoría están relacionados con la función de las células β [47].

Entre las variantes que han sido más estudiadas como factores de susceptibilidad para padecer DT2 se encuentran la de los genes de *KCNJ11* (rs5219), que codifica *Kir6.2* (E23K), en *TCF7L2* (rs12255372), el polimorfismo Gly972Arg en *IRS1*, el polimorfismo Gly1057Asp en *IRS2*, el polimorfismo Trp64Arg en el gen que codifica el receptor β_3 adrenérgico y la

variante promotora -308 G/A en *TNFA* [32]. Además otros genes han sido ampliamente estudiados con relación al origen de la DT2, entre ellos se encuentra el factor nuclear de hepatocitos *HNF4A* el cual codifica para un factor de transcripción conocido como una molécula clave en el desarrollo y la función de las células β . En una exploración de todo el genoma realizada previamente en gemelos japoneses con DT2 se observó la vinculación de DT2 con el cromosoma 20q12-q13, una región en la que se encuentra el gen *HNF4A*, estudios recientes informan asociaciones entre DT2 y polimorfismos en la región promotora P2 de este gen específica para las células β [48, 49].

Las asociaciones genéticas con diversas patologías, permiten una visión profunda de los procesos moleculares subyacentes a la enfermedad, la cual es informativa en el desarrollo de un nuevo diagnóstico y de técnicas terapéuticas para mejorar el cuidado del paciente. Los estudios encaminados a dilucidar el papel que tienen los polimorfismos o mutaciones en el proceso de la enfermedad serán de gran utilidad para el clínico y repercutirán en un buen pronóstico para las personas afectadas.

En este trabajo se han elegido dos genes *HNF4A* y *GCK*, factores muy importantes en el funcionamiento de la célula beta, para la síntesis y secreción de la insulina. *HNF4A* participa además en la homeostasis del colesterol y triglicéridos, así como en el metabolismo de los ácidos biliares, y es un gen clave para la expresión de lipoproteínas como *APOC3* y *MTTP* [50-52] En un estudio realizados por nuestro grupo, en personas mestizas de la Ciudad de México con DT2 se encontró la asociación de una variante de este gen T130I (rs1800961) con esta enfermedad [53].

En el caso de la enzima *GCK*, debido a sus características cinéticas y funcionales que la hacen única para mantener el equilibrio de la glucosa sanguínea en los tejidos donde se expresa y que en la célula beta es primordial para la liberación de la insulina después de recibir una carga de glucosa, la hacen un gen de interés para la búsqueda de polimorfismos que se asocien con factores de riesgo para padecer DT2 [54] o más específicamente que mutaciones de este gen sean la causa directa de diabetes [55]. Por otra parte en nuestro país son escasos los estudios realizados que evalúen la participación de este gene en la DT2 en los pueblos originarios de México.

2.2. HNF4A como gen de riesgo para DT2

El factor hepático nuclear 4A (HNF4A) es un factor de transcripción que se expresa en hígado, riñón, intestino y páncreas, y es un receptor huérfano de la superfamilia de receptores nucleares $2\alpha 1$, NR2A1, aunque existen estudios que sugieren a los ácidos grasos como probables ligandos [56]. La proteína HNF4A presenta una estructura con seis diferentes regiones (referidas de la región A-F). Estas regiones corresponden al dominio funcional: un dominio N-terminal de transactivación (AF-1)(región A/B, aminoácidos 1-50), un dominio de unión al DNA (DBD)(región C, aminoácido 50-116), la función precisa del dominio D (aminoácidos 116-174) y F (aminoácidos 174-465) es incierta. El complejo funcional C-terminal (región E, aminoácidos 128-366) que forma un dominio de unión al ligando (LBD), una interfase de dimerización y un dominio de transactivación (AF2) (figura 4) [57].

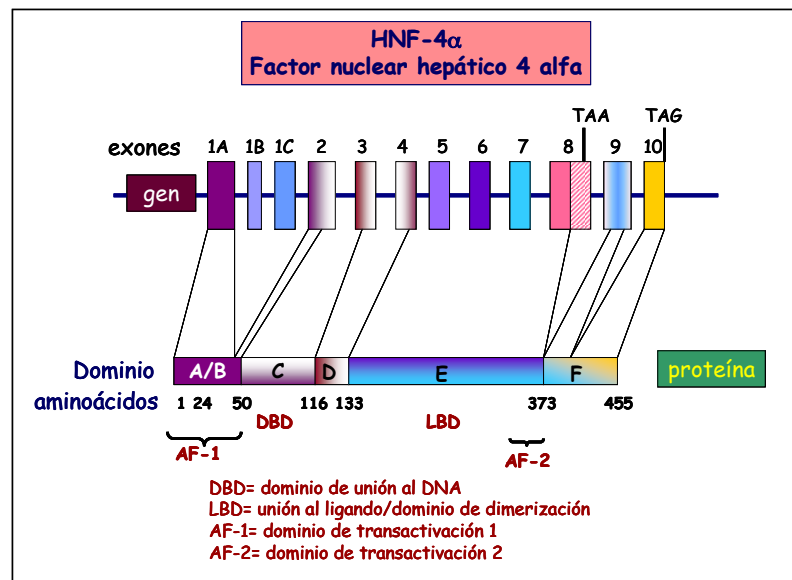


Figura 4. Estructura de HNF4A, gen y proteína [58]

El dominio de unión al DNA de HNF4A consiste en un motivo de dos dedos de Zinc que interactúan específicamente con el elemento de respuesta del DNA de muchos promotores y potenciadores de genes blanco.

Se conoce también que HNF4A participa en los estadios tempranos de la embriogénesis, de acuerdo con los estudios realizados en especies de *Xenopus* y *Drosophila*. En el ratón la

expresión de HNF4A se observa en el endodermo primario en el día 4.5. La pérdida del gen de *HNF4A* en murinos es letal, debido a la disfunción del endodermo visceral [59].

En el humano, el gen de *HNF4A* se localiza en el cromosoma 20 q13.1-13.2. Consiste de 12 exones. Se han identificado 9 diferentes isoformas (HNF4A1-HNF4A9) como se observa en la figura 5, que son productos de variaciones en el procesamiento de su mensajero y el uso de un promotor alternativo (P1 ó P2), las cuales tienen diferente distribución y función tejido específico, estos promotores están físicamente separados por más de 45 kb. El promotor 1 inicia la transcripción del exón 1A (isoformas A1-A6) en el hígado y riñón, mientras que el promotor P2 inicia la transcripción que contiene el exón 1D (isoformas A7-A9). Esta diferencia causa que las isoformas P2 sean más cortas que las isoformas P1 y, lo que es más importante, que carezcan del dominio de interacción del cofactor diseñado como función activadora (AF)-1. Estas diferencias estructurales sugieren que las isoformas P1 y P2 podrían tener distintas funciones al interactuar de manera diferente con cofactores específicos y estar reguladas de manera diferente en contextos específicos de importancia fisiológica. El promotor P2 está predominantemente relacionado a la expresión de HNF4A en los islotes pancreáticos [60].

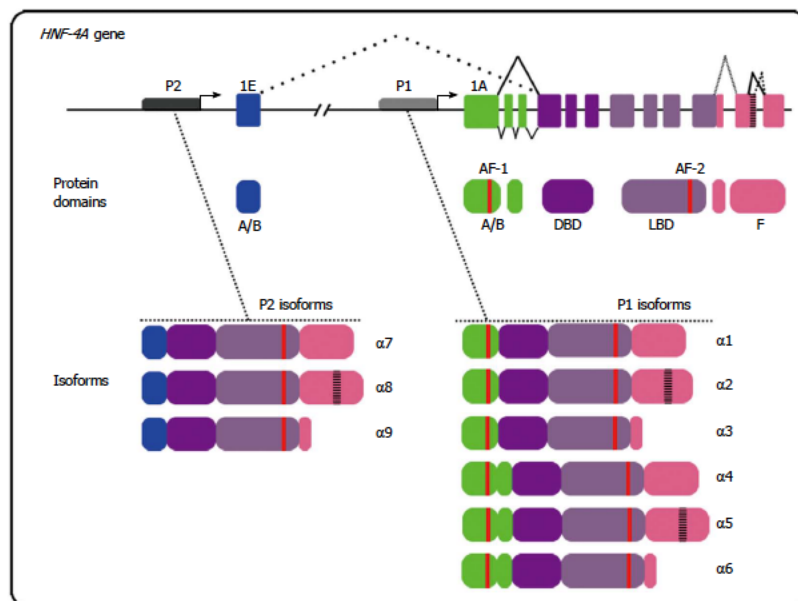


Figura 5. Representación de las isoformas de HNF4A (HNF4A1-HNF4A9). Los promotores P1 o P2 dirigen la transcripción alternativa de HNF4A [60].

HNF4A tiene un papel crítico en el desarrollo, diferenciación celular y metabolismo del páncreas endocrino. Es esencial para el buen funcionamiento del intestino, hígado y riñón. Se han identificado alrededor de 910 genes blanco para *HNF4A* en hígado y 665 en páncreas, y pueden ser agrupados en diferentes categorías de acuerdo con su función, como transportadores de nutrientes y metabolismo, hemostasia sanguínea, función inmune, diferenciación hepática y factores de desarrollo [61]. Los genes blancos mejor caracterizados son aquellos que están relacionados con el transporte de lípidos (gen de apolipoproteínas, *apo CIII*) y el metabolismo de la glucosa (*LPK*, *PEPCK*, y la *Aldolasa B*).

En ratones *knock-out*, la eliminación específica de *HNF4A* en células β , no lleva al desarrollo de diabetes, pero después de la administración intraperitoneal de glucosa los ratones *knock-out* presentan una falla en la secreción de insulina comparados con los ratones de la cepa silvestre [62].

Mutaciones en *HNF4A* son causa de enfermedades en humanos como la diabetes de inicio temprano MODY 1 y la hemofilia [57]. La identificación de factores de transcripción como causa de formas dominantes de DT2, revelan la existencia de una red transcripcional regulatoria, involucrada en el metabolismo (síntesis y secreción de la insulina), en el desarrollo y diferenciación de la célula β . En pacientes con DT2 se ha observado menor expresión del RNAm de *HNF4A*, lo que se ha asociado con un daño en la secreción de insulina [63].

Hasta el día de hoy, se han reportado 72 diferentes mutaciones en el gen de *HNF4A* [64].

En México se han encontrado 3 variantes en el exón 4 de *HNF4A*, N126H/Y, R154G (rs765237979) y T130I (rs1800961) asociadas con la aparición de DT2 [65, 66]. La variante T130I (rs1800961) fue descrita en 1996 por el grupo de Yamagata y cols., en personas no diabéticas de ascendencia no hispana, sin embargo, no fue asociada con la diabetes MODY1, lo cual fue motivo del estudio. La variante T130I consiste en el cambio de Treonina (ACT) por Isoleucina (ATT) en el codón 130 [67].

Posteriormente en el 2003 se llevaron a cabo las pruebas funcionales para evaluar el impacto en la función de la proteína ocasionada por la variante T130I (rs1800961), se encontró la disminución de la función transcripcional en un 46.2% en la línea celular HepG2 [68]. El residuo treonina es un aminoácido conservado del dominio de unión al DNA, en una

región llamada box-A. La región box-A es importante para la homodimerización y para la afinidad de unión al DNA (figuras 6 y 7). El residuo de treonina esta conservado en humano, rata, ratón y *Xenopus* y permite la interacción de dos sitios de fosforilación por PKC, mientras que la sustitución por isoleucina sólo facilita un sitio, afectando la activación de HNF4A [69, 70].

Adicionalmente, la transcripción de *GLUT2* está mediada por *HNF4A*, por lo que la falla de este factor de transcripción podría afectar también el control de la secreción de insulina [71].

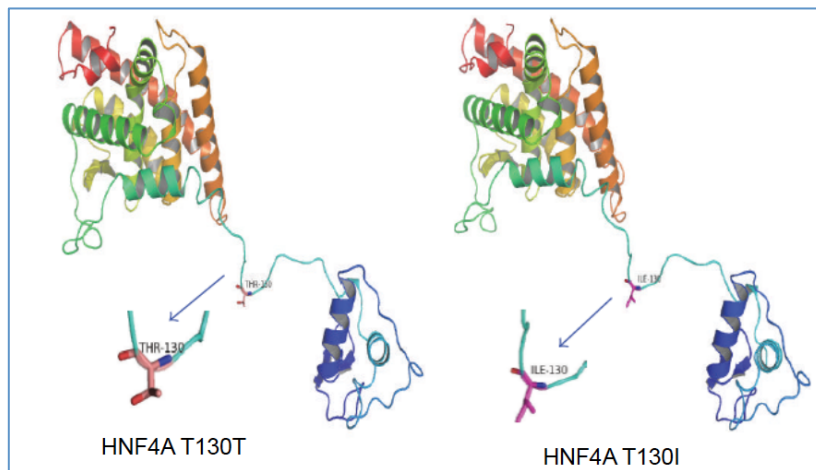


Figura 6. Homología de la proteína de HNF4A con y sin la variante rs1800961 [69].

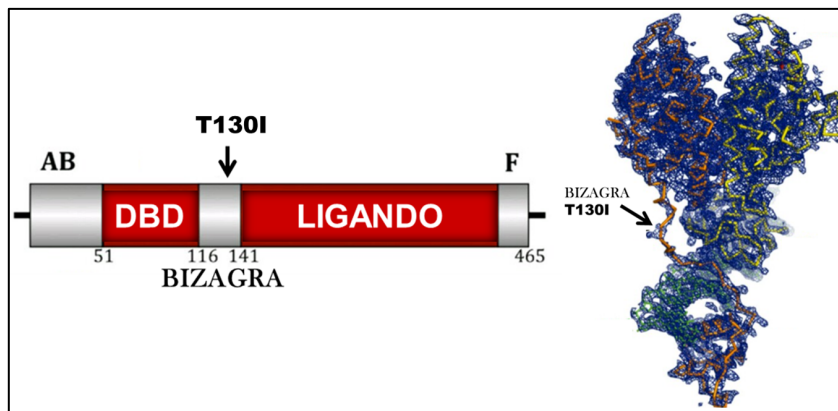


Figura 7. Representación lineal de los dominios de la proteína HNF4A y localización tridimensional de la variante T130I. DBD: dominio de unión a DNA; AB: N terminal; F: C terminal. Modificado de [70].

Previo a este proyecto, nuestro grupo de trabajo mostró que la variante T130I (rs1800961) de *HNF4A* se asocia con el desarrollo de DT2 de aparición temprana (en menores de 40

años) en población mestizo Mexicana [53]. No obstante que se han llevado a cabo estudios en diversas poblaciones, (tabla 1), existe poca información sobre el impacto que tiene esta variante en población indígena mexicana como factor de riesgo para el desarrollo de DT2.

Debido a la importancia funcional que tiene el gen *HNF4A* en la homeostasis de la glucosa, en este trabajo se exploró la posibilidad de encontrar variantes genéticas en la población indígena mexicana las cuales podrían formar parte del fondo diabetogénico en México.

Tabla 1. Frecuencia de la variante T130I (rs1800961) en diferentes poblaciones en personas con y sin diabetes.

Población	N	T130I %	Autor
Controles blancos no hispanos EUA	55	5.0	Yamagata (1996) [67]
Controles Daneses	239	3.8	Moller (1997) [72]
Diabéticos Daneses	509	4.7	Moller (1997) [72]
Controles Chinos	251	7.8	Rissanen (2000) [73]
Diabéticos Chinos	105	0.0	Rissanen (2000) [73]
Controles Japoneses	354	0.8	Zhu (2003) [74]
Diabéticos Japoneses	423	3.5	Zhu (2003) [68]
Controles Daneses	4726	6.6	Ek (2005) [75]
Diabéticos Daneses	1409	8.2	Ek (2005) [75]
Controles mestizos mexicanos	75	5.6	Menjívar (2008) [53]
Diabéticos mexicanos	100	16.0	Menjívar (2008) [53]

2.3 Glucocinasa (GCK) como gen de riesgo para DT2

El primer paso de la glucólisis en los mamíferos es catalizado por cuatro diferentes hexocinasas. Las hexocinasas I-III tienen un peso aproximado de 100 kDa, tienen una relativa afinidad por su sustrato, y son inhibidas por concentraciones fisiológicas de glucosa-6-fosfato (G6P). En contraste la hexocinasa IV, más comúnmente conocida como Glucocinasa (GCK), tiene un peso molecular de alrededor de 50 kDa, una $K_{catS_{0.5}}$ que es similar a la concentración de glucosa sanguínea, y no es afectada significativamente por las concentraciones fisiológicas de G6P [76].

La GCK se expresa en hepatocitos y células β , en acción concertada con GLUT2 permite una entrada rápida de la glucosa independiente de la insulina. Los individuos con mutaciones en el gen *GCK* (*MODY2*) tienen una deficiencia en la detección de las concentraciones de glucosa, lo que conduce a un umbral elevado para el inicio de la secreción de insulina estimulada por glucosa. Esto da como resultado una hiperglucemia pero que no predispone

a las complicaciones de la diabetes. En la célula β , la GCK es un componente clave en el metabolismo de la glucosa. La fosforilación de la glucosa dentro de la célula β está estrechamente acoplada a la secreción de la insulina [77, 78].

La GCK consiste en una cadena o subunidad de 448 aminoácidos que forman una molécula monomérica que consta de 13 hélices alfa y 5 láminas beta que pueden fosforilar glucosa y otras hexosas. La cadena se pliega en dos regiones distintas, un dominio pequeño y grande. La GCK tiene un sitio activo de unión para glucosa y uno para ATP, que es la fuente de energía para la fosforilación (Figura 8). Este sitio de unión activo se encuentra entre los dominios pequeños y grandes [79]. El carboxilo terminal es parte de la hélice alfa 13, que codifica la región que forma la mitad del sitio de unión para la glucosa. Esta enzima se puede modular para formar un complejo inactivo y activo. La conformación inactiva se forma cuando la hélice alfa 13 se ha modulado lejos del resto de la molécula formando un espacio grande. Este espacio es demasiado grande para unir glucosa por lo que se dice que está en forma inactiva. La alternativa es cuando la hélice alfa 13 se modula para formar un espacio más pequeño, lo que activa la proteína [79].

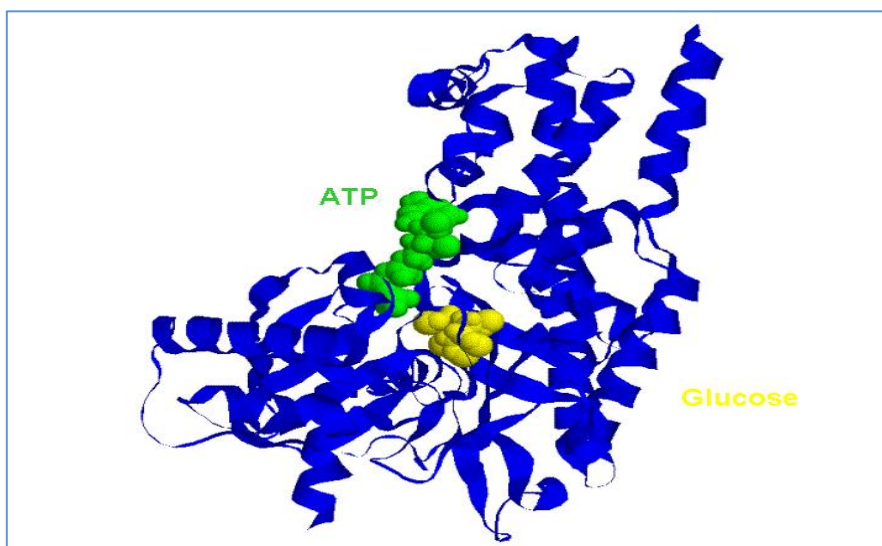


Figura 8: Estructura de GCK cuando se une a ATP y glucosa [79].

El gen de *GCK* se localiza en el cromosoma 7p15-p13, consta de 12 exones que se expresan tanto en la célula β como en el hepatocito (figura 10). A través de un sitio alternativo de transcripción, las formas de GCK en hepatocito y célula β difieren en su extremo N-terminal, codifica para una proteína de 465 aminoácidos con un peso molecular de 52,191 Da. La

presencia de promotores específicos de tejido permite la regulación diferencial dando lugar a tres versiones de diferentes tamaños del exón 1 (a, b y c); el promotor cadena arriba es funcional en el páncreas, mientras que el promotor cadena abajo se usa solo en el hígado. Los exones 1b y 1c se expresan en el hígado y el exón 1a se expresan en la célula β pancreática (Figura 9) [80, 81]. La insulina sobre-regula la expresión de GCK en el hígado [82].

La GCK hepática también es regulada por la proteína reguladora de GK (GKRP). Se sugiere que el incremento en las concentraciones de fructosa-1-P seguida de la ingesta de fructosa rompe la unión GKRP-GK, lo que produce la translocación de GCK del núcleo al citoplasma donde puede promover la captación de la glucosa plasmática [83].

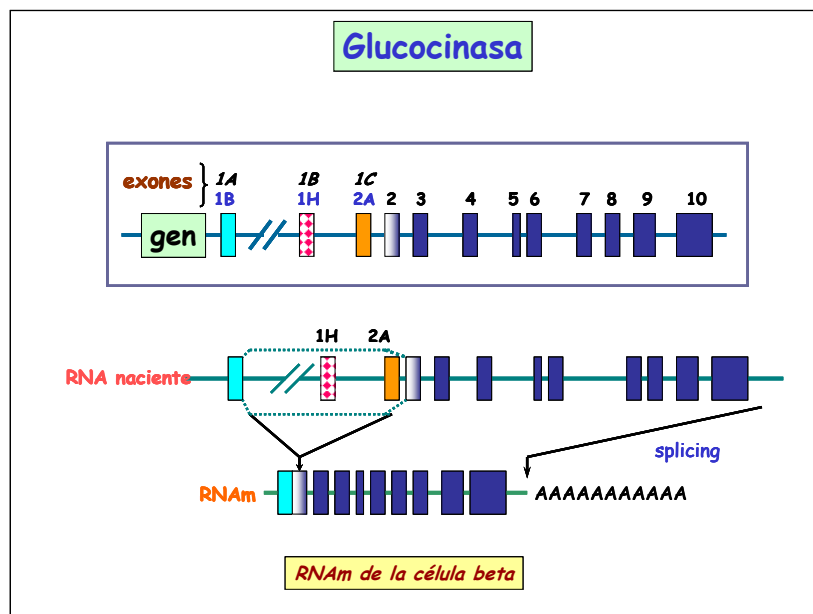


Figura 9. Estructura del gen GCK de la célula β . Incluye los exones 1A y del exón 2 al 10 [58]

Debido al papel central que tiene GCK en la liberación de insulina es entendible que las mutaciones en este gen puedan causar hiperglucemia (GCK; MIM# 138079).

En 1992, 14 de 16 familias con el fenotipo MODY tuvieron un LOD score positivo para el estudio de ligamiento de *GCK* con MODY, así, que las mutaciones relacionadas con este gen que llevaban a producir diabetes se le llamo MODY 2 [84].

Las mutaciones de *GCK* pueden ser bioquímicamente clasificadas como: 1) mutaciones que no producen cambios cinéticos pero que conducen a formas inestables con una actividad

más baja, 2) mutaciones que producen cambios cinéticos, y 3) mutaciones que cambian múltiples parámetros cinéticos de la función enzimática. Las mutaciones en *GCK* hacen que la célula β sea menos sensible a las concentraciones extracelulares de la glucosa. Existen también mutaciones en *GCK* que producen una ganancia de la función de la proteína lo que conlleva a hiperinsulinemia neonatal o familiar provocando una significativa hipoglucemia [85].

Los pacientes que presentan alguna mutación de pérdida de la función en este gen tienen hiperglucemias moderadas en condiciones de ayuno o post-prandiales, la curva de tolerancia a la glucosa es anormal y se ha encontrado un componente hereditario autosómico dominante. A la fecha 130 mutaciones han sido descritas para el gen de la *GCK* en familias tipo MODY [86].

Por otra parte, se ha sugerido también que un defecto en la sensibilidad a la glucosa por el páncreas, causada por una mutación heterocigota en el gen de *GCK*, puede reducir el desarrollo fetal y el peso al nacimiento, causando además hiperglucemia después del nacimiento [87].

No obstante la importancia de *GCK* en la diabetes MODY2, sólo en 2013 se ha informado su relación con la DT2 en un meta-análisis, donde se combinaron 24 estudios incluyendo 88,229 casos y 210, 239 controles los resultados demostraron que existía una asociación modesta entre la variante -30G>A de *GCK* y el riesgo de DT2 [88]

En México se han llevado a cabo pocos estudios respecto a este gen, Del Bosque-Plata en 1997, realizó el primer estudio en la búsqueda de variantes en familias mestizas con características de diabetes MODY; sin embargo no encontraron ninguna variante en la región exónica [89], lo cual no excluye que variantes en este gen estén en la población mexicana. De esta manera y debido a la importancia de esta proteína en la homeostasis de la glucosa, en este trabajo se decidió realizar una búsqueda de variantes genéticas en este gen en población indígena mexicana.

3. Población indígena mexicana

La población mestiza mexicana es el resultado de la mezcla de individuos de diversos orígenes. Los componentes principales de este mestizaje fueron los indígenas mesoamericanos y los españoles que llegaron en el siglo XVI. A ellos se han sumado los

africanos, que llegaron en condición de esclavos y los asiáticos cuyo flujo incrementó a finales del siglo XIX. Estudios realizados por Lisker y cols., señalan que la población mexicana posee una proporción muy grande de ancestría amerindia, que va de 56% en la Ciudad de México hasta 76% en Tlaxcala [90], por lo que es de esperarse que este componente indígena influya en muchos aspectos, incluyendo la susceptibilidad a enfermedades en la población mestiza mexicana.

Se estima que en México hay aproximadamente 25.7 millones de indígenas, lo que constituye cerca del 11% de la población mexicana, lo cual representa al grupo autóctono más grande de América Latina [91]. En el país se hablan 65 lenguas y dialectos. Los Estados del Centro-Sur-Sureste concentran el 80% de los individuos indígenas. Existen grupos etno-lingüísticos con gran número de población (nahuas, mayas mixtecos, zapotecos, otomíes, tzeltales, tzotziles: entre 1,400,000 y 300,000), en contraste con pequeños grupos: kiliwas, paipai, cochimi, kukapás, kikapus, entre otros con menos de 500 habitantes. Los indígenas presentan los mayores índices de marginación del país. En los últimos años, muestran una intensa migración regional, nacional e internacional. En el contexto de la salud, la esperanza de vida en los indígenas, es siete años inferior al resto de la población (69 años indígenas/76 años mestizo) y además presentan altas tasas de desnutrición crónica. La tasa de mortalidad en diabetes, comparada con el mestizo mexicano es de 16 vs 36 por cada 100,000 habitantes [92].

En México son escasos los análisis para conocer la situación social y de salud de los pueblos indígenas, sin embargo se sabe que existen desigualdades en las condiciones de vida, la situación de salud y el acceso a los servicios de salud entre la población indígena y la no indígena [8, 9]. Con respecto al tema de la diabetes, resultados de nuestro grupo revelaron que el 60 % de los indígenas mayas que tiene DT2 desconoce que tiene esta enfermedad [93]. Por otra parte, resultados recientes de Robles-Ordaz muestran la alta prevalencia del estado prediabético de una comunidad Seri de Sonora (47.1%), [94] también resultados del proyecto Maycoba en 2015, muestran el incremento en la prevalencia de diabetes en mujeres Pima de México en un período de 15 años 9.4% (1995) vs. 13.4% (2010) [95].

Lo anterior revela la importancia de realizar estudios genéticos y epidemiológicos en poblaciones socioeconómicamente vulnerables como los pueblos indígenas mexicanos, con

la finalidad de poder brindar un servicio de salud más eficiente y de mayor cobertura en estas comunidades.

3.1 Participación de los genes en la susceptibilidad a la DT2 en población Mexicana

Se ha demostrado que los hispanos tienen una gran sensibilidad para desarrollar enfermedades complejas como la DT2, lo cual se ha atribuido a su componente amerindio [93, 96].

La principal contribución observada del fondo genético a la DT2 en la población mexicana se debe a variantes comunes con un efecto menor. Se conoce, que el 69% de los *loci* asociados tienen MAF (frecuencia del alelo menor) >5% con un OR (razón de momios) entre 1.21 y 3.38. Esto sugiere que la existencia de un conjunto de variantes genéticas con efectos menores tiene un papel importante en la patogénesis de la DT2 en nuestra población (figura 10). Hoy en día, 23 variantes comunes se han relacionado con DT2 en México [97, 98]. Este número, en comparación con los encontrados en poblaciones caucásicas, sugiere que aún hay más variantes comunes por descubrir. Las variantes de susceptibilidad que han sido replicadas en población mestiza mexicana son: rs7903146 (*TCF7L2*), rs10811661 (*CDKN2A/2B*), rs7754840 (*CDKAL1*), rs12779790 (*CDC123/CAMK1D*), rs13266634 (*SLC30A8*), rs74800100 (*LOC387761*), rs7923837 (*HHEX*), rs4402960 (*IGF2BP2*) y rs2237897 (*KCNQ1*). Es importante hacer notar que la variante rs10811661 (*CDKN2A/2B*) ha sido replicada por nuestro grupo de trabajo, encontrando asociación con la aparición de DT2 en población maya [93].

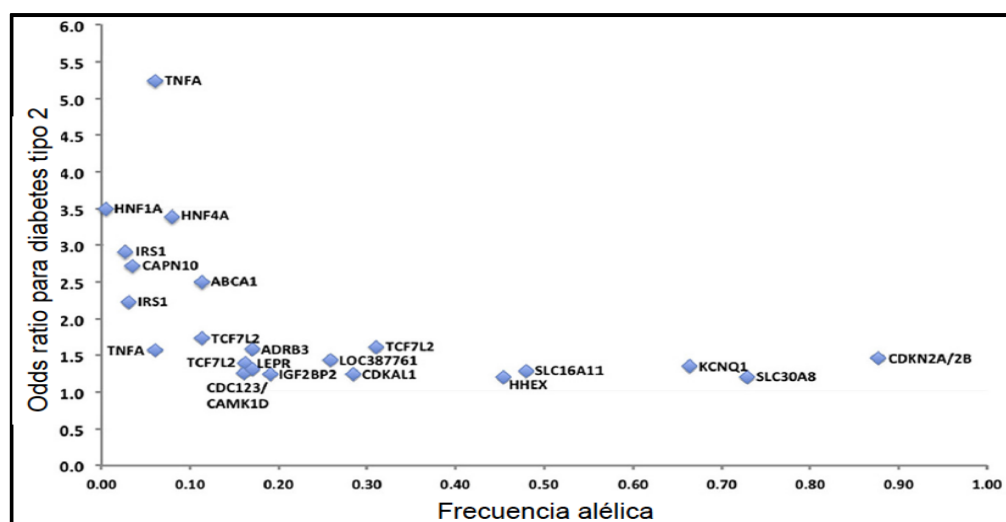


Figura 10. *Loci* asociados a DT2 en población mexicana. (Modificado de [97]).

Aunque se han realizado esfuerzos importantes para el entendimiento de la genética de la DT2 en México, los estudios en población indígena aún son escasos. Sería de gran ayuda la identificación de variantes propias de la población mexicana que ayudarán a la explicación del incremento de la prevalencia de DT2 en nuestro país, además que este tipo de estudio permite la inclusión de comunidades en situación social y económicamente desfavorable en los futuros beneficios que puedan obtenerse de ellos.

II. Justificación

El presente proyecto se plantea en el contexto de un problema nacional de salud, la diabetes tipo 2. Los resultados de la ENSANUT 2016, han mostrado aumento en su incidencia, la cual es mayor que el promedio mundial. Asimismo, existen evidencias que señalan la participación de genes específicos en el desarrollo de la DT2 en la población mexicana.

Se conoce que los genes responsables del fenotipo diabético son distintos y se presentan con diferente frecuencia entre las poblaciones. En México los estudios que evalúan el fondo genético de esta enfermedad en los 62 grupos étnicos aún son pocos y es necesario que se estime el impacto de las variantes ya conocidas y si es posible identificar nuevas variantes propias de la población mexicana que puedan dar una explicación de la alta prevalencia de la DT2 en los mexicanos, sobre todo en las comunidades social y económicamente vulnerables que son las más afectadas por esta patología.

El papel primordial que tienen los genes *HNF4A* y *GCK* en el funcionamiento de la célula beta para mantener la homeostasis de la glucosa les confieren importancia para ser estudiados como factores de riesgo para el desarrollo de DT2, además porque en la población mestiza mexicana se ha reportado una variante de *HNF4A* asociada a la DT2 de aparición temprana (< 40 años).

Este trabajo es parte de una línea de investigación que caracteriza en diferentes grupos indígenas mexicanos las variantes de secuencia en genes asociados al riesgo del desarrollo de DT2.

III. Hipótesis

Existen variantes no sinónimas aún no reportadas en los genes *HNF4A* y *GCK*, las cuales se asocian con factores de riesgo para el desarrollo de la DT2 en población indígena mexicana.

IV. Objetivos

General

Identificar variantes de secuencia no sinónimas en los genes *HNF4A* y *GCK* en grupos indígenas mexicanos y relacionarlas con factores de riesgo para el desarrollo de DT2.

Específicos

- ✓ Identificar variantes de la secuencia exómica de los genes *HNF4A* y *GCK* en los grupos indígenas mazahua y teenek
- ✓ Predecir la probabilidad de que los polimorfismos encontrados tengan impacto funcional en la proteína
- ✓ Establecer la frecuencia de las variantes encontradas en poblaciones indígenas de México
- ✓ Determinar y comparar las frecuencias de las variantes identificadas con otras poblaciones reportadas en el Proyecto Internacional HapMap
- ✓ Establecer la relación estadística de las variantes encontradas con factores de riesgo metabólicos para presentar DT2

V. Material y Métodos

Requerimientos y Consideraciones éticas.

El protocolo fue evaluado y aprobado por el Comité de Ética de Investigación en Humanos del Hospital Juárez de México HJM1402/07.08.22 (Anexo 1). A todos los participantes del estudio se les explicó el proyecto y cada persona firmó o imprimió su huella digital en una carta de consentimiento informado de participación voluntaria (Anexo 2). El muestreo se realizó en población abierta con la condición de que tuvieran un ayuno de por lo menos 8 horas (muestreo por conveniencia). Se obtuvieron datos generales a través de la aplicación de un cuestionario (Anexo 3). En casos necesarios, se requirió la ayuda de un intérprete.

Diseño del proyecto. El proyecto se diseñó en dos fases. En la primera fase se identificaron las variantes no sinónimas de los genes *HNF4A* y *GCK*. Se evaluaron 91 individuos del grupo mazahua del poblado de Jaltepec en el Estado de México y 69 del grupo teenek de Tanzuman y San Isidro del estado de San Luis Potosí, estos grupos fueron los primeros propuestos para realizar esta tesis.

De acuerdo a los resultados de la primera fase, en una segunda fase, se procedió a la búsqueda de la frecuencia de la variante T130I (rs1800961) de *HNF4A*, el cual fue el único cambio no sinónimo encontrado en los mazahua y teenek y del que se conoce que su presencia modifica la funcionalidad de su proteína. La predicción *in silico* de la variante Q26H de *GCK* no predice cambios en la función de la proteína.

La búsqueda de la frecuencia se realizó en los grupos: **cora** (Presidio de los Reyes y San Juan Corapan del Estado de Nayarit), **chatin** (Santos Reyes, Nopala, Estado de Oaxaca), **chontal** (Nacajuca, Estado de Tabasco), **maya** (Hecelchakán, Pocho, Pomuch y Tenabo de Campeche), **mixteco** (Santa Elena Xamiltepec, Estado de Oaxaca), **otomí** (Hidalgo), **purépecha** (Michoacán), **yaqui** (Sonora) y **zapoteco** (Oaxaca). Las muestras de los grupos: purépecha (N = 25), zapoteco (N = 25) y yaqui (N=25), fueron proporcionados por el Dr. Canizales y el Dr. Víctor Acuña y del grupo otomí (N =25) por la Dra. Rosenda Peñaloza, cabe mencionar que para estos grupos sólo se contó con la muestra de DNA

Sujetos:

Criterios de inclusión

1. Individuos con tres o más generaciones residiendo en su población de origen.
2. Hablantes de lengua materna según su grupo indígena, así como considerarse a sí mismos pertenecientes a su grupo étnico
3. Ser mayores de edad
4. Sin restricción de género
5. Estar de acuerdo en participar en el estudio y haber firmado una carta consentimiento informado

Tamaño de muestra

El cálculo de tamaño de muestra para este estudio, se realizó aplicando la siguiente ecuación que es utilizada para estimar el tamaño de muestra para la estimación de proporciones. (Daniel, Bioestadística, Ed. Limusa Wiley).

$$n = \frac{z^2 pq}{d^2} = \frac{(1.96)^2(0.5)(0.5)}{(0.05)^2} = \underline{384.16}$$

z= distribución normal (1.96)

p= 0.5, q= 0.5 y d=0.05.

El resultado de la ecuación indica que se necesitaron evaluar: 384 individuos

Determinaciones antropométricas

A cada individuo se le midió la circunferencia de la cintura (cm) en el punto medio entre el borde inferior de la costilla y la cresta iliaca y la circunferencia de la cadera (cm) en el punto más ancho sobre los trocánteres mayores. Se registró el peso (kg) y la talla (m), y se calculó el índice de masa corporal (IMC) mediante la fórmula que divide el peso entre el cuadrado de la talla (kg/m²) e índice cintura cadera (ICC) dividiendo la medida de la cintura entre la cadera.

Determinaciones bioquímicas

A cada participante se le extrajo una muestra de sangre periférica en ayuno, de aproximadamente 5 ml en un tubo sin anticoagulante y gel separador con la finalidad de

obtener el suero de cada individuo. Se midió la concentración sérica de glucosa, colesterol, y triglicéridos por métodos convencionales con estuches comerciales y de acuerdo con las especificaciones del fabricante. Los niveles de insulina se determinaron por quimioluminiscencia usando el sistema de inmunoensayos Immulite® 2000 de Siemens.

Análisis genético

Extracción de DNA

Se extrajeron 10 ml de sangre periférica, y se depositaron en tubos con 200 µl de EDTA para aislar leucocitos. La obtención del DNA genómico se llevó a cabo por métodos estándar [99]. El DNA se cuantificó espectrofotométricamente (260/280 nm) y se corrieron geles para evaluar pureza e integridad respectivamente. El DNA se utilizó para: a) Evaluar etnicidad a través de marcadores de ancestría, b) realización de la técnica de PCR-SSCP, c) secuenciación y d) RFLP para búsqueda del polimorfismo T130I.

Análisis del polimorfismo conformacional de cadena simple (SSCP)

La SSCP se llevó a cabo para cada uno de los exones de los genes *HNF-4A* y *GCK*. El corrimiento electroforético se realizó en geles al 5.4% y 8%, de poliacrilamida en ambos casos con glicerol al 10% y sin glicerol, se corrieron a 200 V, de 16-24 h a temperatura ambiente dependiendo de la concentración de poliacrilamida. En cada muestra se preparó una reacción estándar de 25 µl conteniendo 1 µg de DNA genómico, 25 µM de cada uno de los dNTP's, 1 µl de dCTP marcado con P³², DMSO, buffer 10X (Tris -HCL 10mM, pH 8.3), KCl 50mM, MgCl₂, y Taq polimerasa 2 U (Perkin Elmer, E.U). Todos los exones se amplificaron en un termociclador (Ependorff-Gradiente, Alemania).

Esta técnica esta basada en el corrimiento electroforético en geles de poliacrilamida de las cadenas de DNA simples, en los cuales se detectan patrones de migración diferentes si se ha generado la sustitución de una base por otra en la cadena de DNA del exón en estudio, el rango de sensibilidad del SSCP varía entre 80–90% para detectar mutaciones en fragmentos menores de 200 pares de bases. El SSCP es un método capaz de identificar la variación de un sólo nucleótido en un segmento de DNA [100].

Oligonucleótidos empleados para amplificación del gen *HNF4A*

Tabla 2. Oligonucleótidos empleados para amplificación de *HNF4A* [67]

Exon	5' - 3'	3' - 5'	pb	TA °C
E1	GGGCACTGGGAGGAGGCAGT	GCCTGTAGGACCAACCTACC	340	65
E1B	TCTGGTGTGCACGACTGCAC	CTGGAGCTGCAGCCTCATAAC	356	64
E2	AAGGCTCCCTTAGATGCCTG	CCACTCAGGGAGAAGACAGACCT	321	67
E3	CCTAGTTCTGTCTAAGAGG	GTCATAAAGTGTGGCTACAG	253	59
E4	CCACCCCCTACTCCATCCCTGT	CCCTCCCGTCAGCTGCTCCA	271	68
E5	GTGCAGGGGACAGAGAATGC	AATCAAGCCAGTCCACGGCTAT	322	66
E6	GCCCAGCGTCACTGAGTTGGCTA	TTGCCTGGGTGAGTGCCATG	234	64
E7	GCACCAGCTATCTTGCCAAC	AGGAGAAGTCTGGCAGAGCG	315	66
E8	CTCCTTGTGTGACACAAGTC	CTCACTGTGTGAGGCCTGTC	407	61
E9	TGGTTGATTGGCCACGCCTG	ATCCTGGTTCTACCTTCTAG	341	59
E10	CATTTACTCCCACAAAGGCT	GACCACGTGATCACCAGGTG	277	59

Tabla 3. Oligonucleótidos empleados para amplificación de los exones de *GCK* [101]

Exon	5' - 3'	3' - 5'	pb	TA °C
E1	TCAAACCCCAAACCAGCCC	AGCAAACACTCCCAGAATGCC	237	61
E1B	AGCTCTCCATCCAAGCAGCC	GCTCTCCCAGTGCAAAG	220	61
E1C	CCTCACATGGCCAACTGCTAC	CTTTGGTGAAAGCGATAAGAGTC	200	59
E2	CGGGGTCAGAAGACAGAAGG	GGGTGCTTCTCCCAGCTAGG	381	61
E3	CCTCCTCCTCTTTGTAATATCC	TAGACAGGTGGCACCTCC	348	59
E4	CATTCAGTGGCCAGGTGTT	GGGGGCTACATTTGAAGGC	320	59
E5	AGCAGCCACGAGGCCTATCTC	GAGAAAGGCAGGCAGTGTGG	196	64
E6	CCAGCACTGCAGCTTCTGTG	GAGCCTCGGCAGTCTGGAAG	176	64
E7	TGCAGCTCTCGCTGACAGTCC	CTCCCATCTGCCGCTGCACC	287	64
E8	CGTGCCTGCTGATGTAATGG	GCCCGAGACCAAGTCTGC	268	64
E9	CTGTGCGGAGCGACACTCAG	CCCCCAAATCTAGGCCAAGG	410	64
E10	GTCGACTGCGTGCAGGGCGC	TGTGGCATCCTCCCTGCGCT	263	64

Secuenciación

De acuerdo con los resultados del corrimiento electroforético de los productos de PCR de cada exón en la SSCP, se secuenciaron las muestras que presentaron diferente patrón de migración. La reacción de secuenciación se llevó a cabo mediante un estuche de AmpliCycle Sequencing de Perkin Elmer (E.U.), y se basa en el método descrito por Sanger [102], se hizo la secuenciación automática en el secuenciador ABI-PRISM Applied Biosystems versión 3.7.

Marcadores de ancestría (AIMs)

Todos los SNPs para ancestría fueron genotipados utilizando el KASPar ensayo (KBioscience Company), utilizando un formato de transferencia homogénea de energía fluorescente. Se seleccionó un panel de 10 marcadores para estimar la ancestría de los individuos participantes, estos marcadores distinguen principalmente la ancestría indígena americana y la europea con un $\delta > 0.44$: rs2695, rs2862, rs3340, rs4884, rs1800498, rs2814778, rs223830, rs203096, rs722098 y rs17203 [103]. Las poblaciones de referencia fueron la YRI y CEU del proyecto HapMap. Como control de calidad se realizó la genotipificación por duplicado en 90 muestras y no se encontraron genotipos discordantes.

Frecuencia de la mutación T130I del gen *HNF4A* por RFLP

El exón 4 de *HNF4A* fue amplificado por PCR con oligonucleótidos específicos 5'-CCACCCCTACTCCATCCCTGT-3', y 5'-CCCTCCCGTCAGCTGCTCCA-3'. La mutación T130I genera un sitio de corte *Sfu* I y es detectada por RFLP (polimorfismo de los fragmentos de restricción) (T: 271 pb, I: 81+190 pb). La digestión para la enzima *Sfu* I se llevo a cabo en un volumen de 25 μ l con 1 μ g de DNA, buffer H (Tris-HCl 500mM, NaCl 1M, MgCl₂ 50mM pH7.5), agua desionizada y una unidad de enzima por μ g de DNA (concentración de la endonucleasa 10 U/ μ l), la mezcla fue incubada a 37 °C toda la noche. Los fragmentos se visualizaron a través de un gel de agarosa al 3% en Buffer TBE 1X utilizando bromuro de etidio.

Análisis de la Extensión de la Homocigidad Haplotípica

Para hacer este análisis, se secuenciaron 2000 pares de bases alrededor de la variante en estudio, para ello se diseñaron oligonucleótidos específicos y se llevó a cabo una secuenciación Sanger, con los datos obtenidos se realizó un alineamiento en la página de UCSC Genome Browser (<http://genome.ucsc.edu/>), se conformó un haplotipo con

rs7269793, rs6093978, rs13041396, rs11574738, rs11574739, rs3212194 y rs3212195 de acuerdo a lo reportado en el Proyecto 1000 genomas (<https://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/>)[104]. Las pruebas basadas en haplotipos para el análisis de selección positiva se calcularon con el paquete R, rehh versión 2.0 [104]. Las variantes genéticas sin información sobre su alelo ancestral fueron excluidos de los análisis. El diagrama de EHH fue obtenido usando el paquete R, rehh. En este análisis solo fueron considerados los Mazahua, Teenek y mestizos grupos de los cuales se contaba con la secuencia de datos.

Análisis de datos

El análisis estadístico se realizó mediante los programas SPSS 10.0 2.00 para la estadística descriptiva y para el análisis de correlación. La predicción de los cambios no sinónimos se llevo a cabo con los programas Polyphen V2.2 [105], SIFT [106] y FATHMM [107]. La correlación de los genotipos con rasgos fenotípicos cuantitativos se llevó a cabo a través del modelo lineal univariado ajustando por posibles confusores como sexo y edad entre otros. La estimación de ancestría se realizó en el programa Adxmiture [108]. Se calculó el equilibrio de Hardy-Weinberg por una prueba de X^2 considerando un valor significativo $p < 0.05$.

VI. Resultados

Ancestría

La estimación de la ancestría posee gran relevancia para conocer la estructura genética de las poblaciones, en este trabajo se evaluó la ancestría como una herramienta para confirmar que los participantes poseen una ancestría mayoritariamente amerindia, y evitar que los resultados obtenidos de las posibles asociaciones no se deban a un efecto de estratificación poblacional.

La tabla 4 muestra los resultados de la estimación de ancestría expresada como porcentaje. Los 11 grupos estudiados presentan una estimación amerindia superior al 96%. Los grupos Teenek y Zapoteco alcanzaron el 98%, lo que sugiere que estas poblaciones conservan aún un fondo genético propio de los primeros pueblos de América.

En México existen 68 grupos indígenas mexicanos a lo largo de todo el país, que pueden clasificarse de acuerdo a la lengua en 11 familias lingüísticas indoamericanas, los individuos de este estudio pertenecen a las familias Yuto-nahua, Oto-mangue, Maya y Tarasco.

En la figura 11 se encuentra la localización geográfica de los 11 pueblos evaluados, estos representan el 16.2 % del total de los grupos indígenas en México.

Tabla 4. Estimación de la ancestría de 11 grupos indígenas mexicanos

Grupo	n=954	ancestría amerindia %	ancestría europea %	Familia Lingüística
Cora	124	97.8	2.2	Yuto-nahua
Chatino	112	97.4	2.6	Oto-mangue
Chontal	141	97.7	2.3	Maya
Maya	109	97.5	2.5	Maya
Mazahua	91	97.6	2.3	Oto-mangue
Mixteco	133	97.7	2.3	Oto-mangue
Otomí	25	96.6	3.4	Oto-mangue
Purépecha	25	97.6	2.4	Tarasco
Teenek	69	98.0	2.0	Maya
Yaqui	25	97.6	2.1	Yuto-nahua
Zapoteco	25	98.1	1.9	Oto-mangue
	n=879	97.6	2.4	



Figura 11. Localización geográfica de los 11 grupos indígenas mexicanos evaluados

Evaluación exómica de *HNFA4* y *GCK* en los grupos mazahua y teenek

HNFA4

Para cada exón se verificó que el tamaño del amplificado correspondiera a lo reportado en la literatura. Para ello se corrieron geles de agarosa al 1% durante 30 minutos a 80 V. Cada uno de los exones analizados tuvieron el tamaño esperado. La figura 12 muestra un ejemplo de la verificación del tamaño de los exones. Para el exón 2 el tamaño esperado es de 321 pb, el exón 4, 271 pb y el exón 5, 322 pb.

Identificación de variantes

Se realizaron 450 reacciones de PCR marcadas con P^{32} y 80 corrimientos electroforéticos (SSCP) de los 10 exones de *HNFA4* para identificar las muestras que tuvieran un cambio en su composición de bases, lo cual se hace evidente por un corrimiento distinto. Como control de calidad y para asegurar que este tipo de ensayo no tuviera alguna interferencia ya sea por factores químicos de las soluciones o físicos como el voltaje, cada corrida se llevo a cabo en 4 condiciones de concentración de acrilamida: 5.4% y 8% de acrilamida y en presencia o ausencia de glicerol. La figura 13 muestra el PCR preparativo para la búsqueda

de variantes en el exón 4 de *HNF4A*, se verifica también que las muestras amplificadas tengan un tamaño de 271 pb.

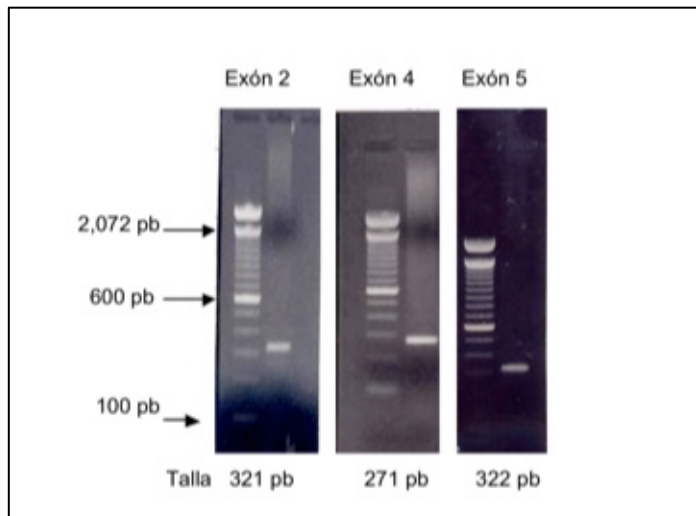


Figura 12. Tamaño de los exones 2, 4 y 5 de *HNF4A*. Gel de agarosa 1%; 80 V/30 minutos; visualización con bromuro de etidio. Marcador: 100 pb

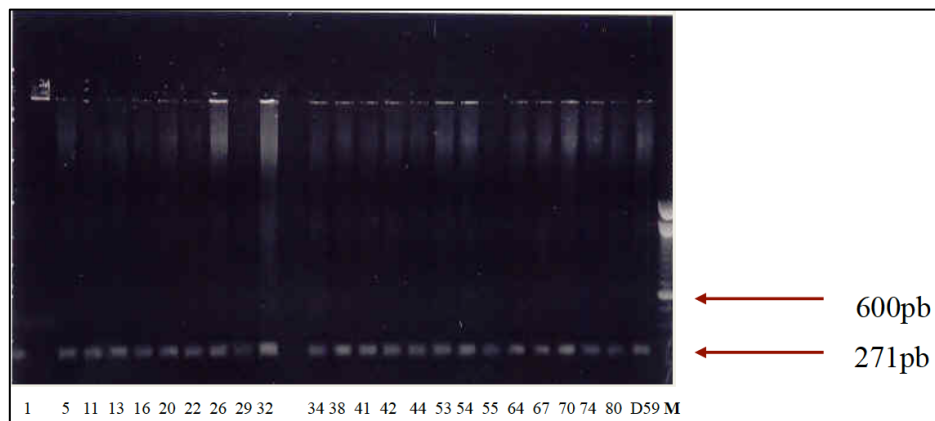


Figura 13. Corrimiento electroforético del PCR preparativo del exón 4 de *HNF4A*. Gel de agarosa 1%; 80 V/30 minutos; visualización con bromuro de etidio, tamaño del exón 271 pb.

En las figuras 14 y 15 se observan dos ejemplos de las autoradiografías obtenidas para identificar las muestras en donde hay cambios en la secuencia de los exones 2 y 4 de *HNF4A* por la técnica de SSCP, como se puede observar en el exón 2 encontramos 3 patrones diferentes y en el exón 4 la muestra MZ20 presenta 3 bandas, mientras que las muestras restantes sólo presentan una banda.

Aquellas muestras en donde se observaron corrimientos diferentes fueron secuenciadas por el método de Sanger.

En las figuras 16, 17 y 18 se muestran las secuencias del exón 2 de las muestras MZ 4, MZ 1 y MZ 5 que tuvieron modificaciones en el corrimiento electroforético.

El electroferograma de la muestra MZ 4 (figura 16) es la reportada como de referencia por el GenBank con número de acceso [AH005099](#). La secuencia de la figura 17 corresponde a la muestra MZ1. Este DNA presenta dos cambios en su secuencia, comparada con el banco de genes. En la posición 242 están presentes las dos bases Timina (T) y Citocina (C), (el banco de genes lo informa como C), lo que nos indica que esta persona es heterocigota para este cambio. En la posición 273, hay un cambio de una Guanina (G) por Adenina (A). Estos dos polimorfismos ya se han reportado con anterioridad (del proyecto Genoma Humano), sin embargo estas modificaciones no corresponden a algún cambio en la secuencia de la proteína. La figura 18 muestra el electroferograma de la muestra MZ5, este DNA presenta un cambio en su secuencia en la posición 273 donde hay un cambio de una G por A como en la MZ1, pero el corrimiento electroforético es diferente de la muestra MZ1 porque es la única modificación presente. Al igual que en la muestra anterior este polimorfismo esta reportado dentro del proyecto genoma humano y no causa ninguna modificación en la proteína.

En la figura 19 se muestra el electroferograma de la muestra MZ20 del exón 2. En ella observamos un cambio en la posición 345 de la secuencia. Se indica por una flecha la posición donde se observa la presencia de 2 nucleótidos. La curva azul indica la presencia de la base **C** en un alelo, y la curva roja superpuesta indica la presencia de la base **T** en el otro alelo (C/T), lo que origina un cambio en el codón 130 de la proteína de HNF4A, Treonina (ACT) por Isoleucina (ATT).

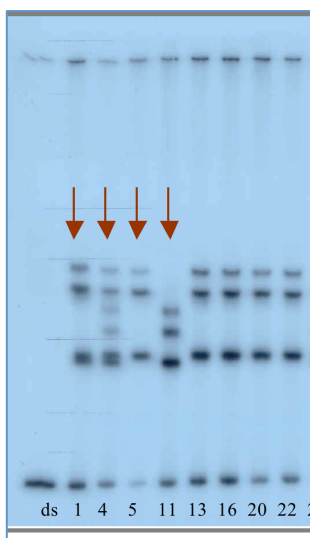


Figura 14. Autorradiografía de la SSCP del exón 2 del gen HNF4A, 8% de acrilamida con glicerol, 200 V/18 hs, exposición 24 h.

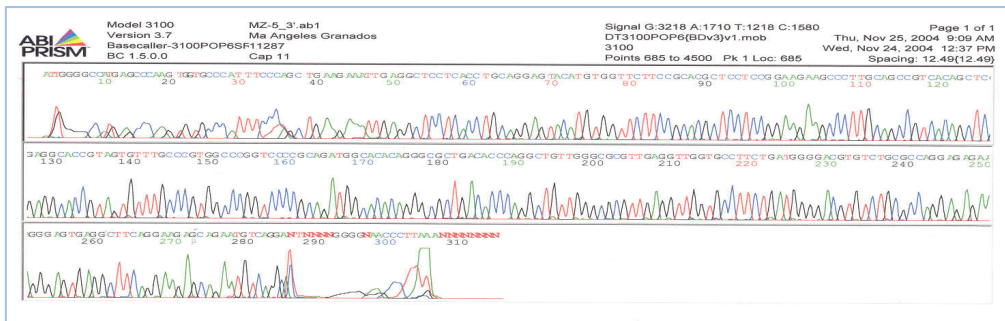


Figura 18. Electroferograma del exón 2 de HNF4, muestra MZ 5.

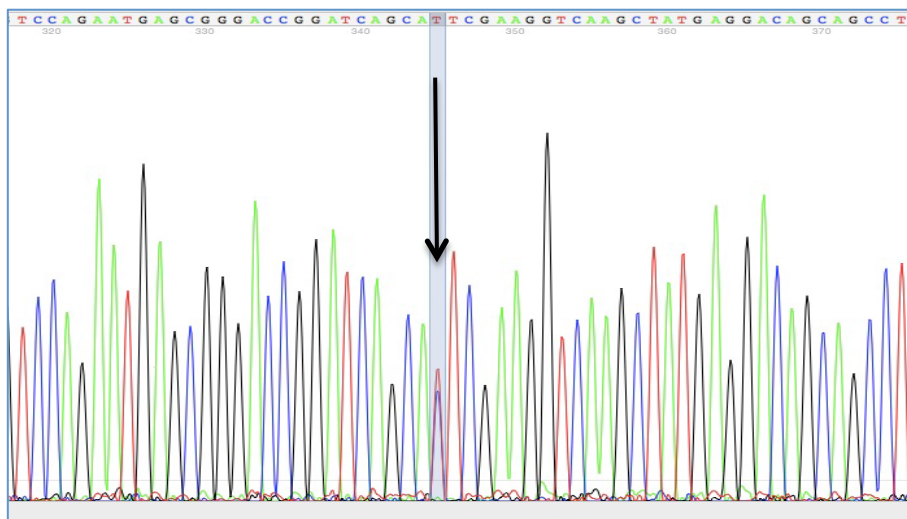


Figura 19. Electroferograma de la muestra MZ20 del exón 4 de *HNF4A*. La flecha indica la posición de la base cambiada en donde se observa la presencia de 2 nucleótidos. La curva roja indica la presencia de timina en un alelo, y la curva verde superpuesta indica la presencia de citosina en el otro alelo (C/T) originando un cambio en el codón 130 de la proteína de *HNF4A* del aminoácido treonina por isoleucina.

En los exones 1, 3, 5, 6, 7, 8, 9 y 10 todas las muestras presentaron un patrón de corrimiento electroforético homogéneo. De esta manera podemos resumir en la tabla 5 los hallazgos de las variantes exómicas encontradas en el gen *HNF4A* en población mazahua y teenek. El único cambio no sinónimo encontrado es el del polimorfismo T130I (rs1800961), el cual ya había sido reportado previamente por nuestro grupo de trabajo en población mestiza mexicana y asociado a la aparición de DT2 en personas menores de 40 años [53], y de la cual ya se reportó que disminuye en 46.2% la función de la proteína [68].

Tabla 5. Variantes de secuencia de *HNF4A* identificados en población Mazahua y Teenek.

Variante	localización	Cambio generado
rs736824	Intrón 2	Ninguno
rs745975	Intrón 2	Ninguno
rs1800961	Exón 4	T>I (T130/I)

GCK

Para cada exón de *GCK* se verificó que el tamaño del amplificado correspondiera a lo reportado en la literatura. Para ello se corrieron geles de agarosa al 1% durante 30 minutos a 80 V. Cada uno de los exones analizados tuvieron el tamaño esperado. Se realizaron 450 reacciones de PCR marcadas con P³² y 80 corrimientos electroforéticos (SSCP) de los 10 exones de *GCK* para identificar las muestras que tuvieran un cambio en su composición de bases en los grupos mazahua y teenek.

La figura 20 muestra un ejemplo del PCR preparativo del exón 1a de *GCK*, y donde se verifica que el exón tenga un tamaño de 237 pb.

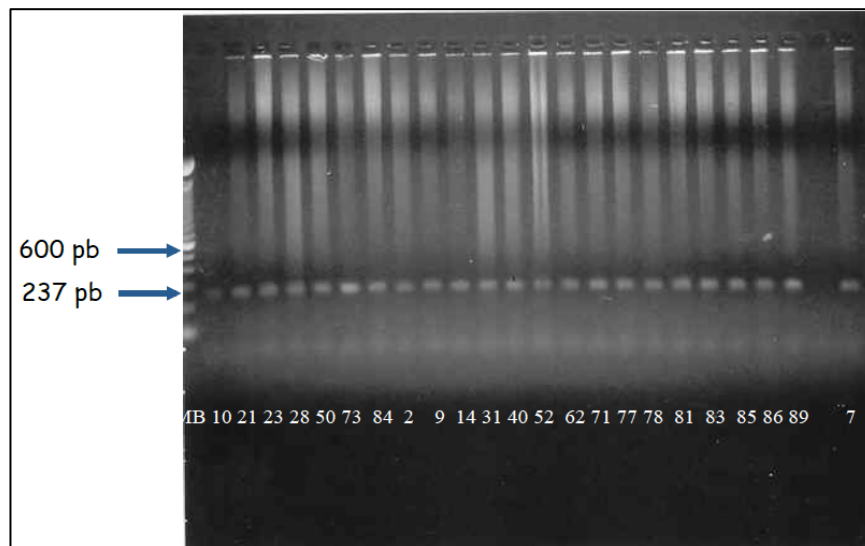


Figura 20. Corrimiento electroforético del PCR preparativo del exón 1a de *GCK*. Gel de agarosa 1%; 80 V/30 minutos; visualización con bromuro de etidio, tamaño del exón 237 pb.

En la figura 21 se observa un ejemplo de la autoradiografía obtenida y se pueden identificar las muestras en donde hay cambios en la secuencia en el exón 1a de *GCK* por la técnica de SSCP. Como puede observarse en la muestra 50 en la dirección de la flecha, se observan 3

bandas mientras que las otras muestras sólo presentan 2, cabe hacer notar que este patrón se conservó en las otras tres condiciones trabajadas (5.4% de acrilamida con glicerol, 8% de acrilamida con y sin glicerol). La muestra se amplificó para su secuenciación posterior.

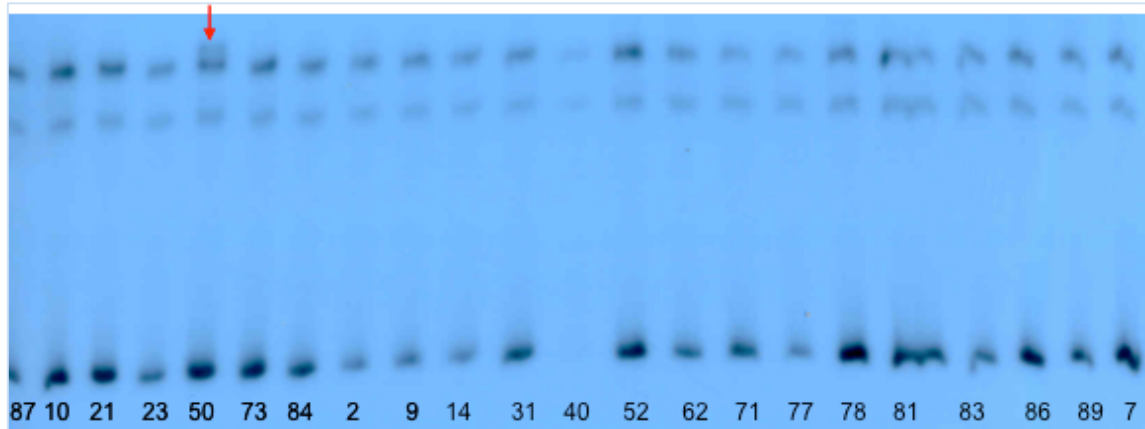


Figura 21. Autoradiografía de la SSCP del exón 1a de *GCK*, gel 5.4% de acrilamida sin glicerol, 180 V/24 h, exposición 24 hs.

La figura 22 muestra la secuenciación de la muestra 50, la diferencia en el corrimiento electroforético se debió al cambio de una C por una T en la tercera posición del codón 3, este cambio es sinónimo pues ambos tripletes codifican para el aminoácido aspártico.

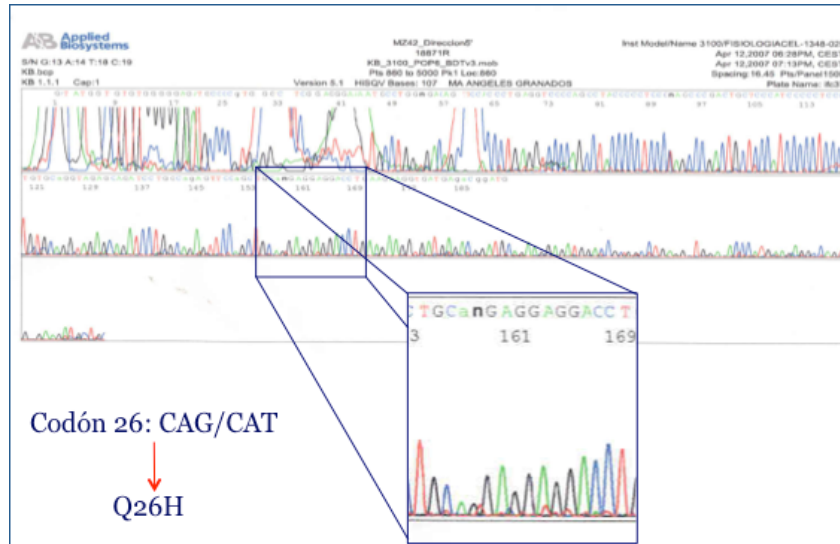


Figura 24. Electroferograma de la secuencia del exón 2 de *GSK* del individuo MZ42. En la posición 159 de la secuencia se encuentran las bases T y G en la misma posición ocasionando un cambio no sinónimo Q26H en la proteína de *GSK*.

Se realizó la predicción *in silico* del polimorfismo Q26H de *GSK* en el programa Polyphen V2.2 [105] y los resultados mostraron que éste no provocaría algún cambio deletéreo en la actividad de la enzima (Figura en anexos). Con el programa SIFT [106] esta variante muestra un puntuación de 0.02 lo que indica que puede modificar la función de la proteína, sin embargo el mismo programa indica que este resultado tiene un bajo nivel de confianza. Por su parte, la predicción por el programa FATHMM [107] indica que no hay cambios en la función.

En la figura 25 se muestra el corrimiento electroforético del exón 6 de *GSK*, revelado con tinción de plata, se busco esta alternativa por no contar ya con P³² debido a la restricción en la compra de este reactivo por su peligrosidad. Se observa que en la muestra 17 hay una banda más localizada entre la banda 3 y 4. Esta muestra corresponde a un individuo de la población teenek.

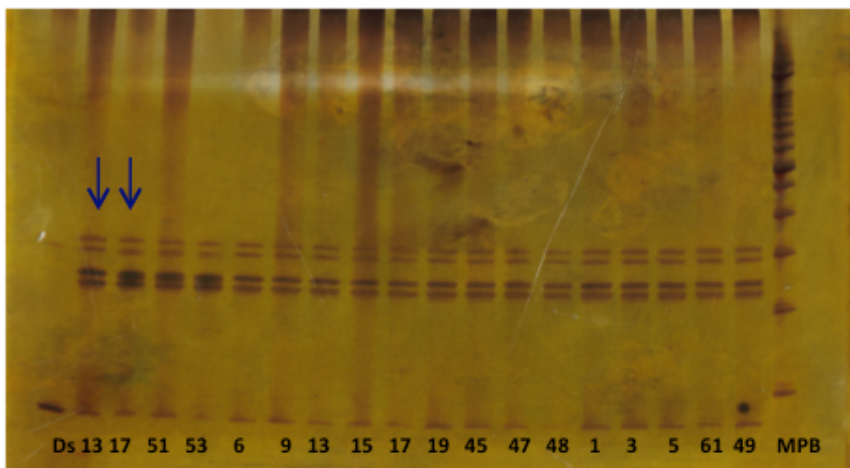


Figura 25. SSCP del exón 6 de *GSK*, gel 8% de acrilamida con glicerol, 180 V/24 h, la electroforesis fue revelada en este caso por tinción de plata.

La muestra TK 17 fue secuenciada y el resultado se muestra en la figura 26. Se observa que en la posición tres en el codón 215 se encuentran simultáneamente las bases T y C este cambio es sinónimo para el aminoácido tirosina (Y)

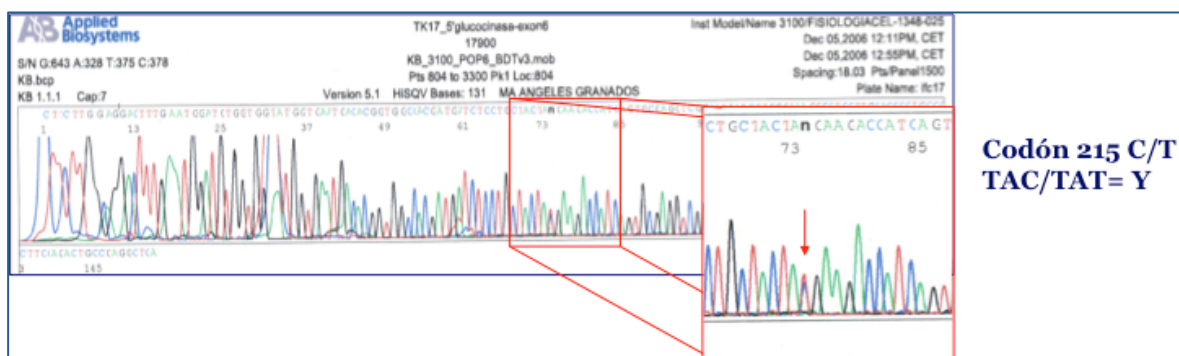


Figura 26. Electroferograma de la secuencia del exón 6 de *GCK* del individuo TK17. En la posición 74 de la secuencia se encuentran las bases T y C en la misma posición ocasionando un cambio sinónimo Y215Y en la proteína de *GCK*.

En los exones 3, 5, 7, 8, 9 y 10 de *GCK* todas las muestras presentaron un patrón de corrimiento electroforético homogéneo. En la tabla se resumen los hallazgos de las variantes exómicas encontradas en el gen *GCK* en población mazahua y teenek. Dos cambios sinónimos uno en la población mazahua D3D (rs144723656) y otro en la población teenek Y215Y (rs rs142553382). Un cambio no sinónimo en población mazahua Q26H (rs193922329), el cual no se reporta relacionado con alguna enfermedad y sólo se encontró en un individuo, la predicción por PolyPhen no indica algún cambio en la proteína.

Tabla 6. Variantes de secuencia de *GCK* en población mazahua y teenek.

Cambio de nucleótido	localización	Cambio generado
C/T Mazahua rs144723656	Exón 1A	D3D
G/T Mazahua rs193922329	Exón 2	Q26H
C/T Teenek rs142553382	Exón 6	Y215Y

Frecuencia de la mutación T130I del gen *HNF4A*

Debido a que nuestro grupo de trabajo ha encontrado que la variante T130I (rs1800961) del gen *HNF4A* es un factor de riesgo para padecer diabetes en la población mestiza mexicana (OR=3.38, CI 95% 1.14-10.03, $\chi^2=4.81$, p=0.028) y que en las poblaciones mazahua y teenek, también encontramos esta variante, se evaluó su frecuencia en otros nueve grupos indígenas: cora, chontal, chatino, maya, mixteco, otomí, purépecha, yaqui y zapoteco. La tabla 7 muestra las frecuencias encontradas en todos los grupos. Comparamos las frecuencias de los grupos indígenas con un grupo de mestizos de un estudio previo [53] encontramos que éstas son estadísticamente diferentes en los grupos cora, chontal y maya. Cuando se compara la frecuencia entre grupos indígenas, los grupos cora, chontal y maya son también diferentes del resto de los grupos.

Tabla 7. Frecuencias alélicas y genotípicas del rs1800961 (T130I) del gen *HNF4A* en población indígena mexicana.

<i>Etnia</i>	<i>T/T</i>	<i>T/I</i>	<i>I/I</i>	<i>Total</i>	<i>T</i>	<i>I</i>	<i>Total</i>	<i>130I (95% CI) frecuencia alélica</i>
Cora	89 (71.7%)	32 (25.8%)	3 (2.4%)	124	210 (84.6%)	38 (15.3%)	248	*p=0.00007 IC= (0.053-0.433)
Chatino	106 (94.6%)	6 (5.4%)		112	218 (97.3%)	6 (2.7%)	224	p=1.00001 IC=(0.279-3.6)
Chontal	98 (70.0%)	41 (29.3%)	2 (1.4%)	141	237 (84.0%)	45 (16.0%)	282	*p=0.00003 IC=(2.44-19.6)
Mixteco	120 (90.2%)	10 (7.5)	3 (2.3%)	133	250 (93.9%)	16 (6.0%)	266	p=0.125 IC=(0.766-7.12)
Maya	81 (74.3%)	24 (22.0%)	4 (3.7%)	109	186 (85.3%)	32 (14.7%)	218	*p=0.00014 IC= (2.17-18.15)
Mazahua	85 (93.4%)	6 (6.6%)		91	176 (96.7%)	6 (3.3%)	182	p=1.00002 IC= (0.345-4.5)
Teenek	61 (88.4%)	8 (11.6%)		69	130 (94.2%)	8 (5.8%)	138	p=0.19287 IC= (0.651-7.516)
Purépecha	23 (92%)	2 (8%)		25	48 (96%)	2 (4%)	50	p=0.74466 IC= (0.270-8.565)
Otomí	23 (92%)	2 (8%)		25	48 (96%)	2 (4%)	50	p=0.74466 IC= (0.270-8.565)
Zapoteca	22 (88%)	3 (12%)		25	47 (94%)	3 (6%)	50	p=0.43324 IC= (0.507-10.787)
Yaqui	25 (100%)	0 (0%)		25	50 (100%)	0 (0%)	50	p=0.8883 IC= (0.017-6.092)
Todos los grupos indígenas	733 (83.4%)	134 (15.2%)	12 (1.4%)	879	1600 (91%)	158 (8.9%)	1758	*p=0.00768 IC= (1.317-9.86)
Mestizo [53]	71 (94.6%)	4 (5.3%)		75	146 (97.3%)	4 (2.7%)	150	

Las frecuencias están en equilibrio Hardy-Weinberg, con excepción del grupo yaqui. La frecuencia del alelo 130I fue comparada con los sujetos mestizos y los grupos indígenas con la prueba de χ^2

Así mismo, la comparación entre grupos muestra que los cora, chontal y maya son los que presentan la frecuencia más alta y los que son significativamente diferente de las demás poblaciones, esto lo podemos observar en la tabla 8.

Tabla 8. Comparación de frecuencias alélicas del polimorfismo T301 (rs1800961) entre grupos indígenas.

	Cora	Chatino	Chontal	Mixteco	Maya	Mazahua	Teenek	Purépecha	Otomí	Zapoteca	Yaqui
Cora		< 0.0001	0.9048	0.0008	0.8970	< 0.0001	0.0052	0.0383	0.0383	0.1128	0.0008
Chatino			< 0.0001	0.0834	< 0.0001	0.7737	0.1635	0.6407	0.6407	0.2141	0.5961
Chontal				0.0002	0.7099	< 0.0001	0.0028	0.0261	0.0261	0.0798	0.0005
Mixteco					0.0020	0.2658	1.0000	0.7485	0.7485	1.0000	0.0855
Maya						0.0001	0.0097	0.0563	0.0563	0.1596	0.0012
Mazahua							0.2869	0.2869	0.6831	0.4095	0.3450
Teenek								1.0000	1.0000	1.0000	0.1119
Purépecha									1.0000	1.0000	0.4949
Otomí										1.0000	0.4949
Zapoteca											0.2424
Yaqui											

P<0.05 se considera significativo. Prueba exacta de Fischer

Realizamos la comparación de estas frecuencias con las reportadas en el Proyecto Internacional HapMap3, tomando el promedio de la frecuencia en toda la población indígena mexicana evaluada para llevar a cabo este análisis (Tabla 9).

Se observa que la frecuencia de la variante rs1800961 es diferente entre los grupos del HapMap3 y los indígenas mexicanos evaluados (P<0.05), y aún más que en los indígenas yoruba de Nigeria esta variante no está presente en su población.

Tabla 9. Comparación de las frecuencias alélicas y genotípicas de la variante T130I (rs1800961) entre los indígenas mexicanos (MEXAME): cora, chontal, chatino, maya, mazahua, teenek, mixteco, purépecha, otomí, zapoteco y yaqui, con las poblaciones del Proyecto Internacional HapMap3 [109].

Población	Frecuencia genotípica		Frecuencia alélica					
	n	CC	n	CT/TT	n	C	n	T
ASW (A)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
CEU (C)	113	0.929	8	0.071*	226	0.965	8	0.035*
CHB (H)	137	0.964	5	0.036*	274	0.982	5	0.018*
CHD (D)	109	0.954	5	0.046*	218	0.977	5	0.023*
GIH (G)	101	0.950	5	0.050*	202	0.975	5	0.025*
JPT (J)	113	0.991	1	0.009*	226	0.996	1	0.004*
LWK (L)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
MEX (M)	58	0.983	1	0.017*	116	0.991	1	0.009*
MKK (K)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
TSI (T)	102	0.971	3	0.029*	204	0.985	3	0.015*
YRI (Y)	63	1.000	0	n/a	126	1.000	0	0*
MEXAME (I)	733	0.834	146	0.166	1600	0.91	158	0.089

*frecuencia comparada con MEXAME $p < 0.005$. Poblaciones: ASW (A): de ascendencia Africana residente en Suroeste de EU, CEU (C): Residentes de Utah de ascendencia del norte y oeste de Europa de la colección CEPH, CHB (H): Chinos Han de Beijing, China, CHD (D): Chinos de Denver, Colorado, GIH (G): Indios Gujarati de Houston, Texas, JPT (J): Japoneses de Tokyo, Japón, LWK (L): Indígenas Luhya de Webuye, Kenya, MEX (M): Ancestría Mexicana en Los Angeles, California, MKK (K): Indígenas Maasai de Kinyawa, Kenya, TSI (T): Población Toscana de Italia, YRI (Y): Indígenas Yoruba de Ibadan, Nigeria, MEXAME (I): cora, chontal, chatin, maya, mazahua, teenek, mixteco, purépecha, otomí, zapoteco y yaqui.

Análisis de la Extensión de la Homocigidad Haplotípica (EHH) para la variante T130I (rs 1800961) del gen *HNF4A*.

En la publicación de los resultados anteriores, nosotros sugerimos la hipótesis de que la variante T130I (rs1800961) podría ser originada en el proceso evolutivo de los pueblos mexicanos, por lo que para reforzar esta teoría nos sugirieron realizar un análisis estadístico que apoyara esta aseveración. Para ello elegimos el análisis de EHH [104]. El EHH es una función de la edad de la mutación, las mutaciones más jóvenes muestran homocigidad en regiones más grandes que las más antiguos. El EHH en una región se puede calcular para un haplotipo central o un alelo SNP central y luego se compara con el de otros haplotipos centrales o el alelo alternativo [110]

La figura 27 muestra el decaimiento de EHH alrededor de la variante T130I (rs1800961) del gen *HNF4A*.

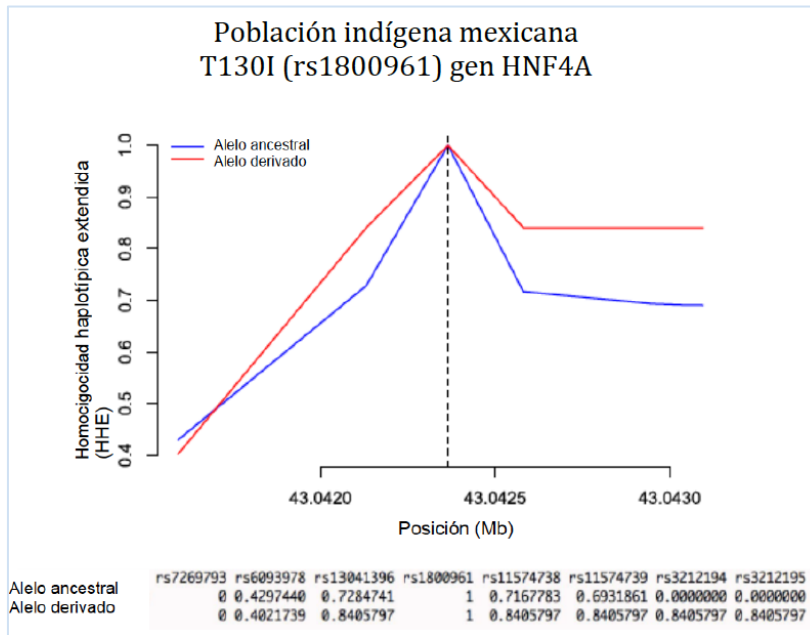
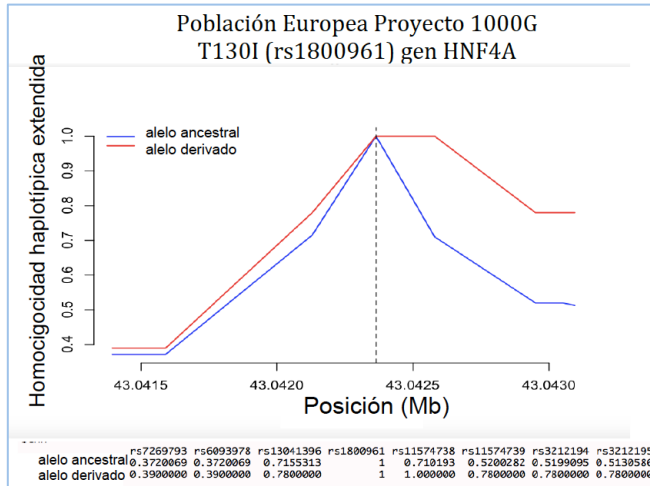


Figura 27. Extensión de la homocigidad haplotípica para la variante T130I (rs1800961) del gen HNF4A en población indígena mexicana

Como puede observarse, del lado derecho de la gráfica, el decaimiento del alelo ancestral junto con los rs3212194 y rs3212195 disminuyeron al 100% de homocigidad, mientras que los mismos polimorfismos, pero con el alelo derivado rs1800961 sólo decayeron un 16% en ese segmento. Río abajo de la variante, el decaimiento con el rs13041396 fue de 27.1% para el alelo ancestral y 16% para el alelo derivado. De acuerdo con la interpretación de Sabeti [110], la presencia de la variante rs1800961, encontrada en alta frecuencia en población indígena mexicana está asociada a una EHH alta, identificándola entonces como una posible señal de selección positiva reciente.

Como comparación y para evaluar como ha sido el comportamiento de la variante T130I (rs1800961) de *HNF4A* en las poblaciones donde ha sido reportada, realizamos un análisis EHH con los datos del proyecto 1000G para la población europea y asiática. La figura 28 muestra la gráfica obtenida. Como puede observarse, la población Europea tiene el mismo patrón que la población indígena mexicana. Sin embargo, en la población asiática los haplotipos formados con el alelo derivado tienen un decaimiento mayor que los formados con el alelo ancestral, lo que sugiere que esta variante tiene más tiempo en esta población.

A)



B)

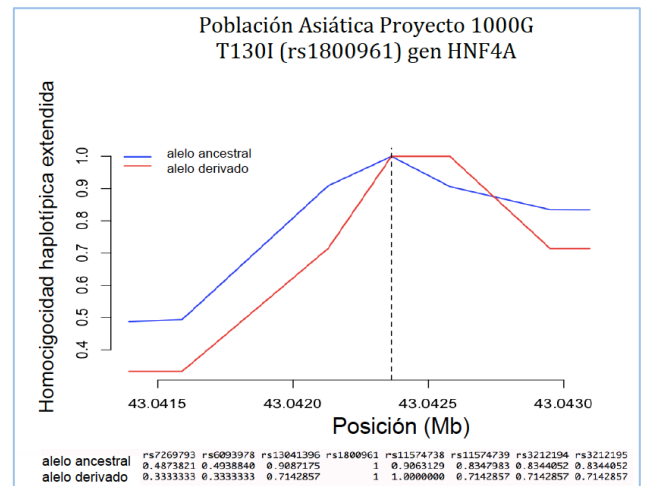


Figura 28. Extensión de la homocigidad haplotípica para la variante T130I (rs1800961) del gen HNF4A en población europea (A) y asiática (B)

Asociación de la variante T130I (rs1800961) del gen HNF4A con rasgos metabólicos.

La asociación de la variante T130I (rs1800961) con factores de riesgo para desarrollar DT2, se llevó a cabo sólo en 7 grupos indígenas, con los que se contaba con muestra sanguínea y datos antropométricos: cora, chatin, chontal, maya, mazahua, mixteco y teenek.

Primero se realizó la caracterización de las poblaciones y los resultados los podemos observar en la tabla 10, en este caso también se llevó a cabo una comparación con el grupo de mestizos de la CdMX [53]. Como puede observarse en el caso de los grupos maya y mixteco el valor de la mediana de la glucosa sanguínea es superior a los valores de referencia (VR) aceptados por la ADA para ser considerados como individuos normoglicémicos (glucosa < 99 mg/dl) [4]. A este respecto en la población evaluada encontramos desde un 4% en los Mazahuas hasta un 26% en los Mixtecos de personas con DT2, el diagnóstico de diabetes se estableció de acuerdo a los criterios de la ADA, glucosa en ayuno > 126 mg/dl o si ellos referían tomar medicamentos hipoglucemiantes [4].

Referente a los valores de colesterol y con excepción del grupo chontal este parámetro se encontró en promedio dentro de lo establecido por el ATP III (colesterol total deseable < 200 mg/dl) [111], sin embargo los triglicéridos tuvieron una gran variabilidad entre los grupos y sólo los cora y mazahua se puede decir que tienen un valor recomendable < 150 mg/dl. Las determinaciones de insulina se reportan en 4 de los 7 grupos estudiados por

aspectos de control de calidad en esta determinación (tiempo de almacenamiento de las muestras). En este parámetro es de hacer notar que el grupo teenek, presenta en promedio VR dentro de los límites establecidos, sin embargo de manera individual, hay personas que exceden en el 25% el límite de referencia superior.

Los valores de IMC indican un estado de sobrepeso y obesidad en todos los grupos evaluados con excepción de los grupos teenek y mixteco, observamos que los grupos cora, maya y mazahua presentan un ICC > 0.94 en los hombres, mientras que en el caso de las mujeres sólo el grupo chontal se encuentra dentro de los VR [112].

Tabla 10. Características bioquímicas y antropométricas de 7 grupos indígenas mexicanos.

Parámetro	Chatin N=114	Chontal N=141	Cora n=124	Maya n=110	Mazahua n=91	Mixteco N=134	Teenek n=68	Mestizo n=75
Edad (años)	47±17	37±13	46±18	43±15	40±12	45±14	43±12	58±10
IMC (kg/m ²) VR 18.5-24.9	25.5±4.5	28.0±5.0	26.0±5.0	*29.7±5.0	27.6±4.0	22.5±4.0	*23.0±4.0	26.0±4.0
ICC (hombre) VR 0.78-0.94	0.80±0.08	0.90±0.10	0.93±0.06	0.93±0.06	0.93±0.06	0.90±0.10	0.90±0.04	0.92±0.06
	(n=25)	(n=16)	(n=36)	(n=37)	(n=4)	(n=15)	(n=21)	(n=20)
ICC (mujer) VR 0.71-0.85	0.90±0.10	0.80±0.10	0.87±0.05	0.85±0.07	0.86±0.05	0.90±0.10	0.87±0.06	0.87±0.06
	(n=89)	(n=125)	(n=88)	(n=73)	(n=87)	(n=119)	(n=47)	(n=55)
Glucosa (mg/dl) VR 55-99 mg/dl	^b 96 (89-109)	⁹⁴ (87-103)	^{c,d} 94 (87-102)	^a 101 (94-117)	^{c,d} 92 (87-102)	^a 105 (94-133)	^c 96 (90-105)	95±14
Colesterol (mg/dl) VR <200 mg/dl	195±37	202±38	^{e,f} 174±35	199±38	^{e,f} 165±47	^{e,f} 177±33	^{e,f} 168±37	202±38
Triglicéridos (mg/dl) VR <150 mg/dl	^h 217 (148-333)	176 (112-249)	^{ij} 144 (101-198)	^{jk} 173 (129-224)	^{h,i,j} 124 (93-187)	^h 248 (165-316)	^k 180 (110-254)	157±52
Insulina (pmol/L) VR 17.8-173 pmol/L			*78±66	*72±66	*30±48		*114±102	39±24

Los datos para glucosa y triglicéridos no tuvieron una distribución normal por lo cual están representados como mediana y los rangos inter-cuartílicos (25%-75%) y en el grupo mestizo como media y ds.

IMC= Índice de masa corporal; ICC= Índice cintura/cadera (VR tomados de Comité de Expertos de la OMS sobre la obesidad [112])

*p< 0.05 vs Mestizo

a p< 0.05 Chontal vs Mixteco y Maya

b p<0.05 chatino vs Mixteco

c p<0.05 Mixteco vs Cora, Mazahua y Teenek

d p<0.05 Maya vs Cora y Mazahua

e p<0.05 Chontal vs Mixteco, Cora, Mazahua y Teenek

f p< 0.05 Chatino vs Mixteco, Cora, Mazahua, Teenek

g p< 0.05 Maya vs Mixteco, Cora, Mazahua y Teenek

h p< 0.05 Chontal vs Chatino, Mixteco, Mazahua; i p<0.05 Chatin vs Cora y Mazahua; j

p<0.05 Mixteco vs Cora Mazahua, Maya; k p<0.05 Mazahua vs Teenek y Maya

La asociación de la variante T130I (rs1800961) con rasgos metabólicos se llevo a cabo por un análisis univariado ajustado por sexo, edad e IMC bajo el modelo dominante por las frecuencias encontradas de homocigotos para la variante. Con la finalidad de establecer una asociación con factores de riesgo para DT2 (ICC, glucosa, triglicéridos y colesterol) sólo se evaluaron individuos con concentraciones de glucosa < 126 mg/dl. Los resultados se muestran en la tabla 11.

Tabla 11. Asociación de la variante T130I (rs1800961), con factores de riesgo para DT2 en población indígena mexicana.

Parámetros	T130T N=558	T130I N=117
ICC (hombres)	0.92±0.06	0.92±0.06
ICC (mujeres)	0.86±0.06	0.85±0.06
Glucosa (mg/dl)	95±12	95±11
Triglicéridos (mg/dl)	195±107	220±156*
Colesterol Total (mg/dl)	182±40	185±36

La relación entre genotipo-factores de riesgo fue evaluada con un análisis univariado ajustado por sexo, edad e IMC bajo el modelo dominante. Los datos se muestran como media ± SD. Un valor de *P < 0.05 fue considerado significativo.

Como puede observarse entre los individuos que presentan la variante T130I (rs1800961) hay una asociación con concentraciones altas de triglicéridos (P<0.05).

Considerando que la frecuencia de la variante T130I (rs1800961) fue significativamente diferente entre los cora, chontal y maya vs el resto de las poblaciones, se realizaron las asociaciones con estos mismos factores de riesgo para cada uno de los grupos tomando en cuenta los individuos con concentraciones de glucosa <126 mg/dl. Los resultados se muestran en la tabla 12. Se observa que la asociación entre los valores altos de triglicéridos se conserva en las poblaciones cora, chontal y chatino, pero se pierde en el resto. En los demás parámetros no hay diferencias significativas.

Tabla 12. Asociación de la variante T130I (rs1800961) del gen *HNF4A*, con factores de riesgo para DT2 por grupo indígena.

rs1800961 T130I	Genotipo	Glucosa (mg/dL)	Colesterol (mg/dl/dL)	Triglicéridos (mg/dL)
Cora	CC (n=86)	92.3 ± 9.5	172.2 ± 35.7	151.5 ± 70.2
	CT/TT (n=32)	95.6 ± 10.9	180.3 ± 34.0	209.0 ± 159.4
	P	0.126	0.298	0.016
Chontal	CC (n=89)	91.8 ± 14.3	204.3 ± 37.7	178.2 ± 100.5
	CT/TT (n=38)	94.3 ± 10.7	191.6 ± 32.8	222.7 ± 140.8
	P	0.214	0.105	0.027
Chatino	CC (n=91)	94.2 ± 11.0	192.0 ± 36.3	230.7 ± 126.7
	CT/TT (n=4)	91.0 ± 8.9	199.0 ± 16.9	513.0 ± 398.9
	P	0.612	0.530	0.000
Maya	CC (n=73)	100.7 ± 9.5	197.1 ± 35.9	183.3 ± 87.0
	CT/TT (n=22)	97.4 ± 9.8	194.3 ± 38.4	191.1 ± 107.7
	P	0.146	0.289	0.607
Mazahua	CC (n=71)	92.2 ± 11.8	165.7 ± 49.9	139.7 ± 75.2
	CT/TT (n=4)	83.0 ± 16.8	157.3 ± 32.0	129.3 ± 28.3
	P	0.153	0.395	0.991
Mixteco	CC (n=93)	99.8 ± 11.3	175.6 ± 31.1	235.3 ± 113.6
	CT/TT (n=8)	92.4 ± 16.7	161.0 ± 33.8	182.8 ± 64.3
	P	0.353	0.323	0.234
Teenek	CC (n=56)	94.7 ± 10.9	168.6 ± 38.7	250.6 ± 321.2
	CT/TT (n=8)	98.6 ± 6.5	169.2 ± 32.0	187.4 ± 102.9
	P	0.346	0.783	0.677

La relación entre genotipo-factores de riesgo fue evaluada con un análisis univariado ajustado por sexo, edad e IMC bajo el modelo dominante. Los datos se muestran como media ± SD. Un valor de *P < 0.05 fue considerado significativo.

VII. Discusión

En México la DT2 ha alcanzado una prevalencia muy alta (15.9%) de acuerdo a lo reportado por la OCDE en 2015 [12], sin embargo, la ENSANUT 2016 de medio camino reporta 9.4%, debido a esta inexactitud en los datos y en las estimaciones de la IDF (México 5º lugar en 2017; México 4º lugar en 2015) (4), se puede evidenciar que podemos estar subestimando el total de la población afectada, y que los gastos que serán erogados por el sector salud en el futuro serán mayores a los calculados, además que la calidad de vida de los individuos también será afectada. Por lo tanto, todos los esfuerzos para dilucidar el origen de la DT2 en población mexicana que se realicen en el ámbito clínico y en el campo de la investigación, serán de suma importancia tanto para la prevención como para el tratamiento de la DT2. Así, la alta prevalencia de esta enfermedad ha llevado a diversos investigadores a evaluar la genética de la DT2 en México, con la finalidad de identificar a los genes involucrados en la aparición de esta patología en poblaciones indígenas y mestizas mexicanas [93, 103, 113].

Resultados de los diversos abordajes utilizados en la investigación genética para asociar variantes específicas con el desarrollo de DT2, sugieren que ciertos grupos étnicos en el mundo poseen un fondo genético que los hacen más susceptibles de padecer diabetes [114, 115].

Ancestría

En particular, se conoce que un componente amerindio aumenta el riesgo de desarrollar enfermedades crónicas como DT2 [116], así el estudio de los polimorfismos en los indígenas mexicanos podría ser un enfoque práctico para desvelar la herencia genética de la diabetes en la población mexicana. Además, la identificación de los factores de riesgo para la salud en estas comunidades es muy importante porque se consideran poblaciones vulnerables debido a su condición socioeconómica. En este contexto, todos los participantes de las comunidades evaluadas se reconocieron como indígenas, cuyos padres y abuelos nacieron en la misma comunidad, que hablan el idioma y conservan sus tradiciones, pero además la ascendencia amerindia fue evaluada a través de AIMS. A este respecto nosotros consideramos que la mayor parte de estratificación poblacional en México es debida al mestizaje que se llevo a cabo a partir de la conquista y que mayoritariamente fue con

individuos europeos. Se ha reportado que se obtienen elevadas correlaciones con 10 AIMs cuando se toman en cuenta sólo dos poblaciones ancestrales para las estimaciones de ancestría [117]. En este trabajo nosotros sólo evaluamos la participación europea y nativoamericana. Sin embargo no se descarta que para tener resultados con más confiabilidad deban de ser utilizados más de 10 AIMs [118].

Aunque, actualmente, no existe población con un linaje puro, sorprendentemente, nuestros resultados (tabla 4) mostraron que el componente amerindio fue del 97.6% en estas comunidades, la contribución amerindia más baja se observó en el grupo otomí (96.4%), y la más alta se observó en el grupo Zapoteco (98.5%). Todos estos hallazgos sugieren una conservación del componente amerindio en estas poblaciones debido a las circunstancias de aislamiento geográfico. La estimación de la ancestría en este trabajo avala la representatividad del componente amerindio en la población indígena estudiada. Como un comparativo, también evaluamos un grupo de mestizos de la Ciudad de México, encontrando una estimación de componente amerindio de 76.8%, lo que concuerda con datos reportados por Lisker y cols. [90].

Variantes encontradas de *HNF4A*

Genes clave en la homeostasis de la glucosa han sido elegidos de primera intención para ser estudiados como factores de riesgo en el desarrollo de DT2. Tal es el caso de *HNF4A*, el cual participa en gran medida en la regulación gluconeogénica hepática y en el metabolismo de los lípidos. En el páncreas, *HNF4A* también es un regulador maestro, que controla aproximadamente el 11% de los genes de los islotes [61].

En una primera etapa, en este trabajo, evaluamos toda la región exómica del gen *HNF4A* en población mazahua y teenek, encontrando únicamente un cambio no sinónimo, la variante T130I (rs1800961) con una frecuencia 6.5 y 11.0% respectivamente, esta variante fue por primera vez reportada por Yamagata y cols., en individuos no hispanos [67]. Adicionalmente encontramos dos variantes intrónicas rs736824 y rs745975 (Tabla 5).

Las dos variantes intrónicas (rs736824 y rs745975) fueron encontradas al inicio del exón 2. En este contexto se conoce por estudios de asociación en otras poblaciones, que la susceptibilidad de padecer ciertas enfermedades complejas se encuentra en los polimorfismos de las zonas intrónicas [119]. Recientemente Marcil y cols. (2015) encontraron los rs736824 y rs745975 como un haplotipo junto con el rs3212183 en

jóvenes Franco-Canadienses con síndrome metabólico. La presencia del rs33212183 no fue detectada porque esa zona no estaba incluida en el fragmento amplificado [120] sin embargo sería interesante buscar la presencia del rs33212183 y de explorar su frecuencia junto con los rs736824 y rs745975 además de conocer si tienen algún impacto en el metabolismo de la población mexicana.

Frecuencia de la variante T130I (rs1800961) del gen *HNF4A*

La variante rs1800961 se traduce en un cambio del aminoácido treonina por una isoleucina (T130I) en la posición 130 de la proteína, localizada entre el dominio de unión a DNA y el dominio de unión a ligando, esta región esta implicada en la dimerización y unión a DNA. El residuo de treonina permite la interacción de dos sitios de fosforilación por PKC, mientras que la sustitución por isoleucina sólo facilita un sitio, afectando la activación de HNF4A [69]. Pruebas funcionales de este polimorfismo, revelan que disminuye 46.2% la actividad de transactivación de la proteína en células Hep-G2, y en cultivos primarios de hepatocitos 27.9% [68].

Estudios de nuestro grupo de trabajo, encuentran que en población mestiza mexicana, esta variante tiene asociación con la enfermedad en diabéticos de aparición temprana (<40 años) [53]. Por lo que debido a la importancia que tiene esta variante en su asociación con DT2 y a la frecuencia con la que se presentó en los grupos mazahua 6.6% y teenek 11.6%, nosotros extendimos la búsqueda de su frecuencia a otros grupos indígenas: cora, chatino, chontal, maya, mixteco, otomí, purépecha, yaqui y zapoteco, encontrando en ellos, con excepción del grupo Yaqui, una frecuencia mayor del 5%.

Es importante mencionar que aunque se han reportado 72 mutaciones de este gen en diversas poblaciones [64], en la población indígena mexicana en este estudio sólo encontramos la variante T130I (rs1800961). Su asociación con la DT2 es controversial, ya que algunos investigadores a pesar de haber evaluado un gran número de individuos en otras poblaciones, encuentran que la frecuencia en personas sin la enfermedad y las afectadas son muy similares, por lo cual no la asocian con la DT2 [67]. Sin embargo, en la población mexicana la diferencia en cuanto a la frecuencia encontrada en individuos diabéticos de aparición temprana es estadísticamente significativa [53]. Por otra parte, también nuestro grupo reportó que en un estudio de asociación, a pesar de haber encontrado una frecuencia alélica del 10.3% en población maya de Yucatán no encontramos

asociación de la variante T130I (rs1800961) con DT2, lo que nos indica el impacto diferencial de esta variante en los diferentes grupos poblacionales [93].

Los resultados encontrados en este trabajo acerca de las frecuencias de la variante T130I (rs1800961) (tabla 7) que están en un amplio intervalo (0-29.3% frecuencias alélicas) en población indígena mexicana sugieren que este polimorfismo es heredado de los primeros pobladores de México. Cabe mencionar que las frecuencias genotípicas observadas en la población indígena mexicana (16.6%) son más altas y significativamente diferentes ($P < 0.005$) (tabla 8) que las encontradas en las poblaciones del proyecto HapMap fase III, y donde además se observa que las poblaciones africanas no presentan esta variante, lo que implica que muy probablemente fuera portada por los primeros pobladores de América (esta variante se encuentra en la población euro-asiática 4%) y que se concentrará en las siguientes generaciones través del tiempo. Además se observa una diferencia significativa entre las frecuencias de los grupos cora, chontal y maya con las otras poblaciones evaluadas (tabla 9). Moreno y cols., ya ha hecho mención a las diferencias genéticas inter-étnicas de los grupos mexicanos, y que estas diferencias pueden tener implicación en aspectos relacionados con la salud [121].

La pregunta que nos planteamos acerca de esta variante ¿es porque este cambio se conservo en las poblaciones amerindias?. Se conoce por ejemplo, que la selección natural provoca mutaciones y polimorfismos en los genes clave y sus productos, modificando vías metabólicas para aprovechar eficientemente los nutrientes disponibles, como en el caso del polimorfismo (C/T-13910) en el gen que codifica la lactasa (*LCT*), el cual se asocia con la capacidad de digerir la leche en adultos europeos (persistencia de lactasa) [122]. Además, se ha sugerido que los procesos de adaptación han forzado la aparición de variantes genéticas que permitieron la resistencia de grandes hambrunas durante viajes largos. Sin embargo, bajo las condiciones actuales, fuera del contexto cazador-recolector transmitidos por antepasados amerindios, las características genéticas desempeñan un papel diferente; por lo tanto, esos beneficios inducidos por un rasgo adaptativo ahora se convierten en patológico bajo condiciones modernas. En este sentido, el consumo de carbohidratos necesario para soportar largas caminatas en busca de comida, ahora favorece la aparición de enfermedades tales como diabetes, obesidad e hipertensión debido al estilo de vida actual [123].

En condiciones fisiológicas, el gen *HNF4A* es clave en la expresión de genes específicos del hígado como: apolipoproteínas, factores séricos, isoformas del citocromo P-450 y las proteínas implicadas en el metabolismo de la glucosa, ácidos grasos y aminoácidos; sin embargo, la caracterización de la diabetes tipo MODY1 reveló su importancia para la secreción de insulina estimulada por glucosa de la β célula pancreática. Por lo tanto, es posible formular la hipótesis de que la variante rs1800961 (T130I) en *HNF4A* podría promover la baja utilización glucosa a nivel celular en un entorno pobre de carbohidratos en la época nómada de los primeros pobladores de México.

Para probar parte de esta hipótesis, realizamos un análisis de HHE (figura 27), observando un decaimiento mayor de la homocigocidad para el alelo ancestral, lo cual nos sugiere que el alelo rs1800961 ha sido sujeto a una selección positiva reciente [110]. Sin embargo esta misma variante tiene un mayor decaimiento en la población asiática lo que sugiere que esta variante es más antigua en esa población (figura28). Cabe mencionar que esta es una primera aproximación en la búsqueda del origen y de la antigüedad que tiene la variante en estudio, y que análisis más precisos y dirigidos deben llevarse a cabo para tener resultados más contundentes en este sentido.

Estudios de Jafar-Mohammadi y Hellwege en el 2011 [124, 125], coinciden en que la variante T130I (rs1800961), se encuentra asociada con un modesto riesgo a la aparición de DT2, pero llama fuertemente la atención que su presencia es una constante en los estudios genéticos realizados con *HNF4A*, lo que nos sugiere que ésta es una variante común de una enfermedad común. Los estudios de asociación de todo el genoma han identificado más de 80 variantes comunes para la DT2, de efecto moderado (el riesgo de DT2 aumenta en un 5-40%). Casi todas estas variantes regulan la secreción de insulina y solo unas pocas regulan la sensibilidad a la insulina y que en contraste la variantes raras y de baja frecuencia con grandes efectos, tienen una contribución limitada al riesgo de DT2 en la población general [126], así entonces podemos darle importancia al hallazgo de la frecuencia encontrada de la variante T130I (rs 1800961) en los pueblos indígenas mexicanos.

Respecto a los estudios previos de la variante T130I (rs1800961) en población mexicana, en el 2008 nuestro grupo de trabajo [53] reportó la asociación de ésta con DT2 de aparición temprana en pacientes mestizos mexicanos. La alta frecuencia del alelo de riesgo (8.9%) (tabla 7) entre las poblaciones indígenas mexicanas estudiadas sugiere la existencia de un

fondo diabetogénico en las poblaciones fundadoras mexicanas y que la rs1800961 podría ser parte de él.

Asociación de la variante T130I (rs1800961) con factores de riesgo metabólicos para desarrollar DT2

Para conocer el impacto de la variante T130I (rs1800961) en la susceptibilidad a desarrollar DT2 en México, su evaluación debe realizarse teniendo en cuenta las características genéticas de nuestra población, tomando en cuenta la aportación de todos aquellos genes que de alguna manera han sido asociados a factores de riesgo para el desarrollo de DT2 en nuestro país, de esta manera buscar que papel juega T130I (rs1800961) como factor de riesgo en la DT2. En este contexto, se conoce que la variante R230C del gen ABCA1 se asoció con una baja concentración de colesterol HDL, mostrando evidencia de selección positiva en nativos americanos [127]. Además, Villarreal-Molina y cols. encontró una asociación de la variante R230C de ABCA1 con la DT2 de inicio temprano en la población mexicana [103]. Por otro lado, se encontró asociación de polimorfismos en los genes HMOX y NQ01 con componentes del síndrome metabólico en las poblaciones indígenas de México [128]. Por su parte, Lara-Riegos y cols., reportaron la asociación de las variantes rs10811661(CDKN2A/2B) y rs9282541 (ABCA1) con DT2 en población maya[93]. La presencia de variantes como T130I en *HNF4A*, R230C en *ABCA1*, variantes en los genes HMOX y NQ01, así como rs10811661, entre otros polimorfismos con efectos metabólicos menores, sugiere fuertemente una contribución genética aditiva en la aparición de enfermedades metabólicas como la DT2, de ahí la importancia de la evaluación genética en poblaciones como las indígenas mexicanas, que no deben ser excluidas y formar parte de los beneficios que puedan traer estos tipos de estudios en la salud de los mexicanos.

La evaluación del estatus metabólico de las poblaciones en estudio fue muy importante y permitió encontrar parámetros que correlacionaron con la presencia de la variante T130I (rs1800961) (tablas 10, 11 y 12). La ENSANUT 2016 ha publicado datos de algunos parámetros como colesterol pero es necesario realizar una evaluación bioquímica más profunda en estas encuestas, porque como observamos en la tabla 10, la mayoría de la población presenta glucosa alterada en ayuno (valores mayores a 99 mg/dl pero menores a 126 mg/dl), y la hiperlipidemia combinada al parecer es muy frecuente en estos grupos indígenas, y en general en la población mexicana [96].

En este estudio, cuando evaluamos el impacto de la variante T130I (rs1800961) con los factores de riesgo para desarrollar DT2 en toda la población indígena con la que contábamos con datos antropométricos y bioquímicos (tabla 11), observamos su asociación con concentraciones altas de triglicéridos ($P= 0.048$), sin embargo cuando se analizan por grupos individuales (tabla 12), esta asociación la encontramos únicamente en los grupos cora, chontal y chatino y al igual que en los resultados de Lara-Riegos, en los mayas no hay asociación con las concentraciones de triglicéridos, lo cual nos habla de impacto diferencial por etnias de esta variante y que algunos otros factores deben estar involucrados. Sin embargo, de manera global podemos especular que la variante T130I (rs1800961) puede ser un factor de riesgo para padecer diabetes, ya que al aumentar los ácidos grasos libres en sangre, aumenta la resistencia a la insulina y también la disfunción de la célula β [129, 130].

HNF4A es un factor de transcripción que influye en el metabolismo de los triglicéridos en plasma a través de un mecanismo aún desconocido. Debido a la importancia de HNF4A en el metabolismo, sería muy interesante evaluar a través de que mecanismo la variante T130I (rs1800961) pueda modular el incremento de triglicéridos que observamos en la población indígena. En este sentido, se han identificado un gran número de genes diana para HNF4A, incluidos los que codifican para apolipoproteínas, factores de coagulación sanguínea y enzimas implicadas en el metabolismo de lípidos, aminoácidos y glucosa [131, 132]. Recientemente se ha informado la participación de HNF4A en la regulación de la diacilglicerol aciltransferasa una enzima que participa en la secreción de lipoproteínas ricas en triglicéridos [133]. Además, en un estudio GWA realizado en población mexicana se reporta la influencia del rs17259126 sobre la actividad de *HNF4A* en uno de sus genes blanco, la proteína transmembrana 241 (TMEM241) ligada a su vez con el incremento de los triglicéridos en los mexicanos [134].

Por otra parte, también los IMC y ICC de la mayoría de los grupos estudiados nos indicarían una tendencia al sobrepeso y a la acumulación de grasa visceral factores independientes que sin embargo siguen siendo un riesgo para el desarrollo de DT2. Con relación a esto, el grupo Mixteco es el que presenta el mayor número de individuos diabéticos, sin embargo, en promedio no son personas con sobrepeso u obesidad, lo que sugiere que existen otros factores que pueden estar participando en la aparición de DT2. A este respecto se conoce que existen mecanismos que pueden mediar la interacción entre los factores ambientales y

el genoma, como la epigenética, y que pueden ser de particular importancia en la patogénesis de la T2D. Por ejemplo, uno de estos mecanismos son los cambios de metilación que tienen lugar en genes candidatos identificados en GWAS [135] . También, existen evidencias que la microbiota intestinal está involucrada en el desarrollo de la DT2, alterando los procesos metabólicos, incluidos el metabolismo de la glucosa y los ácidos grasos [136]. De esta manera, como puede verse, el campo de estudio de la DT2 es muy amplio y puede ser abordado desde diversos puntos de vista.

Respecto a los valores de ICC observamos que las mujeres son las más afectadas en este sentido, ya que todos los grupos evaluados se encuentran fuera de los VR, el ICC es una medida indirecta de la distribución de la grasa en las regiones inferiores y superiores del cuerpo, y se relaciona de forma moderada con los factores de riesgo para desarrollar enfermedades cardiovasculares o metabólicas como la DT2, [137]. La obesidad visceral está asociada con deterioro de la sensibilidad a la insulina aumentado además el riesgo de desarrollar dislipidemia [137].

Los resultados anteriores nos muestran que la población mexicana posee muchos factores de riesgo para desarrollar DT2 y que aunado al componente genético que poseemos, se entiende de alguna manera la prevalencia de esta enfermedad en nuestro país.

Variantes de *GCK*

Glucocinasa es una enzima reguladora, clave para el metabolismo de la glucosa en la célula β pancreática, ya que cataliza la primera reacción de la vía glucolítica y tiene un papel crucial en la regulación de la secreción de insulina [138], más de 195 mutaciones han sido reportadas [86], por lo cual, debido a las características descritas de este gen es que fue elegido para ser evaluado en población mexicana.

A semejanza del gen *HNF4A*, en la primera etapa, realizamos la secuenciación exómica del gen *GCK*, identificando una mutación de sentido equivocado la Q26H (rs193922329) en la población Mazahua y 2 cambios sinónimos: D3D (rs144723656) y Y215Y (rs142553382), esta última en población Teenek (tabla 6). El cambio en la posición 215 fue descrito por Letho y cols., en 1999 [139]. La variante Q26H (rs193922329) ha sido reportada previamente [86], con sustitución Q26X originando un codón de paro, por lo que esta mutación es considerada patogénica, sin embargo en nuestro caso el cambio de base fue en la tercera posición del codón CAG/CAT cambiando una glutamina por una histidina (ambos

aminoácidos polares). Respecto a la persona portadora de la variante Q26H, encontramos que no presenta algún parámetro fuera de los VR. Se realizó la predicción de la funcionalidad de la enzima con el programa PolyPhen V2.2 SIFT y FATHMM y los resultados muestran que es un cambio que no altera la función de la proteína.

Estudios realizados en países europeos sugieren que la prevalencia de mutaciones en este gen, difieren considerablemente entre varios grupos étnicos [86]. En este contexto en Italia, del total de la población afectada con diabetes MODY el 61% corresponde a mutaciones en el gen de *GCK*, en Francia el 60%, en España 41% y en el Reino Unido 20%. En México estudios llevados a cabo en población mestiza diabética, no indican el hallazgo de alguna mutación de este gen que se asocie a la aparición de diabetes en nuestro país [89]. Lo que indica la variabilidad genética que existe en las poblaciones y la importancia de los estudios poblacionales que cada país debe realizar para conocer su fondo genético.

En el caso del cambio Q26H sería interesante recrear la mutación y evaluar si este cambio le confiere a la proteína un cambio en su actividad enzimática ya que, aunque los dos aminoácidos (Q y H) son de naturaleza polar, Q no tiene carga y H posee carga positiva, además, el cambio de estructura química sí podría modificar su actividad. Debido a que este cambio sólo lo encontramos en un sujeto del grupo mazahua sin alteraciones bioquímicas no fue buscado en los otros grupos.

La búsqueda de variantes de *GCK* en nuestra población, nos deja claro que no necesariamente los cambios genéticos de un grupo poblacional serán los mismos que originen la enfermedad en otro, así es importante que cada país realice sus propias evaluaciones.

La identificación de variantes y la evaluación de su asociación con factores de riesgo para la aparición de enfermedades complejas como la DT2 en las poblaciones indígenas es muy importante, porque nos permitiría la obtención de un mapa genético propio de la población indígena y también de la mestiza mexicana, para que se implementen terapias y desarrollen fármacos propios de nuestra población, además permitan predecir el curso de la enfermedad para tomar las mejores decisiones en las medidas preventivas y evitar las complicaciones propias de la diabetes, mejorando así la calidad de vida de los pacientes.

VIII. Conclusiones

1. La evaluación exómica del gene *HNF4A* evidenció la presencia de dos variantes intrónicas en la población indígena mexicana evaluada: rs736824, rs745975, y la variante rs1800961, que genera un cambio no sinónimo en la posición 130 de la proteína (T130I) y esta tiene un impacto estadístico sobre la concentración de triglicéridos.
2. La frecuencia de la variante T130I (rs1800961) del gen *HNF4A* en la población indígena mexicana es la más alta reportada hasta la actualidad.
3. La asociación entre la variante T130I (rs1800961) y las concentraciones elevadas de triglicéridos sugiere que sea un factor de riesgo en la susceptibilidad para el desarrollo de DT2 en las poblaciones mexicanas evaluadas.
4. A través del análisis de EHH se evidenció que la variante T130I (rs1800961) del gen *HNF4A* en la población indígena mexicana, podría ser el resultado de una selección positiva reciente.
5. La evaluación exómica del gene *GCK*, mostró la presencia de dos cambios sinónimos: D3D (rs144723656) y Y215Y (rs142553382) y uno no sinónimo Q26H (rs193922329), en población mazahua y teenek. La predicción informática sobre la variante Q26H no indica cambios en la función de la proteína.

IX. Perspectivas

1. Este estudio evidenció la necesidad de realizar la búsqueda y evaluación de los polimorfismos propios de cada población, ya que cada uno de ellos puede comportarse de manera diferente dependiendo del entorno en la que la comunidad se localice.
2. Se sugiere la búsqueda de las frecuencias de las variantes rs736824, rs745975 en la población mexicana y su relación con factores de riesgo para desarrollar DT2.
3. Diseñar experimentos que expliquen el posible mecanismo por el cual la variante T130I (rs1800961) de HNF4A influya en el metabolismo de lípidos incrementando específicamente los triglicéridos.

X. Bibliografía

1. Organización, et al. *Diabetes*. Diabetes Action Online, 2018. **2018**.
2. Li, M., L.J. Song, and X.Y. Qin, *Advances in the cellular immunological pathogenesis of type 1 diabetes*. J Cell Mol Med, 2014. **18**(5): p. 749-58.
3. Abdul-Ghani, M.A., D. Tripathy, and R.A. DeFronzo, *Contributions of beta-cell dysfunction and insulin resistance to the pathogenesis of impaired glucose tolerance and impaired fasting glucose*. Diabetes Care, 2006. **29**(5): p. 1130-9.
4. American Diabetes Association, A., *Diagnosis and classification of diabetes mellitus*. Diabetes Care, 2014. **37 Suppl 1**: p. S81-90.
5. International, Diabetes, and Federation, *IDF Diabetes Atlas*, IDF, Editor. 2015.
6. Krolewski, A.S., J.H. Warram, and M.B. Freire, *Epidemiology of late diabetic complications. A basis for the development and evaluation of preventive programs*. Endocrinol Metab Clin North Am, 1996. **25**(2): p. 217-242.
7. Arredondo, A., *Type 2 diabetes and health care costs in Latin America: exploring the need for greater preventive medicine*. BMC Med, 2014. **12**: p. 136.
8. Hernández-Ávila, M. *Encuesta Nacional de Salud y Nutrición de Medio Camino*. 2016.
9. Aida, et al. *Diabetes mellitus: La urgente necesidad de reforzar la respuesta en políticas públicas para su prevención y control*. Encuesta Nacional de Salud y Nutrición 2012, 2012.
10. Instituto, et al. *Kilos de más, pesos demenos. Los costos de la obesidad en México*. . 2015. 47.
11. Barraza-Lloréns M, G.-B.n.V., Picó J, García R, Hernández C, Mora F, Athié J, Crable E, Urtiz A, *Carga económica de la diabetes mellitus en México, 2013*, in D.R. © *Fundación Mexicana para la Salud, A.C*. 2015: Mexico.
12. OECD *Cardiovascular Disease and Diabetes: Policies for Better Health and Quality of Care*. 2015.
13. Olaiz-Fernández, G., et al. *Encuesta Nacional de Salud y Nutrición 2006*. ENSANUT, 2006.
14. Nathan, D.M., et al., *Impaired fasting glucose and impaired glucose tolerance: implications for care*. Diabetes Care, 2007. **30**(3): p. 753-9.
15. Williams, J.W., et al., *Gender differences in the prevalence of impaired fasting glycaemia and impaired glucose tolerance in Mauritius. Does sex matter?* Diabet Med, 2003. **20**(11): p. 915-20.
16. Fonseca, V.A., *Defining and characterizing the progression of type 2 diabetes*. Diabetes Care, 2009. **32 Suppl 2**: p. S151-6.
17. Chang, A.M. and J.B. Halter, *Aging and insulin secretion*. Am J Physiol Endocrinol Metab, 2003. **284**(1): p. E7-12.
18. Hu, F.B., *Globalization of diabetes: the role of diet, lifestyle, and genes*. Diabetes Care, 2011. **34**(6): p. 1249-57.
19. Unger, R.H. and Y.T. Zhou, *Lipotoxicity of beta-cells in obesity and in other causes of fatty acid spillover*. Diabetes, 2001. **50 Suppl 1**: p. S118-21.
20. Grant, R.W., A.F. Moore, and J.C. Florez, *Genetic architecture of type 2 diabetes: recent progress and clinical implications*. Diabetes Care, 2009. **32**(6): p. 1107-14.
21. Kahn, S.E., *Clinical review 135: The importance of beta-cell failure in the development and progression of type 2 diabetes*. J Clin Endocrinol Metab, 2001. **86**(9): p. 4047-58.

22. Shah, P., et al., *Lack of suppression of glucagon contributes to postprandial hyperglycemia in subjects with type 2 diabetes mellitus*. J Clin Endocrinol Metab, 2000. **85**(11): p. 4053-9.
23. Samuel, V.T. and G.I. Shulman, *The pathogenesis of insulin resistance: integrating signaling pathways and substrate flux*. J Clin Invest, 2016. **126**(1): p. 12-22.
24. Huang, S. and M.P. Czech, *The GLUT4 glucose transporter*. Cell Metab, 2007. **5**(4): p. 237-52.
25. Herman, M.A. and B.B. Kahn, *Glucose transport and sensing in the maintenance of glucose homeostasis and metabolic harmony*. J Clin Invest, 2006. **116**(7): p. 1767-75.
26. Bays, H., L. Mandarino, and R.A. DeFronzo, *Role of the adipocyte, free fatty acids, and ectopic fat in pathogenesis of type 2 diabetes mellitus: peroxisomal proliferator-activated receptor agonists provide a rational therapeutic approach*. J Clin Endocrinol Metab, 2004. **89**(2): p. 463-78.
27. Freeman, J.S., *Role of the incretin pathway in the pathogenesis of type 2 diabetes mellitus*. Cleve Clin J Med, 2009. **76 Suppl 5**: p. S12-9.
28. Nauck, M.A., B. Baller, and J.J. Meier, *Gastric inhibitory polypeptide and glucagon-like peptide-1 in the pathogenesis of type 2 diabetes*. Diabetes, 2004. **53 Suppl 3**: p. S190-6.
29. Ferrannini, E., *Sodium-glucose transporter-2 inhibition as an antidiabetic therapy*. Nephrol Dial Transplant, 2010. **25**(7): p. 2041-3.
30. DeFronzo, R.A., et al., *Characterization of renal glucose reabsorption in response to dapagliflozin in healthy subjects and subjects with type 2 diabetes*. Diabetes Care, 2013. **36**(10): p. 3169-76.
31. Meyer, C., et al., *Abnormal renal and hepatic glucose metabolism in type 2 diabetes mellitus*. J Clin Invest, 1998. **102**(3): p. 619-24.
32. Dedoussis, G.V., A.C. Kaliora, and D.B. Panagiotakos, *Genes, diet and type 2 diabetes mellitus: a review*. Rev Diabet Stud, 2007. **4**(1): p. 13-24.
33. Queitsch, C., K.D. Carlson, and S. Girirajan, *Lessons from model organisms: phenotypic robustness and missing heritability in complex disease*. PLoS Genet, 2012. **8**(11): p. e1003041.
34. Gibson, G., *Rare and common variants: twenty arguments*. Nat Rev Genet, 2012. **13**(2): p. 135-45.
35. Stranger, B.E., E.A. Stahl, and T. Raj, *Progress and promise of genome-wide association studies for human complex trait genetics*. Genetics, 2011. **187**(2): p. 367-83.
36. Meigs, J.B., L.A. Cupples, and P.W. Wilson, *Parental transmission of type 2 diabetes: the Framingham Offspring Study*. Diabetes, 2000. **49**(12): p. 2201-7.
37. Tattersal, R.B. and S.S. Fajans, *Prevalence of diabetes and glucose intolerance in 199 offspring of thirty-seven conjugal diabetic parents*. Diabetes, 1975. **24**(5): p. 452-62.
38. Pierce, M., H. Keen, and C. Bradley, *Risk of diabetes in offspring of parents with non-insulin-dependent diabetes*. Diabet Med, 1995. **12**(1): p. 6-13.
39. Henkin, L., et al., *Genetic epidemiology of insulin resistance and visceral adiposity. The IRAS Family Study design and methods*. Ann Epidemiol, 2003. **13**(4): p. 211-7.
40. Epidemiología, D.G.d. *Vigilancia Epidemiológica de la Diabetes Mellitus Tipo 2 Resultados Cierre 2016*. 2016.
41. McNeely, M.J. and E.J. Boyko, *Type 2 diabetes prevalence in Asian Americans: results of a national health survey*. Diabetes Care, 2004. **27**(1): p. 66-9.

42. Brancati, F.L., et al., *Incident type 2 diabetes mellitus in African American and white adults: the Atherosclerosis Risk in Communities Study*. JAMA, 2000. **283**(17): p. 2253-9.
43. (NCD-RisC), N.R.F.C., *Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants*. Lancet, 2016. **387**(10027): p. 1513-30.
44. Baier, L.J. and R.L. Hanson, *Genetic studies of the etiology of type 2 diabetes in Pima Indians: hunting for pieces to a complicated puzzle*. Diabetes, 2004. **53**(5): p. 1181-6.
45. Sousa, A.G., et al., *Association between genetics of diabetes, coronary artery disease, and macrovascular complications: exploring a common ground hypothesis*. Rev Diabet Stud, 2011. **8**(2): p. 230-44.
46. Dorajoo, R., J. Liu, and B.O. Boehm, *Genetics of Type 2 Diabetes and Clinical Utility*. Genes (Basel), 2015. **6**(2): p. 372-84.
47. Morris, A.P., et al., *Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes*. Nat Genet, 2012. **44**(9): p. 981-90.
48. Mori, Y., et al., *Genome-wide search for type 2 diabetes in Japanese affected sib-pairs confirms susceptibility genes on 3q, 15q, and 20q and identifies two new candidate Loci on 7p and 11p*. Diabetes, 2002. **51**(4): p. 1247-55.
49. Damcott, C.M., et al., *Polymorphisms in both promoters of hepatocyte nuclear factor 4-alpha are associated with type 2 diabetes in the Amish*. Diabetes, 2004. **53**(12): p. 3337-41.
50. Yin, L., et al., *Hepatic hepatocyte nuclear factor 4 α is essential for maintaining triglyceride and cholesterol homeostasis*. Arterioscler Thromb Vasc Biol, 2011. **31**(2): p. 328-36.
51. Inoue, Y., et al., *Hepatocyte nuclear factor 4alpha is a central regulator of bile acid conjugation*. J Biol Chem, 2004. **279**(4): p. 2480-9.
52. Dai, K. and M.M. Hussain, *NR2F1 disrupts synergistic activation of the MTP gene transcription by HNF-4 α and HNF-1 α* . J Lipid Res, 2012. **53**(5): p. 901-8.
53. Menjívar, M., et al., *High frequency of T130I mutation of HNF4A gene in Mexican patients with early-onset type 2 diabetes*. Clin Genet, 2008. **73**(2): p. 185-7.
54. Weedon, M.N., et al., *A common haplotype of the glucokinase gene alters fasting glucose and birth weight: association in six studies and population-genetics analyses*. Am J Hum Genet, 2006. **79**(6): p. 991-1001.
55. Froguel, P., et al., *Familial hyperglycemia due to mutations in glucokinase. Definition of a subtype of diabetes mellitus*. N Engl J Med, 1993. **328**(10): p. 697-702.
56. Dhe-Paganon, S., et al., *Crystal structure of the HNF4 alpha ligand binding domain in complex with endogenous fatty acid ligand*. J Biol Chem, 2002. **277**(41): p. 37973-6.
57. Sladek, F.M. *Sladek Lab*. 2018 [cited 2018 10/02/2018]; Available from: <http://www.sladeklab.ucr.edu/index.shtml>.
58. Winter, W.E., *Molecular and biochemical analysis of the MODY syndromes*. Pediatr Diabetes, 2000. **1**(2): p. 88-117.
59. Chen, W.S., et al., *Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos*. Genes Dev, 1994. **8**(20): p. 2466-77.
60. Babeu, J.P. and F. Boudreau, *Hepatocyte nuclear factor 4-alpha involvement in liver and intestinal inflammatory networks*. World J Gastroenterol, 2014. **20**(1): p. 22-30.

61. Odom, D.T., et al., *Control of pancreas and liver gene expression by HNF transcription factors*. Science, 2004. **303**(5662): p. 1378-81.
62. Miura, A., et al., *Hepatocyte nuclear factor-4alpha is essential for glucose-stimulated insulin secretion by pancreatic beta-cells*. J Biol Chem, 2006. **281**(8): p. 5246-57.
63. Gunton, J.E., et al., *Loss of ARNT/HIF1beta mediates altered gene expression and pancreatic-islet dysfunction in human type 2 diabetes*. Cell, 2005. **122**(3): p. 337-49.
64. Colclough, K., et al., *Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha and 4 alpha in maturity-onset diabetes of the young and hyperinsulinemic hypoglycemia*. Hum Mutat, 2013. **34**(5): p. 669-85.
65. Aguilar-Salinas, C.A., et al., *Early-onset type 2 diabetes: metabolic and genetic characterization in the mexican population*. J Clin Endocrinol Metab, 2001. **86**(1): p. 220-6.
66. Menjivar, M., et al., *Screening hepatocyte nuclear 40 gene in Mexican families with early onset type 2 diabetes*. 2003, ENDO 2003: The Endocrine Society 85th Annual Meeting: Philadelphia, Pennsylvania.
67. Yamagata, K., et al., *Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1)*. Nature, 1996. **384**(6608): p. 458-60.
68. Zhu, Q., et al., *T130I mutation in HNF-4alpha gene is a loss-of-function mutation in hepatocytes and is associated with late-onset Type 2 diabetes mellitus in Japanese subjects*. Diabetologia, 2003. **46**(4): p. 567-73.
69. Yang, Y., et al., *Identification of HNF4A Mutation p.T130I and HNF1A Mutations p.I27L and p.S487N in a Han Chinese Family with Early-Onset Maternally Inherited Type 2 Diabetes*. J Diabetes Res, 2016. **2016**: p. 3582616.
70. Chandra, V., et al., *Multidomain integration in the structure of the HNF-4α nuclear receptor complex*. Nature, 2013. **495**(7441): p. 394-8.
71. Wang, H., et al., *Hepatocyte nuclear factor 4alpha regulates the expression of pancreatic beta -cell genes implicated in glucose metabolism and nutrient-induced insulin secretion*. The Journal of Biological Chemistry, 2000. **275**(46): p. 7.
72. Møller, A.M., et al., *Studies of the genetic variability of the coding region of the hepatocyte nuclear factor-4alpha in Caucasians with maturity onset NIDDM*. Diabetologia, 1997. **40**(8): p. 980-3.
73. Rissanen, J., et al., *Variants in the hepatocyte nuclear factor-1alpha and -4alpha genes in Finnish and Chinese subjects with late-onset type 2 diabetes*. Diabetes Care, 2000. **23**(10): p. 1533-8.
74. Yang, X., et al., *Systematic genetic and genomic analysis of cytochrome P450 enzyme activities in human liver*. Genome Res, 2010. **20**(8): p. 1020-36.
75. Ek, J., et al., *The functional Thr130Ile and Val255Met polymorphisms of the hepatocyte nuclear factor-4alpha (HNF4A): gene associations with type 2 diabetes or altered beta-cell function among Danes*. J Clin Endocrinol Metab, 2005. **90**(5): p. 3054-9.
76. Ahn, K.J., et al., *Enzymatic properties of the N- and C-terminal halves of human hexokinase II*. BMB Rep, 2009. **42**(6): p. 350-5.
77. Froguel, P., et al., *Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus*. Nature, 1992. **356**(6365): p. 162-4.
78. Hattersley, A.T., et al., *Linkage of type 2 diabetes to the glucokinase gene*. Lancet, 1992. **339**(8805): p. 1307-10.
79. Kamata, K., et al., *Structural basis for allosteric regulation of the monomeric allosteric enzyme human glucokinase*. Structure, 2004. **12**(3): p. 429-38.

80. Magnuson, M.A., *Glucokinase gene structure. Functional implications of molecular genetic studies.* Diabetes, 1990. **39**(5): p. 523-7.
81. Iynedjian, P.B., *Mammalian glucokinase and its gene.* Biochem J, 1993. **293 (Pt 1)**: p. 1-13.
82. Matschinsky, F.M., *Glucokinase as glucose sensor and metabolic signal generator in pancreatic beta-cells and hepatocytes.* Diabetes, 1990. **39**(6): p. 647-52.
83. Ling, B., et al., *Theoretical investigations on the interactions of glucokinase regulatory protein with fructose phosphates.* Comput Biol Chem, 2016. **60**: p. 21-31.
84. Hattersley, A., P. Saker, and P. Patel, *Linkage of maturity-onset diabetes of the young to the glucokinase gene evidence of genetic heterogeneity.* Biochem Soc Trans 1992. **21**: p. 24S.
85. Cuesta-Muñoz, A.L., et al., *Severe persistent hyperinsulinemic hypoglycemia due to a de novo glucokinase mutation.* Diabetes, 2004. **53**(8): p. 2164-8.
86. Gloyn, A.L., *Glucokinase (GCK) mutations in hyper- and hypoglycemia: maturity-onset diabetes of the young, permanent neonatal diabetes, and hyperinsulinemia of infancy.* Hum Mutat, 2003. **22**(5): p. 353-62.
87. Spyer, G., et al., *Influence of maternal and fetal glucokinase mutations in gestational diabetes.* Am J Obstet Gynecol, 2001. **185**(1): p. 240-1.
88. Fu, D., et al., *Genetic polymorphism of glucokinase on the risk of type 2 diabetes and impaired glucose regulation: evidence based on 298,468 subjects.* PLoS One, 2013. **8**(2): p. e55727.
89. del Bosque-Plata, L., et al., *Analysis of the glucokinase gene in Mexican families displaying early-onset non-insulin-dependent diabetes mellitus including MODY families.* Am J Med Genet, 1997. **72**(4): p. 387-93.
90. Lisker, R. and V. Babinsky, *Admixture estimates in nine Mexican Indian groups and five East Coast localities.* Rev Invest Clin, 1986. **38**(2): p. 145-9.
91. Instituto, et al. *Estadísticas del... A propósito del Día Intencional de los Pueblos Indígenas.* 2016.
92. Secretaría, de, and Salud *Salud México 2001, Información para la rendición de cuentas.* 2001.
93. Lara-Riegos, J.C., et al., *Diabetes susceptibility in Mayas: Evidence for the involvement of polymorphisms in HHEX, HNF4alpha, KCNJ11, PPARgamma, CDKN2A/2B, SLC30A8, CDC123/CAMK1D, TCF7L2, ABCA1 and SLC16A11 genes.* Gene, 2015. **565**(1): p. 68-75.
94. Robles-Ordaz, M.D., et al., *Prevalence of prediabetes and modifiable factors in an ethnic group of Mexico: the Comcáac Project.* Public Health Nutr, 2018. **21**(2): p. 333-338.
95. Esparza-Romero, J., et al., *Environmentally Driven Increases in Type 2 Diabetes and Obesity in Pima Indians and Non-Pimas in Mexico Over a 15-Year Period: The Maycoba Project.* Diabetes Care, 2015. **38**(11): p. 2075-82.
96. Aguilar-Salinas, C.A., et al., *Prevalence of dyslipidemias in the Mexican National Health and Nutrition Survey 2006.* Salud Publica Mex, 2010. **52 Suppl 1**: p. S44-53.
97. Sánchez-Pozos, K. and M. Menjívar, *Genetic Component of Type 2 Diabetes in a Mexican Population.* Arch Med Res, 2016. **47**(7): p. 496-505.
98. García-Chapa, E.G., et al., *Genetic Epidemiology of Type 2 Diabetes in Mexican Mestizos.* Biomed Res Int, 2017. **2017**: p. 3937893.
99. Orita, M., et al., *Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction.* Genomics, 1989. **5**(4): p. 874-9.

100. Nollau, P. and C. Wagener, *Methods for detection of point mutations: performance and quality assessment*. IFCC Scientific Division, Committee on Molecular Biology Techniques. Clin Chem, 1997. **43**(7): p. 1114-28.
101. Mantovani, V., et al., *Identification of eight novel glucokinase mutations in Italian children with maturity-onset diabetes of the young*. Hum Mutat, 2003. **22**(4): p. 338.
102. Sanger, F., S. Nicklen, and A.R. Coulson, *DNA sequencing with chain-terminating inhibitors*. Proc Natl Acad Sci U S A, 1977. **74**(12): p. 5463-7.
103. Villarreal-Molina, M.T., et al., *The ATP-binding cassette transporter A1 R230C variant affects HDL cholesterol levels and BMI in the Mexican population: association with obesity and obesity-related comorbidities*. Diabetes, 2007. **56**(7): p. 1881-7.
104. Gautier, M., A. Klassmann, and R. Vitalis, *rehh 2.0: a reimplement of the R package rehh to detect positive selection from haplotype structure*. Mol Ecol Resour, 2017. **17**(1): p. 78-90.
105. Adzhubei, I., D.M. Jordan, and S.R. Sunyaev, *Predicting functional effect of human missense mutations using PolyPhen-2*. Curr Protoc Hum Genet, 2013. **Chapter 7**: p. Unit7.20.
106. Vaser, R., et al., *SIFT missense predictions for genomes*. Nat Protoc, 2016. **11**(1): p. 1-9.
107. Shihab, H.A., et al., *Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models*. Hum Mutat, 2013. **34**(1): p. 57-65.
108. Alexander, D.H., J. Novembre, and K. Lange, *Fast model-based estimation of ancestry in unrelated individuals*. Genome Res, 2009. **19**(9): p. 1655-64.
109. Project, N.H.G. *About the International HapMap Project*. 2012 [cited 2014 Feb 26]; Available from: <http://www.genome.gov/11511175-al-3>.
110. Sabeti, P.C., et al., *Detecting recent positive selection in the human genome from haplotype structure*. Nature, 2002. **419**(6909): p. 832-7.
111. Expert Panel on Detection, E.a., and Treatment of High Blood Cholesterol in Adults, *Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III)*. JAMA, 2001. **285**(19): p. 2486-97.
112. *Obesity: preventing and managing the global epidemic. Report of a WHO consultation*. World Health Organ Tech Rep Ser, 2000. **894**: p. i-xii, 1-253.
113. Canizales-Quinteros, S., et al., *Association of PPARG2 Pro12Ala variant with larger body mass index in Mestizo and Amerindian populations of Mexico*. Hum Biol, 2007. **79**(1): p. 111-9.
114. Long, J.C. and J.G. Lorenz, *Genetic polymorphism and American Indian health*. West J Med, 2002. **176**(3): p. 203-5.
115. Busfield, F., et al., *A genomewide search for type 2 diabetes-susceptibility genes in indigenous Australians*. Am J Hum Genet, 2002. **70**(2): p. 349-57.
116. Gao, X., et al., *Native American ancestry is associated with severe diabetic retinopathy in Latinos*. Invest Ophthalmol Vis Sci, 2014. **55**(9): p. 6041-5.
117. María Gabriela, R., et al. *Evaluación del número mínimo de marcadores para estimar ancestría individual en una muestra de la población argentina*. Revista del Museo de Antropología, 2016. 49-56.
118. Ruiz-Linares, A., et al., *Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals*. PLoS Genet, 2014. **10**(9): p. e1004572.

119. Gao, C., et al., *Exome Sequencing Identifies Genetic Variants Associated with Circulating Lipid Levels in Mexican Americans: The Insulin Resistance Atherosclerosis Family Study (IRASFS)*. *Sci Rep*, 2018. **8**(1): p. 5603.
120. Marcil, V., et al., *Hepatocyte nuclear factor 4 alpha polymorphisms and the metabolic syndrome in French-Canadian youth*. *PLoS One*, 2015. **10**(2): p. e0117238.
121. Moreno-Estrada, A., et al., *Human genetics. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits*. *Science*, 2014. **344**(6189): p. 1280-5.
122. Bersaglieri, T., et al., *Genetic signatures of strong recent positive selection at the lactase gene*. *Am J Hum Genet*, 2004. **74**(6): p. 1111-20.
123. Eaton, S.B. and L. Cordain, *Evolutionary aspects of diet: old genes, new fuels. Nutritional changes since agriculture*. *World Rev Nutr Diet*, 1997. **81**: p. 26-37.
124. Jafar-Mohammadi, B., et al., *A role for coding functional variants in HNF4A in type 2 diabetes susceptibility*. *Diabetologia*, 2011. **54**(1): p. 111-9.
125. Hellwege, J.N., et al., *Examination of Rare Variants in HNF4 α in European Americans with Type 2 Diabetes*. *J Diabetes Metab*, 2011. **2**(145).
126. Stančáková, A. and M. Laakso, *Genetics of Type 2 Diabetes*. *Endocr Dev*, 2016. **31**: p. 203-20.
127. Pavkov, M.E., et al., *Changing patterns of type 2 diabetes incidence among Pima Indians*. *Diabetes Care*, 2007. **30**(7): p. 1758-63.
128. Martínez-Hernández, A., et al., *Association of HMOX1 and NQO1 Polymorphisms with Metabolic Syndrome Components*. *PLoS One*, 2015. **10**(5): p. e0123313.
129. Parhofer, K.G., *Interaction between Glucose and Lipid Metabolism: More than Diabetic Dyslipidemia*. *Diabetes Metab J*, 2015. **39**(5): p. 353-62.
130. Lin, S.X., et al., *Does elevated plasma triglyceride level independently predict impaired fasting glucose?: the Multi-Ethnic Study of Atherosclerosis (MESA)*. *Diabetes Care*, 2013. **36**(2): p. 342-7.
131. Hayhurst, G.P., et al., *Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis*. *Mol Cell Biol*, 2001. **21**(4): p. 1393-403.
132. Sladek, F.M., et al., *Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily*. *Genes Dev*, 1990. **4**(12B): p. 2353-65.
133. Krapivner, S., et al., *DGAT1 participates in the effect of HNF4A on hepatic secretion of triglyceride-rich lipoproteins*. *Arterioscler Thromb Vasc Biol*, 2010. **30**(5): p. 962-7.
134. Rodríguez, A., et al., *Molecular Characterization of the Lipid Genome-Wide Association Study Signal on Chromosome 18q11.2 Implicates HNF4A-Mediated Regulation of the TMEM241 Gene*. *Arterioscler Thromb Vasc Biol*, 2016. **36**(7): p. 1350-5.
135. Davegårdh, C., et al., *DNA methylation in the pathogenesis of type 2 diabetes in humans*. *Mol Metab*, 2018.
136. Crommen, S. and M.C. Simon, *Microbial Regulation of Glucose Metabolism and Insulin Resistance*. *Genes (Basel)*, 2017. **9**(1).
137. Després, J.P., *Intra-abdominal obesity: an untreated risk factor for Type 2 diabetes and cardiovascular disease*. *J Endocrinol Invest*, 2006. **29**(3 Suppl): p. 77-82.
138. Matschinsky, F.M., *Regulation of pancreatic beta-cell glucokinase: from basics to therapeutics*. *Diabetes*, 2002. **51 Suppl 3**: p. S394-404.
139. Lehto, M., et al., *High frequency of mutations in MODY and mitochondrial genes in Scandinavian patients with familial early-onset diabetes*. *Diabetologia*, 1999. **42**(9): p. 1131-7.

Anexos

Anexo 1. Aprobación del comité de ética del Hospital Juárez de México



HOSPITAL JUÁREZ DE MÉXICO
COMISIONES DE INVESTIGACION, ÉTICA Y BIOSEGURIDAD

Agosto 22, 2007.

DRA. GUADALUPE ORTIZ LÓPEZ
COORDINADORA DEL LAB. DE
ENDOCRINOLOGIA MOLECULAR
P R E S E N T E

Muy apreciable Dra. Ortiz:

Muchas gracias por el envío de su protocolo titulado: "ESTUDIO DE LAS VARIANTES POLIMORFICAS DE LOS GENES HNF4A Y SLC30A8 COMO POSIBLES FACTORES DE RIESGO PARA DESARROLLAR DIABETES TIPO 2 EN POBLACIÓN MESTIZA E INDIGENA MEXICANA"

El cual ha sido evaluado por el Comité de Investigación de este Hospital y tengo el gusto de comunicarle que este ha sido aceptado, por lo que podrá Usted contar con todas las facilidades para su desarrollo. Su protocolo ha quedado registrado en la coordinación de protocolos de investigación del Hospital Juárez bajo el folio:

HJM1402/07.08.22

Sin otro en particular, solo quiero aprovechar la ocasión para desearle el mayor de los éxitos en esta empresa.

ATENTAMENTE

Dr. Alejandro Treviño Becerra
Presidente de la Comisión de Investigación

Ccp. Comisión de Investigación.

Av. Instituto Politécnico Nacional No. 5160 Col. Magdalena de las Salinas C.P. 07760 55.57.47.75.60

Anexo 2. Carta consentimiento



HOSPITAL JUÁREZ DE MÉXICO
DIRECCION DE INVESTIGACION
DIVISION DE INVESTIGACION
LABORATORIO DE ENDOCRINOLOGIA MOLECULAR

CARTA CONSENTIMIENTO

“Estudio de las variantes polimórficas de los genes HNF4A y GCK como posibles factores de riesgo para desarrollar diabetes tipo 2 en población mestiza e indígena mexicana”

Lugar y fecha _____

A quien corresponda:

El (la) que suscribe _____

De _____ años, acepto participar en forma libre y voluntaria, donando una muestra de sangre que será utilizada en el proyecto *“Estudio de las variantes polimórficas de los genes HNF4A y GCK como posibles factores de riesgo para desarrollar diabetes tipo 2 en población mestiza e indígena mexicana”*, que se lleva a cabo como parte de un convenio entre el Hospital Juárez de México y la Facultad de Química de la UNAM, y cuyo objetivo es identificar los genes responsables de la diabetes mellitus en población mexicana.

Se me ha informado de la importancia del estudio, así como de los objetivos del mismo y de la forma en que se tomará la muestra de sangre. Dándome la oportunidad para hacer preguntas libremente y estando satisfecho (a) de la información y respuestas recibidas acepto participar en él.

Firma del participante _____

Testigo 1

Testigo 2

Dirección _____

Firma del Investigador _____

Anexo 3. Cuestionario

Cuestionario



“Estudio de las variantes polimórficas de los genes HNF4A y GCK como posibles factores de riesgo para desarrollar diabetes tipo 2 en población mestiza e indígena mexicana

Datos personales:

Nombre: _____ Edad: _____
Ocupación: _____ Estado Civil: _____
Sexo: _____ Dirección _____
Escolaridad: _____ Origen étnico _____

Datos antropométricos:

Peso: _____ kg Talla _____ m IMC _____
Cintura _____ cm Cadera _____ cm ICC _____
Tensión arterial _____
Destrostix _____

Enfermedades:

Diabetes _____ Hipertensión _____ Hiperlipidemia _____
Otra _____

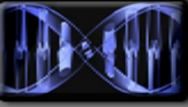
Antecedentes heredofamiliares:

¿Tiene o tuvo algún familiar directo con diabetes mellitus? _____

Observaciones

Ingiere alcohol: _____ poco _____ moderadamente _____ frecuentemente _____
Toma refresco: _____ poco _____ moderadamente _____ frecuentemente _____
Fuma: _____ poco _____ moderadamente _____ frecuentemente _____
Jugos/fruta fresca: _____ poco _____ moderadamente _____ frecuentemente: _____
Café/té/atole: _____ poco _____ moderadamente _____ frecuentemente _____
Pan/Tortillas/ Pan Dulce: _____ poco _____ moderadamente _____ frecuentemente _____
Huevos/frijoles/Carne: _____ poco _____ moderadamente _____ frecuentemente _____
Verduras de hoja verde/zanahorias/chayotes/Papas: _____ poco _____ moderadamente _____ frecuentemente _____
Realiza algún ejercicio: _____ ¿Cual? _____

Anexo 4. Predicción PolyPhen variante Q26H de GCK

 **PolyPhen-2** prediction of functional effects of human nsSNPs

Home About Help Downloads Batch query WHESS.db

PolyPhen-2 report for P35557 Q26H

Query

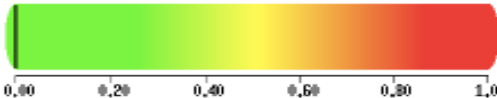
Protein Acc	Position	AA ₁	AA ₂	Description
P35557	26	Q	H	Canonical; RecName: Full=Glucokinase; EC=2.7.1.2; AltName: Full=Hexokinase type IV; Short=HK IV; AltName: Full=Hexokinase-4; Short=HK4; AltName: Full=Hexokinase-D; Length: 465

Results

+ Prediction/Confidence PolyPhen-2 v2.2.2r398

HumDiv

This mutation is predicted to be **BENIGN** with a score of 0.000 (sensitivity: 1.00; specificity: 0.00)



0.00 0.20 0.40 0.60 0.80 1.00

+ HumVar

Details

+ Multiple sequence alignment UniProtKB/UniRef100 Release 2011_12 (14-Dec-2011)

Predicción *in silico* Q26H GCK. PolyPhen V 2.2

Susceptibility background for type 2 diabetes in eleven Mexican Indigenous populations: *HNF4A* gene analysis

M. A. Granados-Silvestre¹ · M. G. Ortiz-López^{1,2} · J. Granados³ · S. Canizales-Quinteros^{1,4} · Rosenda I. Peñaloza-Espinosa⁵ · C. Lechuga⁶ · V. Acuña-Alonzo⁷ · K. Sánchez-Pozos¹ · M. Menjivar^{1,4}

Received: 3 November 2016 / Accepted: 20 June 2017
© Springer-Verlag GmbH Germany 2017

Abstract The genetic risk of developing type 2 diabetes (T2D) increases in parallel with the proportion of Native American ancestry. Mestizo Mexicans have a 70% Native Amerindian genetic background. The T130I polymorphism in the *HNF4A* gene has been associated with early-onset T2D in mestizo Mexicans. Thus, the aim of the present study was to evaluate the frequency and relationship of the T130I variant in the *HNF4A* gene with risk factors for developing T2D in eleven indigenous groups from Mexico. In two groups, all exons of the *HNF4A* gene were directly sequenced; in the remaining the T130I polymorphism was analyzed by restriction fragment length polymorphism. Ancestry informative markers were assessed to confirm the Amerindian component. An additional analysis of EHH was carried out. Interestingly, *HNF4A* gene screening

revealed only the presence of the T130I polymorphism. The range frequency of the risk allele (*T*) in the indigenous groups was from 2.7 to 16%. Genotypic frequencies (T130I/I130I) were higher and significantly different from those of all of the populations included in the HapMap Project ($P < 0.005$). EHH scores suggest a positive selection for T130I polymorphism. Metabolic traits indicate a relationship between the T130I/I130I genotypes with high triglyceride concentrations in the indigenous groups ($P < 0.005$). These results strongly suggest that the high frequency of the T130I polymorphism and its biological relationship with dysfunction in lipid metabolism in Mexican indigenous groups is a risk factor for the developing of T2D in Mexicans.

Keywords Mexican indigenous · T130I polymorphism · Diabetes background · *HNF4A* gene

Communicated by S. Hohmann.

✉ M. Menjivar
menjivar@unam.mx

¹ Departamento de Biología, Facultad de Química, Universidad Nacional Autónoma de México, Edif. F, 2o piso, Lab. 202, Av Universidad 3000, Circuito Interior S/N, Coyoacán, Cd. Universitaria, CP 04510 Mexico City, Mexico

² Laboratorio de Endocrinología Molecular, Hospital Juárez de México, Av. Instituto Politécnico Nacional 5160, Gustavo A. Madero, CP 07760 Mexico City, Mexico

³ Departamento de Trasplantes, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Vasco de Quiroga 15, Belisario Domínguez, Belisario Domínguez Secc 16, CP 14080 Mexico City, Mexico

⁴ Unidad de Genómica de Poblaciones Aplicada a la Salud, Facultad de Química Universidad Nacional Autónoma de México, Instituto Nacional de Medicina Genómica, Periferico Sur 4809, Tlalpan, Arenal Tepepan, CP 14610 Mexico City, Mexico

⁵ Unidad de Investigación Médica en Genética Humana, Centro Médico Nacional, Siglo XXI, Instituto Mexicano del Seguro Social (IMSS), Av. Cuauhtémoc 300, Cuauhtémoc, CP 06720 Mexico City, Mexico

⁶ Centro Cultural Universitario “Casa de las Diligencias”, Universidad Autónoma del Estado de México, Av Benito Juárez Garcia Sur 114, Centro, CP 50000 Toluca, Mexico State, Mexico

⁷ Escuela Nacional de Antropología e Historia, Instituto Nacional de Antropología e Historia, Mexico City, Mexico

Introduction

Many genes have been evaluated as candidates for T2D susceptibility. However, only variants in the *TCF7L2*, *PPARG*, *KCNJ11* and *HNFA4* genes have been extensively replicated in populations around the world, showing their indisputable association with T2D risk (Zeggini 2007). In the particular case of the *HNFA4* gene, it has been implicated in maturity-onset diabetes of the young type 1 (MODY 1) (Mitchell and Frayling 2002; Zhu et al. 2003). *HNFA4* is a member of the nuclear receptor super-family that plays a critical role in embryogenesis and metabolism, by regulating gene expression in pancreatic beta cells, liver and other tissues. The *HNFA4* gene is localized to chromosome 20q13, a region that has demonstrated evidence for linkage with T2D (Sladek et al. 1990; Ghosh et al. 1999). Several genetic studies, mainly in Caucasian and Asian populations, have provided evidence for the association of the variants in *HNFA4* with T2D (Ghosh et al. 1999; Silander et al. 2004; Winckler et al. 2005).

Interestingly, the T130I polymorphism in the *HNFA4* gene was associated with early-onset T2D in the mestizo Mexican population, with 16% as the highest frequency reported so far in diabetic subjects (Menjivar et al. 2008). Mestizo Mexicans are the largest Spanish-speaking population in Mexico; the population resulted mainly from the admixture between Amerindians and Spaniards. The prevalence of T2D is higher in populations of Native American ancestry and Hispanic Americans than in populations of European ancestry (Weiss et al. 1989; Parra et al. 2004). Genetic admixture studies have estimated that Mexicans show a higher proportion of Amerindian genetic components, ranging from 56% in Mexico City to 76% in Tlaxcala (Lisker et al. 1986). Recently, Juarez-Cedillo et al. reported a 69% of Amerindian, 26% of European and 5% of African contribution to the mestizo Mexican population (Juarez-Cedillo et al. 2008).

In Mexico, there are 25.7 million speakers of native languages, corresponding to 21.5% of the total population (Comisión Nacional para el desarrollo de los pueblos indígenas (CDI (2006)) and representing the most numerous group in Latin America.

In less than four decades, T2D has become the most important health problem in Mexico (Villalpando et al. 2010); its prevalence is the highest reported worldwide (14.4%), and approximately 2 of 10 diabetic patients develop the disease before the age of 40 years (Aguilar-Salinas et al. 2001; Barquera et al. 2013). Genetic predisposition is an important risk factor in Mexico; however, little is known about the diabetogenic background in this country.

Because Native American ancestry contributes significantly to the genetic composition of mestizo Mexicans, we hypothesized that it could be implicated in

their genetic susceptibility to develop T2D. Thus, the present work focuses on the molecular analyses of the *HNFA4* gene in eleven Mexican indigenous groups and a sample of a mestizo Mexican population to determine its relevance in the occurrence of T2D in the Mexican population.

Materials and methods

The study design was cross-sectional. Convenience sampling was used because the indigenous groups were located at regions of difficult access and in villages with few inhabitants.

Subjects

All of the participants were of more than three generations of indigenous origin, spoke their native language and preserved traditional folk heritage.

In total, 879 indigenes and 75 mestizos were studied. In the first part of the study, two Mexican indigenous groups and the sample of the mestizo Mexican population were included: Mazahua ($N = 91$), Teenek ($N = 69$) and mestizo ($N = 75$), in which eleven exons of the *HNFA4* gene were screened by direct sequencing.

In the next phase of study, Cora ($N = 124$), Chatino ($N = 112$), Chontal ($N = 141$), Maya ($N = 109$), Mixteco ($N = 133$), Purépecha ($N = 25$), Otomí ($N = 25$), Zapoteco ($N = 25$) and Yaqui ($N = 25$) groups were included to evaluate the frequency of the T130I polymorphism by restriction fragment length polymorphism (RFLP).

Cora is an ethnic group that inhabits a geographic area of northwestern Mexico called the Gran Nayar that includes the two small villages studied: Presidio de los Reyes and San Juan Corapan (Nayarit State). A sample of unrelated individuals of the Mixteco ethnic group, living in Santa Elena Xamiltepec, and the Chatino group from Santos Reyes Nopala village in the State of Oaxaca were also included. From Tabasco Chontal, individuals located at the Nacajuca region were studied. The Maya group included individuals from Hecelchakán, Pocboc, Pomuch and Tenabo villages (State of Campeche). The Mazahua group from the State of Mexico (next to Mexico City) lives in an isolated small village called Jaltepec. The Teenek population is located at the Huasteca region in San Luis Potosí State, and the studied group included subjects from two small villages (Tanzuman and San Isidro).

The mestizo group consisted of unrelated people from Mexico City who showed normal fasting blood glucose levels and were aged 60–65 years (to guarantee that they were not affected with early-onset T2D).

Informed consent was obtained from each volunteer during recruitment, carefully explaining the aims of the study. The protocol and informed consent procedure were approved by the Human Research Ethics Committee of the Hospital Juárez de México in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki.

Blood samples were taken after 12 h of fasting for biochemical, hormonal measurements and DNA isolation, for the Cora, Chatino, Chontal, Maya, Mazahua, Mixteco, Teenek and mestizo groups; however, for the Otomí, Purépecha, Yaqui and Zapoteco groups, only DNA was obtained. Anthropometric data were registered to calculate the body mass index (BMI) and waist-to-hip ratio (WHR).

Biochemical and hormonal determination

Blood glucose, triglycerides, total cholesterol, LDL-C and HDL-C were measured using commercially available standardized methods. The serum insulin concentration was measured by electrochemiluminescence (Elecsys E2010; Roche Diagnostics, Mannheim Germany).

Identification of *HNF4A* gene variants

In the primary screening direct sequencing was performed in two groups (Mazahua and Teenek). Genomic DNA was extracted from leucocytes by standard methods (Miller et al. 1988). The eleven exons (included exon 1a and 1b) and flanking intron regions of the *HNF4A* gene were amplified by PCR using specific primers (Moller et al. 1997) and were sequenced using the BigDye Terminator v1.1 Cycle Sequencing Kit and ABI PRISM DNA sequencer model 3130xl (Applied Biosystems, USA).

Detection of the T130I polymorphism in the *HNF4A* gene

Considering that after screening 235 individuals, only the T130I polymorphism was detected, this SNP was evaluated in the remaining nine groups by RFLP to determine the frequency of this variant. The PCR products were digested using the *S*fuI restriction enzyme (Roche, Penzberg, Germany) at 37 °C overnight and were applied in a 3% agarose gel. To confirm the presence of the T130I polymorphism of *HNF4A* as homozygous or heterozygous forms, direct sequencing was performed in each carrier, as well as in 5% of non T130I carriers.

Ancestry informative markers (AIMs)

In addition, ten AIMs were evaluated (rs4884, rs2695, rs17203, rs2862, rs3340, rs722098, rs203096, rs223830, rs1800498, and rs2814778), distinguishing mainly

Amerindian and European ancestries. SNPs were genotyped by KBiosciences, Hoddesdon, Herts., and UK using the KASPar system (Villarreal-Molina et al. 2007). This is a fluorescence-based allele-specific PCR procedure with improved robustness and discriminating power over conventional ARMS (<http://www.kbioscience.co.uk/chemistry/chemistry-intro.htm>). Genotyping accuracy was assessed by the inclusion of sample duplicates and negative controls. The genotyping call rates of each ancestry informative marker exceeded 95%, and no discordant genotypes were observed in 90 duplicate samples.

Extended haplotype homozygosity analysis (EHH)

Haplotype-based tests for positive selection were calculated using the R package, rehh version 2.0 (Gautier et al. 2017). Genetic variants without information on their ancestral allele were excluded from analyses. The haplotype included the following SNPs: rs7269793, rs6093978 and rs13041396 (upstream); rs13041396, rs11574739, rs3212194 and rs3212195 (downstream). The EHH plot was drawn using the R package, rehh. In this analysis only were considered the Mazahua, Teenek and mestizos groups from which the data sequencing was available.

Statistical analyses

Statistical Package for the Social Science (SPSS, Chicago, IL, USA) for Windows was used. All of the genotype distributions were tested for Hardy–Weinberg equilibrium using a program provided online (<http://ihg2.helmholtz-muenchen.de/cgi-bin/hw/hwt.pl>). Comparison of the allelic frequencies between the indigenous groups and HapMap Project subjects were performed using χ^2 test. The statistical analysis to find the relationship between the T130I polymorphism and metabolic phenotypes was performed only with non-diabetic subjects to avoid confounding effects with comorbidities developed secondary to T2D. For quantitative traits, covariance analysis was performed with age, sex and BMI as covariates, and the genotype was included as a fixed factor in the model (univariate general linear model). The insulin and triglyceride data were transformed to natural logarithm to normalize their distribution. *P* values <0.05 were considered to indicate statistical significance.

Results

The eleven exons of the *HNF4A* gene were screened by direct sequencing in two indigenous groups (Mazahua and Teenek) and in the mestizo Mexican group from Mexico City. Molecular analysis of the *HNF4A* gene revealed the

T130I polymorphism (rs1800961, exon 4) and two intronic variants, rs736824 y rs745975 (exon 2) were found, and this was made possible by the primers design. The frequency found for rs736824 was 45.8%, while for rs745975 it was 4.2%.

The ancestral component estimation of the indigenous and mestizo populations is shown in Table 1. Population structure analyses showed a mean admixture proportion of 97.6% Amerindian ancestry (AMA) and 2.4% European (EUR) ancestry in indigenous individuals and 76.8% AMA and 23.2% EUR in the mestizo population from Mexico City. The Teenek and Zapoteco groups reached the highest proportion of Amerindian ancestry (98.0 and 98.1%, respectively) among the indigenous populations. In Mexico, there are 68 indigenous groups located in different geographical regions that are grouped into eleven linguistic families. In this study, four linguistic families were explored as shown in Table 1, distributed across the country (Fig. 1).

The allele and genotype frequencies of the T130I polymorphism in the indigenous and mestizo groups are summarized in Table 2. The distribution of genotypes for all of the analyzed groups meets Hardy–Weinberg equilibrium except for the distribution observed in the Yaqui group, which does not show the T130I polymorphism. The Maya, Cora and Chontal groups show the highest risk allele frequency (14.7, 15.3 and 16.0%, respectively). The homozygous individuals for the T allele (thymine, which results in the 130I amino acid) were found in the Cora, Chontal, Maya and Mixteco groups.

Comparisons of the allelic distribution between the eleven Mexican ethnic groups and those reported previously in the HapMap Project are shown in Table 3. A

significant increased frequency of the T allele is observed in all of the studied groups compared with the Caucasian, Asian and African populations ($P < 0.005$). An additional analysis of EHH was carried out. The Fig. 2 shows the haplotype patterns, the EHH decay with distance of ancestral allele is slower than derived allele as it could be seen in the EHH statistics for both the ancestral and derived allele of the rs1800961 focal SNP.

The effect of the T130I polymorphism in the biochemical parameters in indigenous subjects is shown in Table 4. The T130I/I130I carriers without T2D presented significantly higher triglyceride concentrations than non-carriers ($P = 0.049$).

The anthropometric and biochemical variables of the indigenous groups were compared with those of the mestizo group. The mean BMI placed both groups into the overweight category ($26 \pm 5 \text{ kg/m}^2$ in indigenous; $26 \pm 4 \text{ kg/m}^2$ in mestizo Mexicans). Additionally, the population showed a high prevalence of central obesity, with values of WHR in indigenous women and men (0.86 ± 0.06 ; 0.92 ± 0.06 , respectively) and in mestizo women and men (0.88 ± 0.07 , 0.89 ± 0.1 , respectively) above the cut-off point according to the World Health Organization (Organization 2000).

It was also observed that 56.7% of the indigenous subjects showed concentrations of triglycerides above the estimated cut-off point ($<150 \text{ mg/dl}$). Significantly higher concentrations of triglycerides were observed in indigenous than in mestizo subjects from Mexico City [167 (111–248) vs 155 (127–211) mg/dl, respectively, $P < 0.05$], adjusted by sex and BMI. The insulin concentrations were lower in mestizo Mexicans than in indigenous individuals (6 ± 4 vs $14 \pm 13 \text{ mUI/ml}$, respectively) ($P < 0.005$).

Table 1 Ancestry of eleven Mexican indigenous groups and a sample of a mestizo Mexican population

Population	(<i>N</i> = 954)	Ancestry Amerindian (%)	Ancestry European (%)	Linguistic family (%)
Cora	13	97.78	2.21	Yuto-nahua
Chatino	11.7	97.41	2.58	Oto-mangue
Chontal	14.8	97.73	2.26	Maya
Maya	11.4	97.50	2.50	Maya
Mazahua	9.5	97.60	2.33	Oto-mangue
Mixteco	14	97.70	2.29	Oto-mangue
Otomí	2.6	96.64	3.35	Oto-mangue
Purépecha	2.6	97.59	2.40	Tarasco
Teenek	7.2	98.02	1.97	Maya
Yaqui	2.6	97.60	2.10	Yuto-nahua
Zapoteco	2.6	98.05	1.94	Oto-mangue
Mestizo*	7.9	76.84	23.2	

* Mestizo individuals from Mexico City



Fig. 1 Localization of Mexican indigenous groups studied

Table 2 Genotype and allelic frequencies of T130I polymorphism of *HNF4A* gene in eleven Mexican indigenous groups and a sample of a mestizo Mexican population

Ethnicity	TTT	TII	III	Total	T	I	Total	P
Mazahua	85 (93.4%)	6 (6.6%)		91	176 (96.7%)	6 (3.3%)	182	1.00002 (0.345–4.5)
Teenek	61 (88.4%)	8 (11.6%)		69	130 (94.2%)	8 (5.8%)	138	0.19287 (0.651–7.516)
Cora	89 (71.7%)	32 (25.8%)	3 (2.4%)	124	210 (84.6%)	38 (15.3%)	248	0.00007* (0.053–0.433)
Chatino	106 (94.6%)	6 (5.4%)		112	218 (97.3%)	6 (2.7%)	224	1.00001 (0.279–3.6)
Chontal	98 (70.0%)	41 (29.3%)	2 (1.4%)	141	237 (84.0%)	45 (16.0%)	282	0.00003* (2.44–19.6)
Maya	81 (74.3%)	24 (22.0%)	4 (3.7%)	109	186 (85.3%)	32 (14.7%)	218	0.00014* (2.17–18.15)
Mixteco	120 (90.2%)	10 (7.5)	3 (2.3%)	133	250 (93.9%)	16 (6.0%)	266	0.125 (0.766–7.12)
Otomí	23 (92%)	2 (8%)		25	48 (96%)	2 (4%)	50	0.74466 (0.270–8.565)
Purépecha	23 (92%)	2 (8%)		25	48 (96%)	2 (4%)	50	0.74466 (0.270–8.565)
Yaqui	25 (100%)	0 (0%)		25	50 (100%)	0 (0%)	50	0.8883 (0.017–6.092)
Zapoteca	22 (88%)	3 (12%)		25	47 (94%)	3 (6%)	50	0.43324 (0.507–10.787)
All indigenous	733 (83.4%)	134 (15.2%)	12 (1.4%)	879	1600 (91%)	158 (8.9%)	1758	0.00768* (1.317–9.86)
Mestizo	71 (94.6%)	4 (5.3%)		75	146 (97.3%)	4 (2.7%)	150	

Observed and expected phenotypes are in Hardy–Weinberg equilibrium, except Yaqui group. The frequency of the 130I allele, were compared between the Mexican-mestizo subjects and Amerindian groups. Comparison of allelic frequencies between Mexican-Mestizo and indigenous subjects was performed with a χ^2 test

* $P < 0.05$. A value of $P < 0.05$ was considered significant

Table 3 Comparison of T130I polymorphism in *HNF4A* gene. Genotype and allele frequencies of Amerindian groups (MEXAME): Cora, Chontal, Chatin, Maya, Mazahua, Teenek, Mixteco, Purepecha, Otomi, Zapoteco, Yaqui, relative to HapMap Project populations

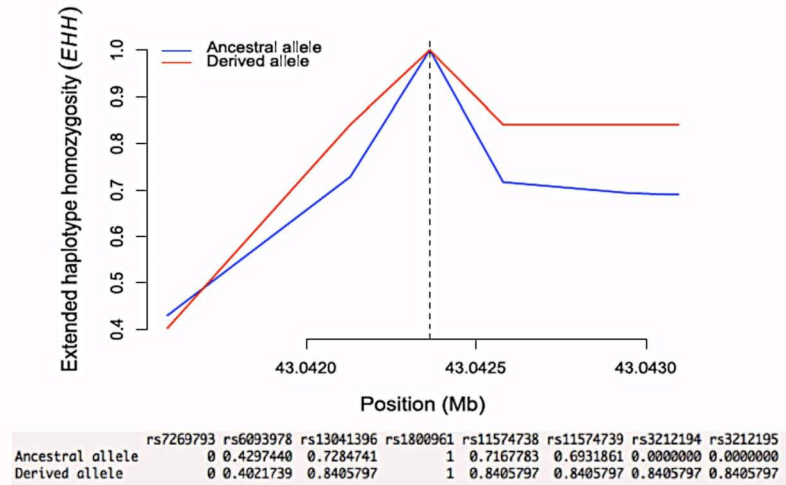
Population	Genotype				Allelic frequency			
	<i>n</i>	CC	<i>n</i>	CT/TT	<i>n</i>	<i>C</i>	<i>n</i>	<i>T</i>
ASW (A)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
CEU (C)	113	0.929	8	0.071*	226	0.965	8	0.035*
CHB (H)	137	0.964	5	0.036*	274	0.982	5	0.018*
CHD (D)	109	0.954	5	0.046*	218	0.977	5	0.023*
GIH (G)	101	0.950	5	0.050*	202	0.975	5	0.025*
JPT (J)	113	0.991	1	0.009*	226	0.996	1	0.004*
LWK (L)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
MEX (M)	58	0.983	1	0.017*	116	0.991	1	0.009*
MKK (K)	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
TSI (T)	102	0.971	3	0.029*	204	0.985	3	0.015*
YRI (Y)	63	1.000	0	n/a	126	1.000	0	0*
MEXAME (E)	733	0.834	146	0.166	1600	0.91	158	0.089

The values shown in bold are the genotypic and allelic frequencies of indigenous groups (MEXAME)

Population descriptors: ASW (A): African ancestry in Southwest USA, CEU (C): Utah residents with Northern and Western European ancestry from the CEPH collection, CHB (H): Han Chinese in Beijing, China, CHD (D): Chinese in Metropolitan Denver, Colorado, GIH (G): Gujarati Indians in Houston, Texas, JPT (J): Japanese in Tokyo, Japan, LWK (L): Luhya in Webuye, Kenya, MEX (M): Mexican ancestry in Los Angeles, California, MKK (K): Maasai in Kinyawa, Kenya, TSI (T): Tuscan in Italy, YRI (Y): Yoruban in Ibadan, Nigeria, MEXAME (E): Cora, Chontal, Chatin, Maya, Mazahua, Teenek, Mixteco, Purepecha, Otomi, Zapoteco, Yaqui

* Genotype and allele frequencies were compared to MEXAME $P < 0.005$. A value of $P < 0.05$ was considered significant

Fig. 2 Extended haplotype homozygosity of *HNF4A* haplotypes. The EHH decay shows evidence of positive selection for rs1800961 in *HNF4A* gene



Discussion

Presently, estimating how the genetic background affects population susceptibility to develop T2D is an enormous challenge. To gain a better understanding of the genetic component of diabetes in the Mexican population, the *HNF4A* gene was studied in eleven different indigenous

communities and a mestizo group. Interestingly, our results showed only one non-synonymous variant, T130I, is present at a high frequency in the studied population. Multi-centric studies have shown the clinical significance of the T130I polymorphism in the *HNF4A* gene, although with low frequency. By contrast, in this work, an elevated frequency of the T130I polymorphism was revealed, strongly

Table 4 Anthropometric and biochemical parameters stratified by T130I polymorphism in healthy Mexican indigenous

Parameters	T130T N = 558	T130I/T130I N = 117
BMI (kg/m ²)	26 ± 5	26 ± 5
WHR (male)	0.92 ± 0.06	0.90 ± 0.07
WHR (female)	0.86 ± 0.06	0.85 ± 0.06
Fasting glucose (mg/dl)	95 ± 12	95 ± 11
Triglyceride (mg/dl)	195 ± 107	220 ± 156*
Total cholesterol (mg/dl)	182 ± 40	185 ± 36
HDL cholesterol (mg/dl)	38 ± 12	39 ± 10
LDL cholesterol (mg/dl)	107 ± 38	103 ± 33
Insulin (mU/ml)	15 ± 13	14 ± 14

The relationship between genotypes and metabolic traits was evaluated with a univariate analysis adjusted by gender and BMI under the dominant model. Insulin and triglycerides values were log transformed. Data are mean ± SD

* $P < 0.05$. A value of $P < 0.05$ was considered significant

suggesting a relationship between the T130I polymorphism and susceptibility to T2D in the Mexican population.

The Mexican population is made up of indigenous and mestizo groups. The mestizo group is the result of a Caucasian (25%), Amerindian (70%), and African (5%) admixture (Juarez-Cedillo et al. 2008). The mestizos represent approximately 90% of the total inhabitants, and the remaining (10%) is constituted by Mexican indigenes, representing approximately 12.4 million people (Cédulas básicas de información de los pueblos indígenas de México 2017). Considering the information mentioned above, the study of the polymorphisms in indigenous Mexicans could be a practical approach to unravel the genetic inheritance of diabetes in the Mexican population. Moreover, the identification of health risk factors in these communities is very important because they are considered vulnerable populations because of their socioeconomic status. However, participants from the indigenous communities recognized themselves as indigenous, whose parents and grandparents were born in the same community, spoke the traditional language, and retained their traditions; the genetic ancestry was confirmed by AIMs. Although, presently, no population exists with a pure lineage, surprisingly, our results showed that the Amerindian component ranged from 97.6% in these communities, the lowest Amerindian contribution was observed in the Otomi group (96.4%), and the largest was observed in the Zapotec group (98.5%) (Table 1). All of these findings suggest a conservation of the Amerindian component in these populations because they did not have a deep genetic admixture due to isolation circumstances.

Previously, our group showed that the T130I polymorphism of the *HNF4A* gene was associated with the development of early-onset T2D in the mestizo Mexican population (Menjívar et al. 2008). To date, approximately 72 polymorphisms

of the *HNF4A* gene have been reported. However, only a few variants in the *HNF4A* gene have been associated with T2D in several populations around the world (Colclough et al. 2013). Thus, we decided to search for new variants in the *HNF4A* gene. In the first part of the study, molecular analysis of the eleven exons of the *HNF4A* gene in two indigenous populations (Mazahua and Teenek) and a mestizo group were performed. Interestingly, in this stage of the study, only one non-synonymous variant was identified, the T130I polymorphism. In addition, two intronic variants, rs736824 y rs745975 (exon 2) were found. Recently Marcil et al. showed that rs736824 and rs745975 are part of a haplotype with rs3212183, which was associated with metabolic syndrome in French-Canadian youth (Marcil et al. 2015). In the present study, the frequency of rs736824–rs745975 was 4.2%, (rs3212183 was not found), however further evaluations are required to elucidate its role as a risk factor for T2D in the Mexican indigenous population.

In the second part of the study, the T130I polymorphism was evaluated in another nine indigenous groups from Mexico (Cora, Chatino, Chontal, Maya, Mixteco, Otomi, Purepecha, Yaqui, and Zapoteco). The frequency of the risk allele 130I ranged from 2.7% in Chatino to 16% in Chontal, with a mean of 8.9% for all groups. However, in the mestizo Mexican population, this allele represented only 2.7% (Table 2). It is important to note that the 130I allele frequency observed in the 11 indigenous groups was higher and significantly different from that in the mestizo group ($P < 0.05$). The high frequency of the allele 130I allele in the Mexican indigenous population and high percentage of the Amerindian component in the mestizo population strongly suggest that this polymorphism is inherited from the first settlers from Mexico.

It is worth mentioning that the frequencies observed in the Mexican indigenous population 8.9% (Table 3) were also higher and significantly different ($P < 0.005$) than those found in the HapMap Project Phase III populations (Table 3), but African populations do not present this polymorphism, implying that this polymorphism was originated in the early settlers of America and that it was concentrated through generations. The presence of the T130I polymorphisms in indigenous populations suggests that this polymorphism in Mexico area inhabitants could be the result of an evolutive process. In this context, we carried out an analysis of EHH scores, considering the data of sequenced individuals. The results revealed a slower EHH decay of derived allele than ancestral allele (Fig. 2). Previous studies have shown, that a classic signal of strong selection is that because the favored allele increases in frequency very fast, it tends to sit on an unusually haplotype of low diversity, which is the case of rs1800961 (Voight et al. 2006). For instance, natural selection causes mutations and polymorphisms in key genes and their products, modifying metabolic

pathways to efficiently dispose of the available nutrients, such as the SNP in the gene encoding lactase (*LCT*) (C/T-13910) that is associated with the ability to digest milk in Europeans adults (lactase persistence). Moreover, it has been suggested that adaptive processes have forced the appearance of genetic variants that allowed the resistance of great famines and long journeys. However, under the current conditions, outside the hunter-gatherer context transmitted by Amerindian ancestors, genetic characteristics play a different role; thus, those benefits induced by an adaptive trait now become pathological under modern conditions. In this sense, the consumption of carbohydrates needed to endure long walks searching for food now favors disease occurrences such as diabetes, obesity and hypertension (Eaton and Cordain 1997). The *HNF4A* gene defines the expression of liver-specific genes encoding apolipoproteins, serum factors, cytochrome P-450 isoforms, and proteins involved in the metabolism of glucose, fatty acids, and amino-acids; however, the clinical characterization of MODY1 subjects revealed that primary defects impaired glucose-stimulated insulin secretion from pancreatic β -cell. Thus, it is possible to hypothesize that T130I in *HNF4A* gene mutation could promote glucose low utilization at the cellular level in an environment poor of carbohydrates during the nomadic step.

Ancestry is an important variable that influences an individual's health in several ways. In particular, it has been reported that an ancestral component has an impact in the risk for development of chronic diseases such as T2D (Flegel et al. 1991; Gao et al. 2014). For instance, Pima Indian groups from the United States and Mexico show a high prevalence of T2D, strongly suggesting a genetic susceptibility background in Amerindian groups (Schulz et al. 2006; Pavkov et al. 2007; Rong et al. 2009; Lara-Riegosa et al. 2015). In this context, the ATP-binding cassette transporter A1 (*ABCA1*) R230C variant was associated with low HDL-cholesterol concentration, showing evidence for positive selection on Native Americans (Pavkov et al. 2007). Additionally, Villarreal-Molina et al. found an association of the R230C variant of *ABCA1* with early-onset T2D in the Mexican population (Villarreal-Molina et al. 2008). On the other hand, the association of polymorphisms in the *HMOX* and *NQO1* genes with components of metabolic syndrome in Mexican indigenous populations has been reported (Martinez-Hernandez et al. 2015). The presence of SNPs such as T130I in *HNF4A*, R230C in *ABCA1* and polymorphisms in the *HMOX* and *NQO1* genes, among other polymorphisms with minor metabolic effects, strongly suggests an additive genetic contribution of these polymorphisms in the appearance of metabolic diseases such as T2D, a finding that can be even more important under malnutrition conditions in low-income populations such as indigenous Mexican populations. The identification

of haplotypes involved in health and disease in each population will allow the development of new diagnostic technologies to lay the foundations for personalized medicine.

Regarding the T130I polymorphism, the functional effect has been defined clearly. This variant has been shown to have a deleterious effect on protein activity based on the studies of the transcriptional regulation of *HNF4A* target genes in cell lines and primary mouse hepatocytes (Navas et al. 1999; Zhu et al. 2003). These studies showed that T130I-*HNF4A* showed reduced transcriptional activity in HepG2 cells and primary cultured mouse hepatocytes (27–78% of wild type, $P < 0.05$). Likewise, the *HNF4A* gene is implicated in pancreatic development and the control of insulin secretion, in part by regulating GLUT-2 transcription (Wang et al. 2000; Rodriguez-Trejo et al. 2012). In addition to its primary role as a hepatic gene regulator (Hayhurst et al. 2001), disruption of the *HNF4A* gene can lead to autosomal-dominant Maturity-Onset Diabetes of the Young (MODY 1), a non-insulin-dependent form of diabetes (Yamagata et al. 1996). Thus, to be a carrier of T130I is an undeniable risk factor to develop T2D.

Jafar-Mohammadi et al. evaluated an extensive collection of rare and common *HNF4A* gene variants to assess their prevalence and association in a T2D case-control study in a large sample. They found that the T130I variant (frequency 6.5%) was associated with T2D, with an additive P value of 2.1×10^{-5} (Jafar-Mohammadi et al. 2011). Similar results were obtained in an evaluation of the T130I variant performed by Hellwege et al., with a sample of more than 2000 European American T2D cases and controls. The findings showed that the odds ratios were in the same direction, and they were similar in magnitude for the additive model of Jafar-Mohammadi's study, although the P value did not reach significance for an association (Hellwege et al. 2011). Interestingly, both studies suggested the a constant, although there was a modest risk of the T130I variant in the *HNF4A* gene for the development of T2D that resembles TCF7L2, whose polymorphisms have been consistently associated with T2D around the world (Cauchi et al. 2007). In the same way, Menjivar et al. reported the association of the T130I variant in mestizo Mexicans with an early onset of T2D, indicating a role of the variant in the pathogenesis of this disease in Mexico (Menjivar et al. 2008). The high frequency of the risk allele (8.9%) among the indigenous populations studied suggests a diabetogenic background in Mexican founder populations (Table 2). To understand the impact of the T130I variant in the susceptibility to develop T2D, their evaluation should be performed considering the different genetic background among populations.

On the other hand, dyslipidemia has also been associated as a risk factor for the development of T2D; in this sense, high circulating levels of triglycerides lead to increased free

fatty acid concentrations, which may induce insulin resistance and β -cell dysfunction (Parhofer 2015). Lipid disorders are frequent in Mexican adults (Aguilar-Salinas et al. 2010). In addition, many studies have also shown that low HDL cholesterol and high triglyceride concentration are independent risk factors for development T2D and impaired fasting glucose (Wilson et al. 2007; Lin et al. 2013). Moreover, high HWR values indicate visceral obesity, which is associated with an increased cardiovascular risk and likelihood to develop diseases such as T2D. Anthropometric and biochemical results in this study showed a value of HWR above the cut-off point in 86.4% of indigenous women and an HDL-C concentration <40 mg/dl in 58.6% of the entire population, suggesting a high risk to develop T2D in these communities due to impaired lipid metabolism. Interestingly, the relationship found between the T130I variant and high triglyceride concentration in non-diabetic subjects (Table 4) provides additional evidence about the genetic risk to develop T2D in populations with Amerindian ancestry. Moreover, the gene-environment interaction between an obesogenic environment (where indigenous populations live) and being a carrier of the T130I polymorphism (*HNF4A* is an essential gene for the hepatic response to changes in nutritional status) increase the genetic impact on vulnerable populations such as indigenous populations to develop T2D (Adamson et al. 2006).

This is a first study in the discovery phase carried out in the Mexican indigenous populations to evaluate if the *HNF4A* gene is implicated in the pathogenesis of T2D. The frequency of the T130I variant in eleven Mexican groups and its relationship with metabolic traits, particularly with triglycerides in non-diabetic subjects, showed a significant correlation ($P = 0.049$). These results provide the basis for further studies of association with this variant in Mexican indigenous with T2D.

Finally, the presence of the T130I polymorphism and their correlations with metabolic traits observed in the studied Mexican indigenous populations place this polymorphism as a risk factor for developing T2D. In addition, the high frequency of the T130I polymorphism in Mexican indigenous populations reveals a diabetogenic background, which contributes to the high prevalence of T2D in the Mexican population.

Acknowledgements We thank the Health Departments of each of the participating States.

Compliance with ethical standards

Funding This study was funded by CONACYT (scholarship 201106) and PAPIIT-DGAPA (Grant no. IN208507).

Conflict of interest Granados-Silvestre MA declares that she has no conflict of interest; Ortiz-López MG declares that she has no con-

flict of interest; Granados J declares that he has no conflict of interest; Canizales-Quinteros S declares that he has no conflict of interest; Rosenda I Peñaloza-Espinosa declares that she has no conflict of interest; Lechuga C declares that she has no conflict of interest; Acuña-Alonzo V declares that he has no conflict of interest; Sánchez-Pozos K declares that she has no conflict of interest and Menjívar M declares that she has no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent: Informed consent was obtained from all individual participants included in the study.

References

- Adamson AW, Suchankova G, Rufo C, Nakamura MT, Teran-Garcia M, Clarke SD, Gettys TW (2006) Hepatocyte nuclear factor-4 α contributes to carbohydrate-induced transcriptional activation of hepatic fatty acid synthase. *Biochem J* 399:285–295
- Aguilar-Salinas CA, Reyes-Rodríguez E, Ordonez-Sanchez ML, Torres MA, Ramirez-Jimenez S, Dominguez-Lopez A, Martinez-Francois JR, Velasco-Perez ML, Alpizar M, Garcia-Garcia E, Gomez-Perez F, Rull J, Tusie-Luna MT (2001) Early-onset type 2 diabetes: metabolic and genetic characterization in the Mexican population. *J Clin Endocrinol Metab* 86:220–226
- Aguilar-Salinas CA, Gomez-Perez FJ, Rull J, Villalpando S, Barquera S, Rojas R (2010) Prevalence of dyslipidemias in the Mexican National Health and Nutrition Survey 2006. *Salud Publica Mex* 52(Suppl 1):S44–S53
- Barquera S, Campos-Nonato I, Aguilar-Salinas C, Lopez-Ridaura R, Arredondo A, Rivera-Dommarco J (2013) Diabetes in Mexico: cost and management of diabetes and its complications and challenges for health policy. *Global Health* 9:3
- Cauchi S, El Achhab Y, Choquet H, Dina C, Krempler F, Weitgasser R, Nejjari C, Patsch W, Chikri M, Meyre D, Froguel P (2007) TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. *J Mol Med* 85:777–782
- Cédulas básicas de información de los pueblos indígenas de México (2017). <http://www.cdi.gob.mx/>
- Colclough K, Bellanne-Chanelot C, Saint-Martin C, Flanagan SE, Ellard S (2013) Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha and 4 alpha in maturity-onset diabetes of the young and hyperinsulinemic hypoglycemia. *Hum Mutat* 34:669–685
- Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CDI) (2006) Indicadores de la población indígena. Sobre desarrollo humano de los pueblos indígenas de México 2006. <http://www.gob.mx/cms/uploads/attachment/file/35732/cdi-informe-desarrollo-humano-pueblos-indigenas-mexico-2006.pdf>. Accessed 3 July 2017
- Eaton SB, Cordain L (1997) Evolutionary aspects of diet: old genes, new fuels. Nutritional changes since agriculture. *World Rev Nutr Diet* 81:26–37
- Flegal KMET, Harris MI, Haynes SG, Juarez RZ, Knowler WC, Perez-Stable EJ, Stern MP (1991) Prevalence of diabetes in Mexican Americans, Cubans, and Puerto Ricans from the Hispanic Health and Nutrition Examination Survey, 1982–1984. *Diabetes Care* 14:628–638
- Gao X, Gauderman WJ, Marjoram P, Torres M, Chen YD, Taylor KD, Rotter JI, Varma R (2014) Native American ancestry is

- associated with severe diabetic retinopathy in Latinos. *Invest Ophthalmol Vis Sci* 55:6041–6045
- Gautier M, Klassmann A, Vitalis R (2017) rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. *Mol Ecol Resour* 17:78–90
- Ghosh S, Watanabe RM, Hauser ER, Valle T, Magnuson VL, Erdos MR, Langefeld CD, Balow J Jr, Ally DS, Kohtamaki K, Chines P, Birznieks G, Kaleta HS, Musick A, Te C, Tannenbaum J, Eldridge W, Shapiro S, Martin C, Witt A, So A, Chang J, Shurtleff B, Porter R, Kudelko K, Unni A, Segal L, Sharaf R, Blaschak-Harvan J, Eriksson J, Tenkula T, Vidgren G, Ehnholm C, Tuomilehto-Wolf E, Hagopian W, Buchanan TA, Tuomilehto J, Bergman RN, Collins FS, Boehnke M (1999) Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. *Proc Natl Acad Sci USA* 96:2198–2203
- Hayhurst GP, Lee YH, Lambert G, Ward JM, Gonzalez FJ (2001) Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. *Mol Cell Biol* 21:1393–1403
- Hellwege JN, Hicks PJ, Palmer ND, Ng MC, Freedman BI, Bowden DW (2011) Examination of rare variants in HNF4 alpha in European Americans with type 2 diabetes. *J Diabetes Metab* 2(145):1000145. doi:10.4172/2155-6156.1000145
- Jafar-Mohammadi B, Groves CJ, Gjessing AP, Herrera BM, Winckler W, Stringham HM, Morris AP, Lauritzen T, Doney AS, Morris AD, Weedon MN, Swift AJ, Kuusisto J, Laakso M, Altschuler D, Hattersley AT, Collins FS, Boehnke M, Hansen T, Pedersen O, Palmer CN, Frayling TM, Gloyn AL, McCarthy MI (2011) A role for coding functional variants in HNF4A in type 2 diabetes susceptibility. *Diabetologia* 54:111–119
- Juarez-Cedillo T, Zuniga J, Acuna-Alonzo V, Perez-Hernandez N, Rodriguez-Perez JM, Barquera R, Gallardo GJ, Sanchez-Arenas R, Garcia-Pena Mdel C, Granados J, Vargas-Alarcon G (2008) Genetic admixture and diversity estimations in the Mexican Mestizo population from Mexico City using 15 STR polymorphic markers. *Forensic Sci Int Genet* 2:e37–e39
- Lara-Riegos JC, Ortiz-López MG, Peña-Espinoza BI, Montúfar-Robles I, Peña-Rico MA, Sánchez-Pozos K, Granados-Silvestre MA, Menjivar M (2015) Diabetes susceptibility in Mayas: evidence for the involvement of polymorphisms in HHEX, HNF4α, KCNJ11, PPARγ, CDKN2A/2B, SLC30A8, CDC123/CAMK1D, TCF7L2, ABCA1 and SLC16A11 genes. *Gene* 565:68–75
- Lin SX, Berlin I, Younger R, Jin Z, Sibley CT, Schreiner P, Szklo M, Bertoni AG (2013) Does elevated plasma triglyceride level independently predict impaired fasting glucose? the multi-ethnic study of atherosclerosis (MESA). *Diabetes Care* 36:342–347
- Lisker R, Perez-Briceno R, Granados J, Babinsky V, de Rubens J, Armendares S, Buentello L (1986) Gene frequencies and admixture estimates in a Mexico City population. *Am J Phys Anthropol* 71:203–207
- Marcel V, Amre D, Seidman EG, Boudreau F, Gendron FP, Menard D, Beaulieu JF, Sinnett D, Lambert M, Levy E (2015) Hepatocyte nuclear factor 4 alpha polymorphisms and the metabolic syndrome in French-Canadian youth. *PLoS One* 10:e0117238
- Martinez-Hernandez A, Cordova EJ, Rosillo-Salazar O, Garcia-Ortiz H, Contreras-Cubas C, Islas-Andrade S, Revilla-Monsalve C, Salas-Labadia C, Orozco L (2015) Association of HMOX1 and NQO1 polymorphisms with metabolic syndrome components. *PLoS One* 10:e0123313
- Menjivar M, Granados-Silvestre MA, Montufar-Robles I, Herrera M, Tusie-Luna MT, Canizales-Quinteros S, Aguilar-Salinas CA, Ortiz-Lopez MG (2008) High frequency of T130I mutation of HNF4A gene in Mexican patients with early-onset type 2 diabetes. *Clin Genet* 73:185–187
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16:1215
- Mitchell SM, Frayling TM (2002) The role of transcription factors in maturity-onset diabetes of the young. *Mol Genet Metab* 77:35–43
- Moller AM, Urhammer SA, Dalgaard LT, Reneland R, Berglund L, Hansen T, Clausen JO, Lithell H, Pedersen O (1997) Studies of the genetic variability of the coding region of the hepatocyte nuclear factor-4alpha in Caucasians with maturity onset NIDDM. *Diabetologia* 40:980–983
- Navas MA, Munoz-Elias EJ, Kim J, Shih D, Stoffel M (1999) Functional characterization of the MODY1 gene mutations HNF4(R127W), HNF4(V255M), and HNF4(E276Q). *Diabetes* 48:1459–1465
- Organization WH (2000) Obesity: preventing and managing the global epidemic report of a WHO Consultation (WHO Technical Report Series 894). In: 894 (ed) WHO Technical Report Series 894. WHO, Geneva, p 252
- Parhofer KG (2015) Interaction between glucose and lipid metabolism: more than diabetic dyslipidemia. *Diabetes Metab J* 39:353–362
- Parra EJ, Hoggart CJ, Bonilla C, Dios S, Norris JM, Marshall JA, Hamman RF, Ferrell RE, McKeigue PM, Shriver MD (2004) Relation of type 2 diabetes to individual admixture and candidate gene polymorphisms in the Hispanic American population of San Luis Valley, Colorado. *J Med Genet* 41:e116
- Pavkov ME, Hanson RL, Knowler WC, Bennett PH, Krakoff J, Nelson RG (2007) Changing patterns of type 2 diabetes incidence among Pima Indians. *Diabetes Care* 30:1758–1763
- Rodriguez-Trejo A, Ortiz-Lopez MG, Zambrano E, Granados-Silvestre Mde L, Mendez C, Blondeau B, Breant B, Nathanielsz PW, Menjivar M (2012) Developmental programming of neonatal pancreatic beta-cells by a maternal low-protein diet in rats involves a switch from proliferation to differentiation. *Am J Physiol Endocrinol Metab* 302:E1431–E1439
- Rong R, Hanson RL, Ortiz D, Wiedrich C, Kobes S, Knowler WC, Bogardus C, Baier LJ (2009) Association analysis of variation in/near FTO, CDKAL1, SLC30A8, HHEX, EXT2, IGF2BP2, LOC387761, and CDKN2B with type 2 diabetes and related quantitative traits in Pima Indians. *Diabetes* 58:478–488
- Schulz LO, Bennett PH, Ravussin E, Kidd JR, Kidd KK, Esparza J, Valencia ME (2006) Effects of traditional and western environments on prevalence of type 2 diabetes in Pima Indians in Mexico and the U.S. *Diabetes Care* 29:1866–1871
- Silander K, Mohlke KL, Scott LJ, Peck EC, Hollstein P, Skol AD, Jackson AU, Deloukas P, Hunt S, Stavrides G, Chines PS, Erdos MR, Narisu N, Conneely KN, Li C, Fingerlin TE, Dhanjal SK, Valle TT, Bergman RN, Tuomilehto J, Watanabe RM, Boehnke M, Collins FS (2004) Genetic variation near the hepatocyte nuclear factor-4 alpha gene predicts susceptibility to type 2 diabetes. *Diabetes* 53:1141–1149
- Sladek FM, Zhong WM, Lai E, Darnell JE Jr (1990) Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes Dev* 4:2353–2365
- Villalpando S, Shamah-Levy T, Rojas R, Aguilar-Salinas CA (2010) Trends for type 2 diabetes and other cardiovascular risk factors in Mexico from 1993–2006. *Salud Publica Mex* 52(Suppl 1):S72–S79
- Villarreal-Molina MT, Aguilar-Salinas CA, Rodriguez-Cruz M, Riano D, Villalobos-Comparan M, Coral-Vazquez R, Menjivar M, Yescas-Gomez P, Konigsoerg-Fainstein M, Romero-Hidalgo S, Tusie-Luna MT, Canizales-Quinteros S (2007) The ATP-binding cassette transporter A1 R230C variant affects HDL cholesterol levels and BMI in the Mexican population: association with obesity and obesity-related comorbidities. *Diabetes* 56:1881–1887

- Villarreal-Molina MT, Flores-Dorantes MT, Arellano-Campos O, Villalobos-Comparan M, Rodriguez-Cruz M, Miliar-Garcia A, Huertas-Vazquez A, Menjivar M, Romero-Hidalgo S, Wacher NH, Tusie-Luna MT, Cruz M, Aguilar-Salinas CA, Canizales-Quinteros S, Metabolic Study G (2008) Association of the ATP-binding cassette transporter A1 R230C variant with early-onset type 2 diabetes in a Mexican population. *Diabetes* 57:509–513
- Voight BF, Kudaravalli S, Wen X, Pritchard JK (2006) A map of recent positive selection in the human genome. *PLoS Biol* 4:e72
- Wang H, Maechler P, Antinozzi PA, Hagenfeldt KA, Wollheim CB (2000) Hepatocyte nuclear factor 4alpha regulates the expression of pancreatic beta-cell genes implicated in glucose metabolism and nutrient-induced insulin secretion. *J Biol Chem* 275:35953–35959
- Weiss KM, Ulbrecht JS, Cavanagh PR, Buchanan AV (1989) Diabetes mellitus in American Indians: characteristics, origins and preventive health care implications. *Med Anthropol* 11:283–304
- Wilson PW, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB Sr (2007) Prediction of incident diabetes mellitus in middle-aged adults: the Framingham offspring study. *Arch Intern Med* 167:1068–1074
- Winckler W, Graham RR, de Bakker PI, Sun M, Almgren P, Tuomi T, Gaudet D, Hudson TJ, Ardlie KG, Daly MJ, Hirschhorn JN, Groop L, Altshuler D (2005) Association testing of variants in the hepatocyte nuclear factor 4alpha gene with risk of type 2 diabetes in 7,883 people. *Diabetes* 54:886–892
- Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, Cox NJ, Fajans SS, Signorini S, Stoffel M, Bell GI (1996) Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature* 384:458–460
- Zeggini E (2007) A new era for Type 2 diabetes genetics. *Diabet Med* 24:1181–1186
- Zhu Q, Yamagata K, Miura A, Shihara N, Horikawa Y, Takeda J, Miyagawa J, Matsuzawa Y (2003) T130I mutation in HNF-4alpha gene is a loss-of-function mutation in hepatocytes and is associated with late-onset Type 2 diabetes mellitus in Japanese subjects. *Diabetologia* 46:567–573