



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE QUÍMICA

**IDENTIFICACIÓN DE VARIANTES EN EL NÚMERO DE COPIAS Y SU ASOCIACIÓN CON
SÍNDROME METABÓLICO EN POBLACIÓN AMERINDIA MEXICANA**

TESIS

QUE PARA OBTENER EL TÍTULO DE

QUÍMICO FARMACÉUTICO BIÓLOGO

PRESENTA

DAVID GIRÓN VILLALOBOS



CIUDAD DE MÉXICO

2021



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.

JURADO ASIGNADO:

PRESIDENTE: **Profesor: Marisol López López**

VOCAL: **Profesor: Alberto Ortega Vázquez**

SECRETARIO: **Profesor: Angélica Graciela Martínez Hernández**

1er. SUPLENTE: **Profesor: Javier Axosco Marín**

2° SUPLENTE: **Profesor: Manuel Gutiérrez Aguilar**

SITIO DONDE SE DESARROLLÓ EL TEMA:

LABORATORIO DE INMUNOGENÓMICA Y ENFERMEDADES METABÓLICAS, INSTITUTO NACIONAL DE MEDICINA GENÓMICA (INMEGEN)

ASESOR DEL TEMA:

Dra. Angélica Graciela Martínez Hernández

SUSTENTANTE (S):

David Girón Villalobos

Índice

Resumen	1
1 Introducción	2
1.1 Síndrome metabólico	2
1.1.1 Diagnóstico	2
1.1.2 Prevalencia	3
1.1.3 Fisiopatología	4
1.1.4 Factores ambientales	5
1.1.5 Factores genéticos	7
1.1.5.1 DNA, organización y tipos de variantes	7
1.1.5.1.1 Variaciones en número de copias (CNV)	7
1.1.5.1.2 CNV y sus efectos en el metabolismo	8
1.1.5.1.2.1 CNV y sus efectos en el metabolismo en México	8
2. Objetivos	10
2.1 Objetivo general	10
2.2 Objetivos particulares	10
3. Hipótesis	10
4. Pregunta de investigación	10
5. Metodología	11
5.1 Población de estudio	11
5.2 Análisis bioquímico	11
5.3 Criterios de diagnóstico	11
5.4 Estudios genómicos:	12
5.4.1 Detección de CNVs.	12
5.4.2 Análisis en <i>Chromosome Analysis Suite (ChAS)</i>	12
5.4.3 Análisis en <i>PennCNV</i> y <i>CNVruler</i>	14
6. Resultados	15

6.1 Datos generales de la población	15
6.2 Análisis de variantes en número de copias (CNV)	16
7. Discusión	24
7.1 Aneuploidías encontradas en la población	26
8. Conclusiones	27
9. Bibliografía	28

Índice de figuras

Figura 1.	Interfaz gráfica de <i>ChAS</i>	13
Figura 2.	Individuo con complemento cromosómico 47, XXY.	15
Figura 3.	Individuo femenino con ganancia del cromosoma X, 47, XXX.	16
Figura 4.	Número de copias por cromosoma en la población amerindia mexicana.	17

Índice de tablas

Tabla 1.	Criterios para el diagnóstico de síndrome metabólico.	3
Tabla 2.	Factores de riesgo para el desarrollo de síndrome metabólico.	6
Tabla 3.	Variantes en número de copias asociadas a riesgo de síndrome metabólico.	19
Tabla 4.	Variantes en número de copias asociadas a protección de síndrome metabólico.	20
Tabla 5.	Frecuencias de las CNV identificadas en este estudio.	22

Abreviaturas

AHA	Asociación Americana del Corazón
CC	Circunferencia de cintura
C-HDL	Colesterol asociado a lipoproteína de alta densidad
CNV	Variaciones en número de copias
DNA	Acido desoxirribonucleico
DT2	Diabetes tipo 2
ECV	Enfermedad cardiovascular
Glu	Glucosa en ayuno
IDF	Federación Internacional de Diabetes
IL-6	Interleucina 6
NHLBI	Instituto Nacional de Corazón, Pulmón y Sangre
OMS	Organización Mundial de la Salud
PA	Presión arterial
RI	Resistencia a insulina
ROS	Especies reactivas de oxígeno
Smet	Síndrome metabólico
SNS	Sistema nervioso simpático
SNV	Variantes de un solo nucleótido
TA	Tejido adiposo
TG	Triglicéridos
<i>TLR2</i>	Receptor de tipo Toll 2
<i>TLR4</i>	Receptor de tipo Toll 4
<i>TNF-α</i>	Factor de necrosis tumoral alfa

Resumen

El síndrome metabólico (Smet) es una combinación de anomalías metabólicas (resistencia a la insulina, exceso de grasa abdominal, entre otros) que resulta en un marcador de riesgo para el desarrollo de enfermedades cardiovasculares (ECV). El diagnóstico de Smet consiste en la medición de parámetros metabólicos (glucosa en ayuno, colesterol HDL, triglicéridos) y físicos (circunferencia de cintura y presión arterial), donde tres o más componentes anormales representan un diagnóstico positivo de Smet.

El desarrollo de esta entidad metabólica está determinado por diferentes factores incluyendo los genéticos, estos han sido útiles para explicar la complejidad y variabilidad del Smet. Sin embargo, la investigación en esta rama se ha concentrado más hacia variantes de un solo nucleótido (SNV) dejando a un lado otros tipos de variantes genéticas como las variaciones en número de copias (CNV). Debido a lo anterior, el objetivo principal de este trabajo fue buscar una asociación entre las CNV y su asociación con Smet.

Con los datos metabólicos, físicos y de microarreglos de 641 muestras de amerindios mexicanos de diferentes etnias, se realizó una prueba estadística para determinar la asociación entre CNVs, Smet y sus componentes. El llamado de variantes de realizó con el software *PennCNV* y para la asociación estadística *CNVruler*. Además, el software *ChAS* permitió analizar individualmente de manera gráfica el cariotipo de cada sujeto de estudio.

Los resultados mostraron una asociación significativa de las CNV con Smet, donde fue interesante encontrar que algunas de estas variantes contienen en su secuencia genes que podrían estar involucrados en el desarrollo Smet. Por otro lado, en el análisis individual fueron encontrados dos pacientes con aberraciones cromosómicas asociadas al cromosoma sexual X.

1. Introducción

1.1 Síndrome metabólico

El Síndrome Metabólico (Smet) es una combinación de diferentes anormalidades que tienen una estrecha relación con la resistencia a la insulina (RI), depósitos de grasa ectópica (principalmente abdominal) y un grado de inflamación crónica ¹. En la actualidad, el Smet, es considerado un problema de salud pública ya que aproximadamente una cuarta parte de la población mundial lo padece y es un factor de riesgo para diabetes tipo 2 (DT2) y enfermedades cardiovasculares (ECV)².

1.1.1 Diagnóstico

En las dos últimas décadas se han propuesto diferentes definiciones del Smet, las cuales están basadas en parámetros metabólicos y en datos epidemiológicos. Dentro de esas clasificaciones las principales son la de la OMS que tiene como criterio principal, hiperglucemia en sangre y la de la Federación Internacional de Diabetes (IDF) que tiene como componente central la circunferencia de cintura relacionada a grasa abdominal. Estas definiciones fueron unificadas en 2009 y en la actualidad la más utilizada es la propuesta por la AHA-NHLBI (*American Heart Association- National Institute of Heart, Lung and Blood*) estableciendo que la presencia de tres o más componentes es diagnóstico de Smet (Tabla 1) ³.

Tabla 1. Criterios para el diagnóstico de Smet de acuerdo con la AHA-NLBI.

Componente	Punto de corte
Circunferencia de cintura elevada (cm)	>102/>88 hombres/mujeres
Triglicéridos (mg/dL)	≥150
Colesterol HDL (mg/dL)	Mujeres: <50 Hombres: <40
Presión arterial (mmHg)	Sistólica: ≥130 Diastólica: ≥85
Glucosa en ayuno (mg/dL)	≥100

1.1.2 Prevalencia

La prevalencia de Smet varía de una población a otra, ya que la edad, el género y la etnicidad contribuyen a su prevalencia. En Estados Unidos, según la última encuesta nacional de salud y nutrición, en 2012 se reportó un incremento en la prevalencia del Smet, del 25% en 1998 al 34.2% en el 2012 ⁴.

En Latinoamérica, en el 2011, se reportó una prevalencia del 24.9% ⁵ y en México, los últimos reportes mencionan que el 48% de la población mexicana padece Smet ⁶. De acuerdo a la última encuesta nacional de salud y nutrición (ENSANUT) en 2018, el 75.2% de población mexicana mayor a 20 años tiene sobrepeso u obesidad, estrechamente relacionado con circunferencia de cintura elevada, uno de componentes del Smet. El riesgo que esto representa para la población mexicana es alto tomando en cuenta que el Smet es un marcador de riesgo para el desarrollo de ECV y DT2 que son las principales causas de muerte en nuestro país según el Instituto Nacional de Estadística y Geografía (INEGI) en el 2020.

1.1.3 Fisiopatología

El Smet es una enfermedad compleja, en donde la interacción entre los factores genéticos y ambientales contribuyen a su desarrollo y al riesgo para desencadenar DT2 y ECV⁷. El Smet ocurre comúnmente en personas con un índice de masa corporal (IMC) mayor de 30 kg/m²⁸, por lo que el tejido adiposo (TA) juega un papel importante para su desarrollo.

El TA está formado por una mezcla de diferentes células, adipocitos, células estromales, preadipocitos, células del sistema inmune y del endotelio, que no solo contribuye para el desarrollo del Smet, si no también, se ha demostrado tener influencia sobre la hiperglucemia y dislipidemia. En la actualidad, el TA es considerado como un órgano endocrino y parácrino que en situaciones de estrés, causadas por el consumo excesivo de nutrientes, lleva a hipertrofia e hiperplasia de adipocitos y genera una hipoxia promoviendo la liberación de citocinas proinflamatorias, como el IL-6 y el TNF-alfa y la liberación de PAI-1, un inhibidor protrombótico, del activador del plasminógeno⁹⁻¹¹ En presencia de elevados niveles de glucosa (Glu), la insulina promueve la captura de Glu, incrementando la síntesis de glucógeno y disminuyendo la gluconeogénesis. La insulina, es una hormona polipeptídica secretada por las células beta de los islotes pancreáticos de Langerhans y actúa vía receptores de glicoproteínas localizados principalmente en hígado, musculo esquelético y adipocitos. Cuando los tejidos sensibles a la insulina como el TA y el hígado disminuyen su respuesta, se incrementa la cantidad de Glu circulante en plasma conduciendo a una resistencia a la insulina (RI)¹². En el TA, también se incrementa la formación de ácidos grasos libres por la inhibición de la lipólisis. Estos ácidos grasos inhiben la activación de la captura de Glu en el músculo y promueven la gluconeogénesis y lipogénesis en el hígado. Como resultado, se forma un estado hiperinsulinémico para tratar de mantener los niveles de Glu normales en la sangre¹³. Los ácidos grasos libres, también pueden promueven la inflamación, ya que son un ligando para receptores *TLR4* y *TLR2* que participan en la respuesta inmune innata, activando las vías de NFκ-B y JNK que incrementan la producción de quimiocinas promoviendo la inflamación de macrófagos en el TA¹⁴.

Finalmente, también se ha observado una relación entre la hipertensión y los diferentes componentes del Smet ¹⁵⁻¹⁸, ya que la hiperinsulinemia y la hiperglucemia activan el sistema renina-angiotensina incrementando la expresión de Angiotensina II (ATII) y el receptor Angiotensina I (ATI) el cual, en pacientes con RI, pueden contribuir al desarrollo de hipertensión. Existe evidencia de que la RI y la hiperinsulinemia activan el sistema nervioso simpático (SNS), resultando un incremento en la reabsorción de sodio por el riñón, por lo que el corazón aumenta el gasto cardíaco y las arterias responden con una vaso constricción que resulta en hipertensión ¹⁹.

1.1.4 Factores ambientales

El Smet es una entidad multifactorial, donde los factores genéticos y ambientales juegan un papel importante. Los factores ambientales (tabaco, alcohol, actividad física, etc.) (Tabla 2) ²⁰ y el estilo de vida son fundamentales para explicar el desarrollo y las diferentes prevalencias del Smet. Estos factores están fuertemente asociados a un estilo de vida occidental caracterizado por una dieta alta en calorías e inactividad física ²¹.

Se ha identificado como otro posible factor para el desarrollo de Smet a las especies reactivas de oxígeno (ROS). Un exceso de ROS puede causar estrés oxidativo, dando como resultado efectos tóxicos en células y tejidos, mediante el incremento de oxidación en carbohidratos, lípidos y proteínas. De los factores más reconocidos como fuentes de estrés oxidativo están el cigarro y la contaminación ambiental, este último reconocido por la *American Heart Association* (AHA) como un factor de riesgo para el desarrollo de ECV ²².

Tabla 2. Factores de riesgo para el desarrollo de Smet.

Factores de riesgo
Antecedentes familiares
Tabaquismo
Edad
Antecedentes de obesidad
Inactividad física
Estado postmenopáusico
Consumo alto de bebidas azucaradas
Consumo excesivo de alcohol
Dieta "occidental" rica en grasa.

1.1.5 Factores genéticos

Con el propósito de explicar la variabilidad, la complejidad y la relación que existe en el Smet y sus componentes, se han realizados diversos estudios genómicos para tratar de aportar conocimiento sobre las causas de su desarrollo y la variación que existe entre individuos para desencadenar la aparición de esta entidad. De modo que, se han encontrado que factores genéticos, como son las SNV y CNV, están asociadas al desarrollo de síndrome metabólico y/o de alguna enfermedad asociada al metabolismo²³.

1.1.5.1 DNA, organización y tipos de variantes

1.1.5.1.1 Variaciones en número de copias (CNV)

En eucariontes la secuencia de ADN (2'-ácido desoxirribonucleico) está organizada y delimitada dentro del núcleo celular y el máximo nivel de compactación del DNA son los cromosomas. El tamaño del genoma humano haploide es de 3.2×10^9 pares de bases, está organizado en cromosomas lineales con extensión variable de 45 millones a 279 millones de pares de bases. Los genes están localizados en el DNA y en humanos comprenden menos del 2% del genoma codificante²⁴. Todos los humanos compartimos la misma la secuencia de DNA (99.9%), sin embargo existen variantes genéticas que hacen que cada individuo sea fenotípicamente diferente y proporcionan identidad a cada persona; además, estas variantes aportan valiosa información sobre la predisposición a diferentes enfermedades²⁵. Las variantes genéticas se clasifican en variantes de un solo nucleótido (SNV, antes polimorfismos de un solo nucleótido), y variantes estructurales, dentro de este último grupo se encuentran, las variantes en el número de copias (CNV), inserciones, deleciones, translocaciones y aneuploidias²⁶

Es importante menciona, que la gran mayoría de los estudios en variación genética humana se han enfocado en las SNV, y han aportado información importante sobre la susceptibilidad genética a enfermedades comunes. Sin embargo, gracias a la implementación de nuevas tecnologías genómicas y bioinformáticas, se ha encontrado que existen otros tipos de variaciones genéticas, como las CNV, que podrían estar relacionadas a la predisposición de las entidades metabólicas.

1.1.5.1.2 CNV y sus efectos en el metabolismo

Las CNV son variantes estructurales del genoma con tamaños ²⁵ de 50 pb a varias Mb que pueden incrementar los niveles de mRNA y de la proteína correspondiente. Las CNV se pueden clasificar de acuerdo a su frecuencia en raras, que se encuentran en el 1% de la población, y en comunes en más de 1% de la población. En años recientes, los microarreglos han permitido mejorar la comprensión del estudio de las CNV y han arrojado valiosa información de la contribución de estas en la susceptibilidad para desarrollar diferentes enfermedades. Un ejemplo de ello, es una CNV que se encuentra en la región que comprende al gen *LDLR*, la cual induce un fenotipo más severo de hipercolesterolemia, una enfermedad autosómica dominante que afecta a 1:500 personas. ²⁷. Otro estudio, evidenció que CNV en el exón 2 del gen *LEPR* están asociadas con un incremento de Glucosa en ayuno, colesterol total y diabetes tipo 2 ²⁸. En población europea y asiática se ha mostrado, que la CNV en el gen *AMY1* se asocia con el riesgo para desarrollar obesidad ²⁹ y, en población Han del sur de China, una deleción en *LPA* correlaciona con una disminución en el riesgo a enfermedad arterial coronaria, sugiriendo que la presencia de esta variante podría ser un nuevo biomarcador para la prevención de esta entidad ³⁰.

1.1.5.1.2.1 Las CNV y sus efectos en el metabolismo en México

En México, la prevalencia de DT2 (una enfermedad asociada a síndrome metabólico ³¹) es de las más altas a nivel mundial, por lo que se han realizado diversos estudios con la finalidad de profundizar en este padecimiento en nuestra población. En 2017, se encontró que las personas con DT2 tienen una proporción mayor en CNV en comparación con los no-diabéticos. Estas CNV se localizaron en genes con previa asociación a DT2 como *AMY2B*, *LPP*, *ARL15*, *HFE*, *CDKAL1*, *RASGRP1* ³². En niños mexicanos, un incremento en CNV en el gen *ARGHGEF4* se relacionó con obesidad abdominal y, un menor número de copias en el gen *INS* mostró una asociación con incremento en la circunferencia de cintura e IMC³³.

En población indígena mexicana existen pocos reportes. En población infantil maya de Yucatán se han reportado variantes en el gen *FTO* con asociación significativa para sobrepeso únicamente en varones.³⁴ Otro estudio, también en población maya, evidenció que las CNV en el gen *CYP2D6* son proporcionales al fenotipo metabolizador de los sujetos, ya que este gen codifica una enzima importante en el metabolismo de fármacos³⁵. Un estudio similar en *CYP2D6* fue realizado en amerindios tepehuanos para detectar la variabilidad en el número de copias de esa región.³⁶

En el gen *AMY1*, las CNV fueron asociadas con obesidad en niños y adultos mestizos mexicanos y en población indígena. Aún no se sabe el mecanismo por el cual esta variante podría estar asociada con obesidad, pero se ha propuesto que pudiese estar implicado con la selección natural por el aumento a una dieta alta en almidón³⁷. Por lo anterior, es necesario incrementar el conocimiento de cómo las CNV están correlacionadas con entidades metabólicas en población mexicana, incluyendo indígenas y mestizos.

2. Objetivos

2.1 Objetivo General

Identificar variantes en el número de copias (CNV) y su asociación con síndrome metabólico en población amerindia mexicana.

2.2 Objetivos Específicos

1. Identificar la presencia de CNV en población amerindia mexicana
2. Conocer la frecuencia de CNV en la población amerindia y compararla con las diferentes poblaciones
3. Establecer la asociación de las CNV identificadas con el síndrome metabólico.
4. Realizar la búsqueda en literatura de variantes genéticas y genes contenidos en las CNV identificadas con asociación a síndrome metabólico.

3. Hipótesis

Si existen CNV en población indígena mexicana, entonces es probable que estén asociadas con la presencia del síndrome metabólico.

4. Pregunta de investigación

¿Existen CNV asociadas al síndrome metabólico en población indígena mexicana?

5. Metodología

5.1 Población de estudio

Se incluyeron 641 indígenas mexicanos pertenecientes a la Cohorte MAIS (Análisis metabólico en una muestra indígena). Las muestras fueron colectadas en 73 comunidades indígenas y los criterios de inclusión fueron: individuos que se auto reconocieron como indígenas, que nacieran en la misma comunidad que sus padres y abuelos, por último, que sus abuelos y padres hablaran alguna lengua nativa³⁵. Se excluyeron a todos los individuos que no aceptaron firmar la carta de consentimiento y que sus padres y abuelos no hablaran una lengua indígena. Los individuos con historia clínica y pruebas bioquímicas incompletas fueron eliminados.

5.2 Análisis bioquímico

Se colectaron muestras de sangre en ayuna de cada participante. Esta muestra fue utilizada para los estudios bioquímicos: glucosa, colesterol HDL (C-HDL) y triglicéridos y para el análisis genómico. También se les realizó una historia clínica completa, en la cual se midió circunferencia de cintura entre la última costilla y la cresta iliaca y la presión arterial (PA) siguiendo la norma NOM-030-SSA2- 1999.

5.3 Criterios de diagnóstico

Para realizar el diagnóstico de Smet se siguieron los criterios de la AHA (*American Heart Association*) y del NHLBI (*National Heart, Lung and Blood Institute*) (Tabla 1). Un individuo se diagnosticó con Smet (caso) cuando presentó tres o más componentes metabólicos. Dentro de la población, 300 pacientes tuvieron diagnóstico positivo a Smet, mientras que los 341 individuos restantes se tomaron como controles para este estudio.

5.4 Estudios genómicos

5.4.1 Detección de CNVs

Se realizó el análisis de las CNV a partir de datos generados por un microarreglo tipo *Genome Wide Human SNP array 6.0*, que contiene ~906, 600 SNVs y ~946, 000 sondas para detectar CNV. Estos marcadores están distribuidos por todo el genoma humano y fueron elegidos de acuerdo a la base de datos del *HapMap Project*. Este microarreglo fue proporcionado por el laboratorio de Inmunogenómica y Enfermedades Metabólicas del INMEGEN.

5.4.2 Análisis en ChAS (Chromosome Analysis Suite)

Los archivos .CEL derivados de los microarreglos fueron convertidos por el *GeneChip Command Console Software* a un formato CNCHP para ser analizados con el programa *ChAS* y con el algoritmo CN5 de *Affymetrix*. *ChAS* es una herramienta diseñada para detectar y analizar de manera fácil, información relevante sobre alteraciones cromosómicas a partir de información obtenida de microarreglos, de modo que, permite visualizar gráficamente las aberraciones, ganancias y deleciones de cada muestra (Figura 1).



Figura 1. Interfaz gráfica de *ChAS* y representación de los cromosomas. Los segmentos azules y rojos simbolizan regiones de ganancias y pérdidas, respectivamente.

5.4.3 Análisis en *PennCNV* y *CNVruler*

El llamado de variantes para el análisis de las CNV a partir también de los archivos .CEL se realizó con el algoritmo *PennCNV* ³⁸. El programa utiliza el algoritmo *Birdseed* que produce la conversión de la señal que se encuentra en el archivo .CEL para finalmente obtener CNV. Este algoritmo genera un archivo que divide las señales individuales y realizar el llamado de las CNV utilizando los valores de *Log R Ratio (LRR)* y *B Allele Frequency (BAF)* para cada variante. Finalmente, el llamado de variantes fue realizado con la versión del genoma hg18.

Debido a que las CNV pueden presentarse en distintos tamaños y estados (número de copias), su estudio resulta más complejo en comparación a otro tipo de variaciones. *CNVruler* provee un algoritmo capaz de realizar la clasificación y agrupamiento de las CNV encontradas en los diferentes individuos dependiendo de su tamaño y posición en el genoma, este agrupamiento produce regiones de CNV (CNVR). Este algoritmo también permitió realizar el análisis de casos (n = 300) y controles (n = 341) de los datos clínicos con una prueba de regresión logística con un *minor threshold* de 0.05 para el síndrome metabólico ³⁹.

De los segmentos obtenidos con significancia estadística se realizó una búsqueda dirigida en las bases de datos (*DGV* y *DECIPHER*) para conocer, si las variantes encontradas contaban con reporte previo. Debido a que *DECIPHER* no cuenta con la construcción del genoma hg 18, se utilizó la herramienta *liftOver* del UCSC <https://genome.ucsc.edu/cgi-bin/hgLiftOver> para transpolar una secuencia de una construcción a otra más reciente.

6. Resultados

6.1 Datos generales de la población

Del total de muestras analizadas (641), el 47% (300) de los individuos presentaron Smet. El componente más frecuente fueron los valores bajos de colesterol HDL (69%), seguido de los niveles altos de triglicéridos (60%) y el menos frecuente los niveles elevados de glucosa (34%).

En nuestra población se encontraron dos individuos con un cromosoma sexual extra, uno de ellos con 47, XXY perteneciente a la etnia tarahumara y el otro 47, XXX de la etnia otomí (Figuras 2-3). El resto de la población mostró en complemento cromosómico 46, XX o 46, XY según el caso.

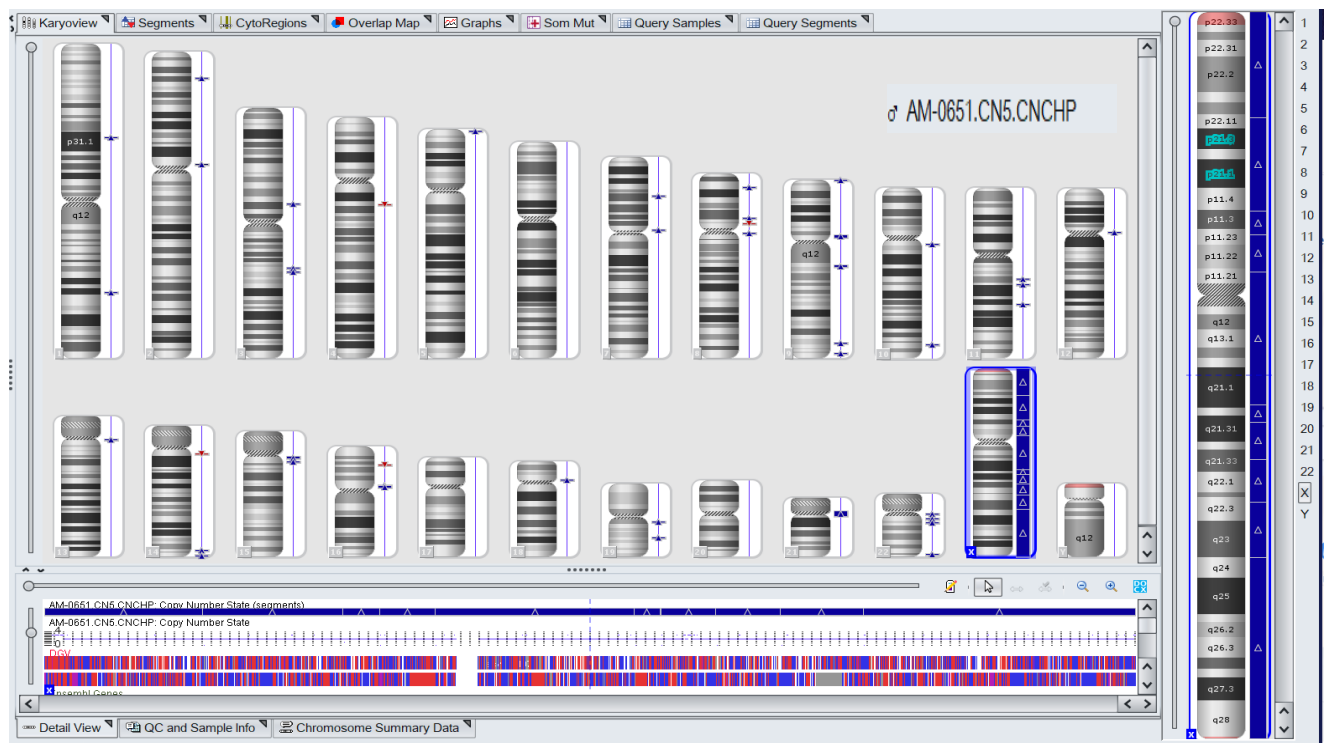


Figura 2. Individuo con complemento cromosómico 47, XXY.

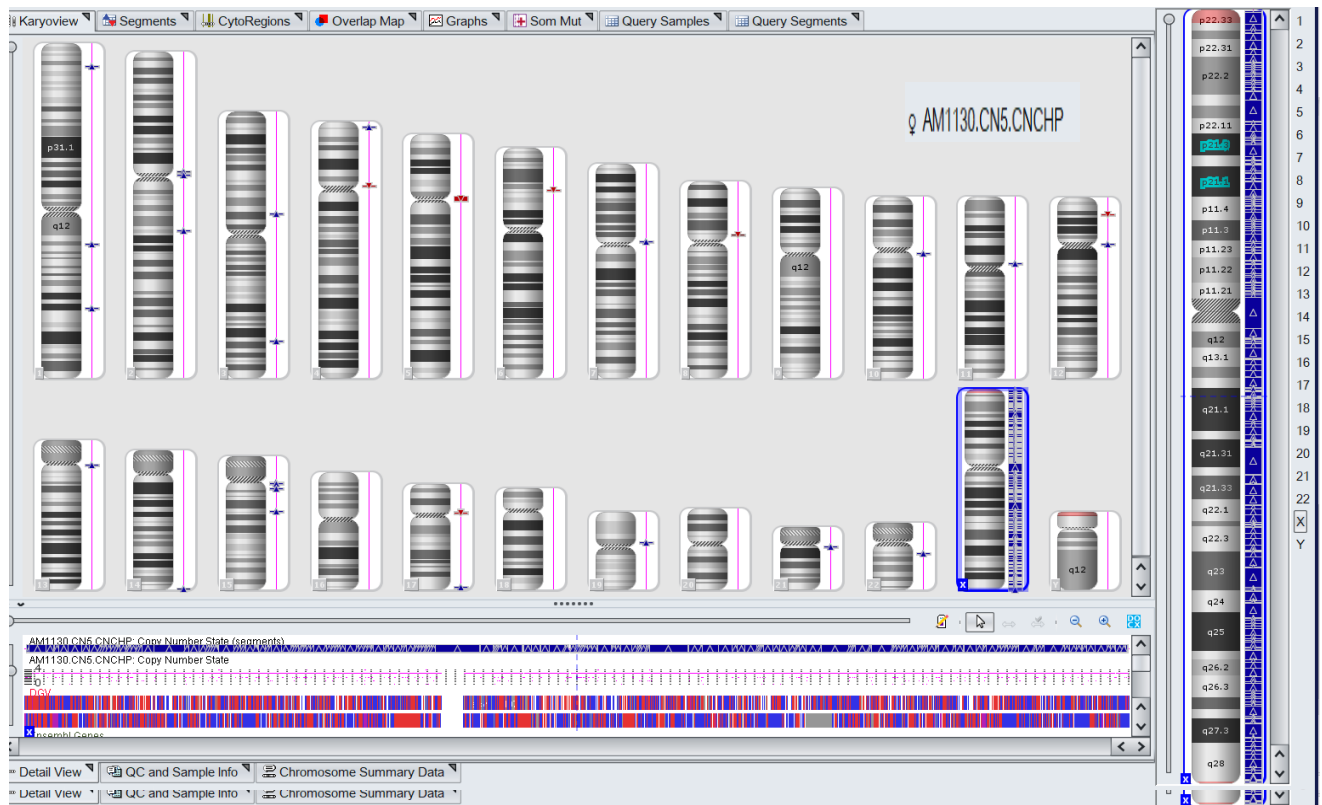


Figura 3. Individuo femenino con ganancia del cromosoma X, 47, XXX

6.2 Análisis en variantes en número de copias

Se detectaron más de 35 mil variantes en número de copias en la población, con un valor promedio de 55 variantes por persona. Interesantemente, se observó que en la población indígena existe más pérdidas que ganancias genómicas (Figura 4).

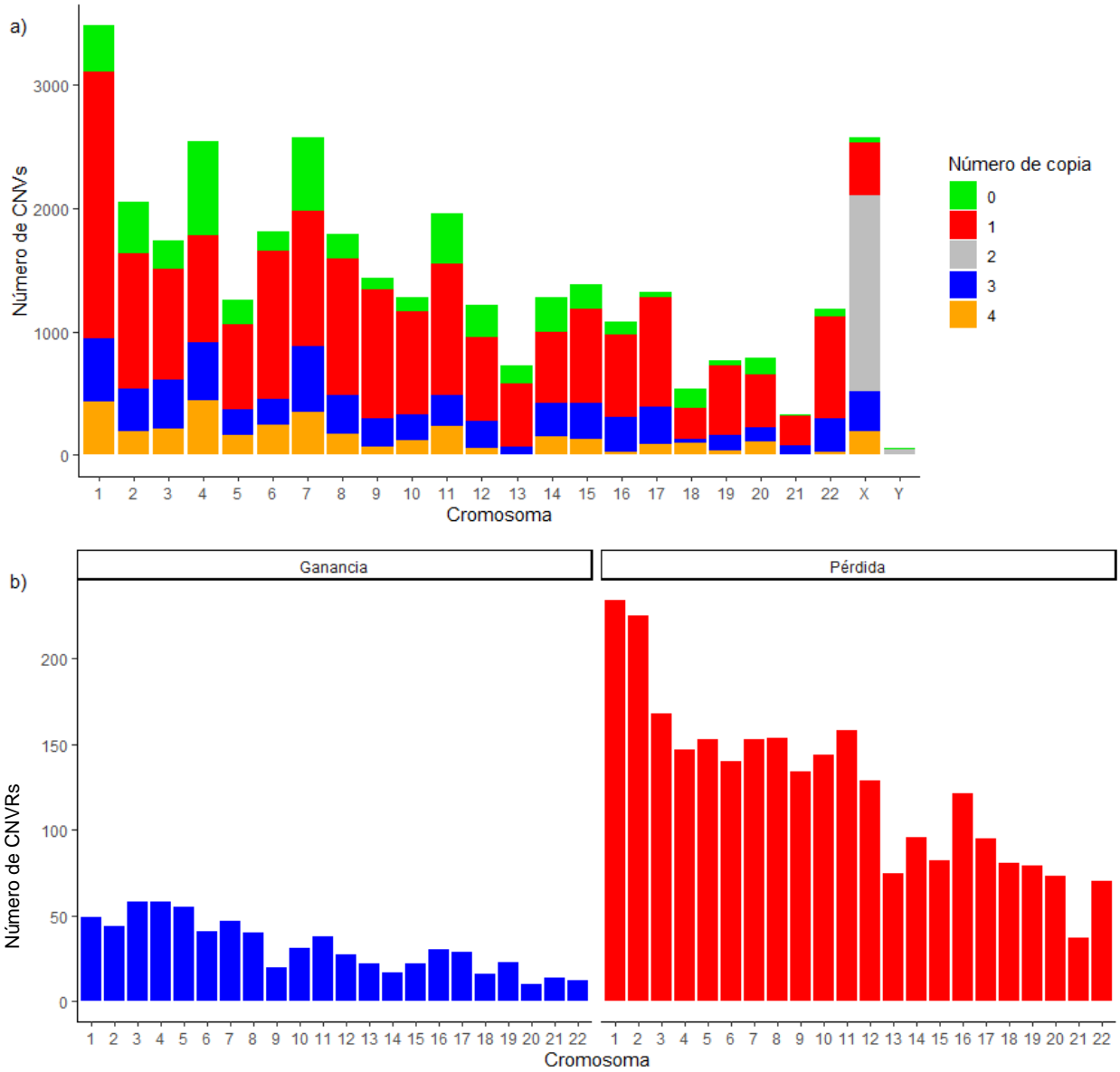


Figura 4. a) Distribución de las CNV totales identificados por cromosoma en la población amerindia mexicana. b) Número de copias por cromosoma en la población amerindia mexicana.

El análisis de asociación mostro que seis variantes presentaron un riesgo significativo para desarrollar Smet (Tabla 3). 5/6 variantes mostraron una pérdida de CNV y se encontraron en los cromosomas 1p, 4p, 7q 9p, 14q. Por otro lado, observamos una sola variante con ganancia en el cromosoma 2q.

Este análisis permitió observar que existen 11 CNV asociadas a protección a Smet (Tabla 4). Todas las variantes corresponden a deleciones en los cromosomas 2q, 4p, 8q, 10q, 15q, 16p, 19p, 20q y 22q.

Interesantemente, cuando comparamos las frecuencias de las CNV encontradas en la población indígena con lo reportado previamente, observamos que, en todos los casos, la población indígena mexicana mostró la mayor frecuencia en relación a otros estudios (Tabla 5).

Tabla 3. Variantes en número de copias asociadas a riesgo de desarrollo de síndrome metabólico (n = 641)

ID CNV	Citobanda	Coordenadas	Tipo de CNV	Tamaño (kb)	Genes	P	OR (IC95%)
CNVR_P106	1p13.2	112,397,436-112,599,414	Delección	202	-	1×10^{-3}	1.85(1.27-2.68)
CNVR_G83	2q22.3	146,574,839-147,410,802	Ganancia	836		4×10^{-3}	1.80(1.20-2.25)
CNVR_P2109	14q32.33	105,592,904-106,014,879	Delección	422	-	2×10^{-2}	1.62(1.07- 2.45)
CNVR_P648	4p16.1	9,639,892-10,027,768	Delección	388	<i>SLC2A9, WDR1</i>	9×10^{-3}	1.54(1.11-2.14)
RO_2820	9p11.2	45,805,240-45,931,125	Delección	126	-	3×10^{-2}	1.75(1.04-2.95)
RO_2354	7q35	142,989,141-143,083,821	Delección	95	<i>TCAF2</i>	4×10^{-2}	1.70(1.02-2.84)

Tabla 4. Variantes en número de copias asociadas a protección de síndrome metabólico (n = 641)

ID CNV	Citobanda	Coordenadas	Tipo de CNV	Tamaño (kb)	Genes	P	OR (IC95%)
CNVR_P 2638	20q13.33	60,116,677-60,943,041	Delección	826	<i>PPP1R3D, CDH26, CDH4, TAF4, LSM14B, PSMA7, SS18L1, MTG2, HRH3, OSBPL2, ADRM1, LAMA5, RPS21, CABLES2, RBBP8NL, GATA5, SLCO4A1, NTSR1, MRGBP, OGFR, COL9A3</i>	4×10^{-4}	0.30(0.15-0.58)
CNVR_P 2741	22q13.32	47,179,043-47,543,599	Delección	364	<i>FAM19A5</i>	1×10^{-3}	0.35(0.19-0.66)
CNVR_P 2194	16p13.3	240,642-1,815,767	Delección	1575	<i>AX1N1, LMF1, SSTR5, CACNA1H, GNPTG, CLCN7</i>	5×10^{-3}	0.45(0.26-0.78)
CNVR_P 2684	22q11.21	17,947,776-18,992,706	Delección	1045	<i>GP1BB, TBX1, COMT, RTN4R</i>	5×10^{-3}	0.47(0.27-0.80)

CNVR_P 643	4p16.1	7,464,168- 7,868,785	Delección	407	<i>SORCS2, AFAP1, PSPAL1</i>	2x10 ⁻³	0.36(0.19-0.69)
CNVR_P 2510	19p13.12	15,124,385- 15,645,152	Delección	521	<i>NOTCH3, CYP4F22</i>	5x10 ⁻³	0.40(0.20-0.76)
CNVR_P 1370	8q24.3	142,274,419- 142,821,687	Delección	547	<i>DENND3, SLC45A4, MROH5, PTP4A3</i>	1x10 ⁻²	0.49(0.27-0.86)
CNVR_P 1551	10q11.21	42,631,395- 43,089,910	Delección	458	<i>BMS1, RET, CSGALNA CT2, RASGEF1 A</i>	2x10 ⁻²	0.48(0.26-0.88)
CNVR_P 457	2q37.3	240,749,808- 241,847,112	Delección	1097	<i>CAPN10, KIF1A, AGXT, PPP1R7</i>	3x10 ⁻²	0.50(0.28-0.94)
CNVR_P 2735	22q13.31	45,041,218- 45,250,436	Delección	209	<i>TTC38, GTSE1, TRMU, CELSR1</i>	3x10 ⁻²	0.50(0.26-0.96)
CNVR_P 2161	15q24.3	75,573,958- 76,038,320	Delección	464	<i>LINGO1</i>	4x10 ⁻²	0.50(0.26-0.96)

Tabla 5. Frecuencias en otras poblaciones de las CNV identificadas en este estudio.

Citobanda	Frecuencia en este estudio (%)	Frecuencia en otra población (%)	Referencia
1p13.2	26.37	7.50	Conrad (2010) ⁴⁰
2q22.3	20.44	1.05	Wong (2007) ⁴¹
2q37.3	7.64	0.90	Jakobsson (2008) ⁴²
4p16.1_riesgo	44.46	0.78	Pinto (2007) ⁴³
4p16.1_protección	7.33	0.38	Locke (2006) ⁴⁴
7q35	11.70	1.60	Campbell (2011) ⁴⁵
8q24.3	8.89	1.12	Jakobsson (2008) ⁴²
9p11.2	11.23	0.99	Vogler (2010) ⁴⁶
10q11.21	7.96	1.85	Redon (2006) ⁴⁷
14q32.33	18.88	4.90	McCarroll (2006) ⁴⁸
15q24.3	6.55	1.05	Wong (2007) ⁴¹
16p13.3	9.52	4.30	Jakobsson (2008) ⁴²
19p13.12	7.02	1.05	Wong (2007) ⁴¹

20q13.33	7.333	4.51	Jakobsson (2008) ⁴²
22q11.21	10.45	3.22	Park (2010) ⁴⁹
22q13.31	6.40	0.23	Jakobsson (2008) ⁴²
22q13.32	7.64	1.05	Wong (2007) ⁴¹

7. Discusión

Como parte principal de este trabajo, se realizó un estudio tipo caso-control en población amerindia mexicana para encontrar una relación causal entre la presencia de CNV y el desarrollo de síndrome metabólico. Los resultados mostraron 17 variantes asociadas con la protección o con el riesgo para desarrollar Smet.

Los resultados de asociación mostraron 17 *loci* con asociación significativa a Smet. De acuerdo con una búsqueda intencionada en las diferentes bases de datos (DGV), estas variantes tienen reportes previos, es decir, no son CNV nuevas. Esta búsqueda intencionada mostró que la mayoría de los reportes se realizaron para la obtención de variantes en poblaciones controles, por ejemplo, los estudios hechos por Redon⁴⁵, Locke⁴², y Campbell⁴³ utilizaron la información obtenida del HapMap Project, para realizar una búsqueda de variantes estructurales que podrían ser comunes en el genoma humano.

Por otro lado, la comparación de las frecuencias de las CNV identificadas en nuestra población evidenció una frecuencia mayor en relación a lo previamente reportado ³⁸⁻⁴⁶. Esto podría deberse a la metodología usada para realizar el traslape de las variantes en la población, ya que una CNVR comprende una región donde se encuentran inmersas otras CNV. En cambio, en la base de datos DGV se reporta la frecuencia para una determinada variante, pero se ignoran todas las otras CNVs que se encuentran en la misma región.

En este estudio se encontraron 6 CNV asociadas a tener un riesgo para desarrollar Smet. Nuestros resultados concuerdan con la literatura ya que en 4p16.1 también se ha reportado una deleción que contiene a los genes *WDR1* y *SLC2A9* previamente asociados con el aumento de niveles de ácido úrico y Smet en población china de Taiwán ⁵⁵.

En 7q35 fue identificada una CNV que en su secuencia está comprendido el gen *TCAF2*. Interesantemente, esta región ha sido reportada como variable, es decir, se han identificado duplicaciones segmentarias (DS) que han llevado a la formación de

reestructuraciones del *locus* en diferentes poblaciones y que son considerados cambios evolutivos ⁵⁶. También se sabe, que las DS son puntos de partida para la recombinación no homologa, un mecanismo de formación de CNVs ⁵⁷. Aunque no se encontraron asociaciones de esta región con alguna entidad metabólica en la literatura, este podría ser el primer reporte en población indígena mexicana de una delección, producto de un cambio evolutivo que además podría conferir riesgo al desarrollo de Smet.

En la CNV identificada en 2q22.3 se encuentran también las SNV rs222826 y rs222827, las cuales han sido previamente asociadas en población de Estados Unidos de América ⁵⁸ con niveles bajos de HDL-C y niveles altos de triglicéridos (parámetros de diagnóstico del Smet), además con el desarrollo de otras enfermedades, como lo son enfermedad coronaria, paro cardíaco, diabetes y cáncer. Las SNV rs222826 y rs222827 se encuentran en una región desierta de genes en el genoma, lo que sugiere que este tipo de regiones tienen secuencias regulatorias involucradas en la expresión de genes cercanos como los son *ZEB2* y *ACVR2A* que participan en vías relacionadas con la inflamación, envejecimiento, carcinogénesis, entre otras funciones biológicas.

Otra CNV para riesgo en Smet fue identificada en 1p13.2, la región intergénica que comprende esta CNV ha sido asociada previamente con presión arterial extremadamente alta en población europea ⁵⁹, otro componente de Smet.

Por otra parte, se encontraron 11 variantes significativas asociadas a protección a Smet. Una de las CNVRs de mayor interés fue una delección en 22q13.32, la cual involucra al gen *FAM19A5*, este gen expresa para una nueva adipocina, la FAM19A5 y, se ha demostrado que una regulación negativa de *FAM19A5* podría relacionarse con el desarrollo de enfermedades cardiometabólicas ⁶⁰. Estos datos no concuerdan con lo observado en nuestro trabajo, ya que nosotros encontramos una pérdida de esta región, por lo que es probable que exista una menor cantidad de proteína expresada, en este sentido, son necesarios más estudios para profundizar en la función de esta adipocina.

Finalmente, la variante asociada a protección en 22q11.21 contiene al gen *TBX1* que de acuerdo a reportes previos la expresión de este es requerido para la regulación de la señalización de insulina en tejido adiposo⁶¹. Lo anterior es una observación importante

porque de acuerdo con la etiología del Smet la resistencia a la insulina juega un papel crucial en su desarrollo. Por lo que, se sugiere que la existencia de una pérdida en la región que codifica a *TBX1* esté relacionada con la protección a síndrome metabólico.

7.1 Aneuploidías encontradas en la población

Las aneuploidías están definidas como un cambio en el número del complemento cromosómico normal, es decir, pueden estar presentes o deletados 1,2 ...22 cromosomas y, dependiendo el tipo de cromosoma, serán las características clínicas de los pacientes. Las aneuploidías más comunes son el síndrome de Down (47, XX+21 o 47, XY+21) y el síndrome de Turner 45, X⁵⁰.

Interesantemente, en este trabajo se encontró que, en 641 individuos indígenas aparentemente sanos, dos individuos presentaron aneuploidías de los cromosomas sexuales (47, XXY y 47, XXX). En hombres, el síndrome de Klinefelter (SK) o 47, XXY ocurre en aproximadamente 1 de cada 600-650 individuos⁵¹. El fenotipo en estos pacientes es muy variable, en general cursan con estatura alta, hipogonadismo, infertilidad o disfunción sexual, así como, déficit cognitivos y de lenguaje⁵². En relación con la infertilidad, se ha reportado que la mayoría de los individuos 47,XXY nacen con espermatogonia y la pérdida de las células espermatogoniales se da progresivamente en la pubertad⁵³ y el 90% de los pacientes cursa con azoospermia. Profundizando en el historial clínico de este paciente, él pertenece a la etnia tarahumara, tiene 57 años con una estatura de 1.62cm y reportó que tiene dos hijos, no hay más datos clínicos relacionados a SK. Es de llamar la atención, la presencia de dos hijos y de acuerdo a la literatura, únicamente los individuos con mosaicos del SK pueden tener descendencia, por lo que es muy probable, que este paciente sea un mosaico SK, sin embargo, habría que hacer dirigidamente un cariotipo para confirmar la presencia de las dos líneas celulares. La otra alternativa es que sea un caso de no paternidad. En relación con la paciente 47, XXX es una mujer de 57 años, de la etnia otomí que sí reportó en su historia clínica infertilidad, lo que coincide con la literatura⁵⁴.

8. Conclusiones

1. Se observó una frecuencia 0.3% de aneuploidías en una población indígenas aparentemente sana.
2. Se identificaron 17 variantes asociadas a Smet en población indígena mexicana
3. Se identificó, por primera vez, una delección CNV en 7q35 en población indígena mexicana producto de un cambio evolutivo y que confiere un riesgo al desarrollo de Smet.
4. Se identificó la participación de FAM19A5 una nueva adipocina involucrada en el desarrollo de Smet y enfermedades cardio metabólicas, aunque es necesario realizar más estudios al respecto.
5. Se identificó una variante asociada a protección a Smet que contiene al gen *TBX1* relacionado con la regulación de la señalización de insulina en tejido adiposo.

9. Bibliografía

1. Aguilar-salinas, C. A. & Viveros-ruiz, T. Recent advances in managing / understanding the metabolic syndrome [version 1 ; peer review : 3 approved] Referee Status : **8**, (2019).
2. Swarup S, Goyal A, Grigороva Y, Z. R. Metabolic Syndrome. *StatPearls Publ.* (2020).
3. Alberti, K. G. M. M. *et al.* Harmonizing the Metabolic Syndrome. *Circulation* **120**, 1640–1645 (2009).
4. Moore, J. X., Chaudhary, N. & Akinyemiju, T. Metabolic syndrome prevalence by race/ ethnicity and sex in the united states, national health and nutrition examination survey, 1988-2012. *Prev. Chronic Dis.* **14**, 1–16 (2017).
5. Márquez-Sandoval, F. *et al.* The prevalence of metabolic syndrome in Latin America: A systematic review. *Public Health Nutr.* **14**, 1702–1713 (2011).
6. Gutiérrez-Solis, A. L., Datta Banik, S. & Méndez-González, R. M. Prevalence of Metabolic Syndrome in Mexico: A Systematic Review and Meta-Analysis. *Metab. Syndr. Relat. Disord.* **16**, 395–405 (2018).
7. De Edelenyi, F. S. *et al.* Prediction of the metabolic syndrome status based on dietary and genetic parameters, using Random Forest. *Genes Nutr.* (2008) doi:10.1007/s12263-008-0097-y.
8. Grundy, S. M. Metabolic syndrome update. *Trends Cardiovasc. Med.* **26**, 364–373 (2016).
9. Catrysse, L. & van Loo, G. Inflammation and the Metabolic Syndrome: The Tissue-Specific Functions of NF-κB. *Trends Cell Biol.* **27**, 417–429 (2017).
10. Oda, E. Metabolic syndrome: Its history, mechanisms, and limitations. *Acta Diabetol.* **49**, 89–95 (2012).

11. Kaur, J. Retracted: A Comprehensive Review on Metabolic Syndrome. *Cardiol. Res. Pract.* **2019**, 1–1 (2019).
12. Xu, H., Li, X., Adams, H., Kubena, K. & Guo, S. Etiology of metabolic syndrome and dietary intervention. *Int. J. Mol. Sci.* **20**, 1–19 (2019).
13. Rochlani, Y., Pothineni, N. V., Kovelamudi, S. & Mehta, J. L. Metabolic syndrome: Pathophysiology, management, and modulation by natural compounds. *Ther. Adv. Cardiovasc. Dis.* **11**, 215–225 (2017).
14. Saltiel, A. R., Olefsky, J. M., Saltiel, A. R. & Olefsky, J. M. Inflammatory mechanisms linking obesity and metabolic disease Find the latest version : Inflammatory mechanisms linking obesity and metabolic disease. *J. Clin. Invest.* **127**, 1–4 (2017).
15. Morse, Stephen A.;Zhang, Rubin;Thakur, Vashu; Reisin, E. Hypertension and the metabolic syndrome. *Am. J. Med. Sci.* **6**, 303–310 (2005).
16. Lusis, A. J., Attie, A. D. & Reue, K. Metabolic syndrome: From epidemiology to systems biology. *Nat. Rev. Genet.* **9**, 819–830 (2008).
17. Matfin, G. The metabolic syndrome: What's in a name? *Ther. Adv. Endocrinol. Metab.* **1**, 39–45 (2010).
18. Johansen, C., Kathiresan, S. & Hegele, R. Genetic determinants of plasma triglycerides. *J. Lipid Res.* (2011) doi:10.1194/jlr.R009720.
19. Malhotra, A., Kang, B. P. S., Cheung, S., Opawumi, D. & Meggs, L. G. Angiotensin II promotes glucose-induced activation of cardiac protein kinase C isozymes and phosphorylation of troponin I. *Diabetes* (2001) doi:10.2337/diabetes.50.8.1918.
20. McCracken, E., Monaghan, M. & Sreenivasan, S. Pathophysiology of the metabolic syndrome. *Clin. Dermatol.* **36**, 14–20 (2018).
21. Mendrick, D. L. *et al.* Metabolic Syndrome and Associated Diseases: From the Bench to the Clinic. *Toxicol. Sci.* **162**, 36–42 (2018).

22. Hutcheson, R. & Rocic, P. The metabolic syndrome, oxidative stress, environment, and cardiovascular disease: The great exploration. *Exp. Diabetes Res.* **2012**, (2012).
23. Nazarenko, M. S. *et al.* Genomic structural variations for cardiovascular and metabolic comorbidity. *Sci. Rep.* **7**, 1–9 (2017).
24. Krebs, J. E., Goldstein, E. S. & Kilpatrick, S. T. *Lewin's Genes XII*. (Jones & Bartlett Learning, 2018).
25. Feuk, L., Carson, A. R. & Scherer, S. W. Structural variation in the human genome. *Nat. Rev. Genet.* **7**, 85–97 (2006).
26. Eichler, E. E. Genetic Variation, Comparative Genomics, and the Diagnosis of Disease. *N. Engl. J. Med.* **381**, 64–74 (2019).
27. Pollex, R. L. & Hegele, R. A. Genomic copy number variation and its potential role in lipoprotein and metabolic phenotypes. *Curr. Opin. Lipidol.* **18**, 174–180 (2007).
28. Jeon, J. P. *et al.* Copy number variation at leptin receptor gene locus associated with metabolic traits and the risk of type 2 diabetes mellitus. *BMC Genomics* **11**, (2010).
29. Falchi, M. *et al.* Low copy number of the salivary amylase gene predisposes to obesity. *Nat. Genet.* **46**, 492–497 (2014).
30. Wu, Z. *et al.* Copy number variation of the Lipoprotein(a) (LPA) gene is associated with coronary artery disease in a southern Han Chinese population. *Int. J. Clin. Exp. Med.* **7**, 3669–3677 (2014).
31. Punthakee, Z., Goldenberg, R. & Katz, P. Definition, Classification and Diagnosis of Diabetes, Prediabetes and Metabolic Syndrome. *Can. J. Diabetes* **42**, S10–S15 (2018).
32. Ascencio-Montiel, I. D. J. *et al.* Characterization of large copy number variation in Mexican type 2 diabetes subjects. *Sci. Rep.* **7**, 1–7 (2017).

33. Antúñez-Ortiz, D. L. *et al.* Copy Number Variations in Candidate Genes and Intergenic Regions Affect Body Mass Index and Abdominal Obesity in Mexican Children. *Biomed Res. Int.* **2017**, (2017).
34. González-Herrera, L. *et al.* Genetic variation of FTO: rs1421085 T>C, rs8057044 G>A, rs9939609 T>A, and copy number (CNV) in Mexican Mayan school-aged children with obesity/overweight and with normal weight. *Am. J. Hum. Biol.* **31**, 1–13 (2019).
35. López-López, M. *et al.* CYP2D6 genetic polymorphisms in Southern Mexican Mayan Lacandones and Mestizos from Chiapas. *Pharmacogenomics* **15**, 1859–1865 (2014).
36. Sosa-Macías, M., Dorado, P., Alanis-Bañuelos, R. E., Llerena, A. & Lares-Asseff, I. Influence of CYP2D6 deletion, multiplication, -1584C→G, 31G→A and 2988G→A gene polymorphisms on dextromethorphan metabolism among Mexican tepehuanos and mestizos. *Pharmacology* **86**, 30–36 (2010).
37. León-Mimila, P. *et al.* Low Salivary Amylase Gene (AMY1) Copy Number Is Associated with Obesity and Gut Prevotella Abundance in Mexican Children and Adults. *Nutrients* **10**, 1–12 (2018).
38. Li, M. *et al.* PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* **17**, 1665–1674 (2007).
39. Kim, J. H. *et al.* CNVRuler: A copy number variation-based case-control association analysis tool. *Bioinformatics* **28**, 1790–1792 (2012).
40. Conrad, D. F. *et al.* Origins and functional impact of copy number variation in the human genome. *Nature* **464**, 704–712 (2010).
41. Wong, K. K. *et al.* A Comprehensive Analysis of Common Copy-Number Variations in the Human Genome. *Am. J. Hum. Genet.* **80**, 91–104 (2007).
42. Jakobsson, M. *et al.* Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* **451**, 998–1003 (2008).

43. Pinto, D., Marshall, C., Feuk, L. & Scherer, S. W. Copy-number variation in control population cohorts. *Hum. Mol. Genet.* **16**, 168–173 (2007).
44. Locke, D. P. *et al.* Linkage Disequilibrium and Heritability of Copy-Number Polymorphisms within Duplicated Regions of the Human Genome. *Am. J. Hum. Genet.* **79**, 275–290 (2006).
45. Campbell, C. D. *et al.* Population-genetic properties of differentiated human copy-number polymorphisms. *Am. J. Hum. Genet.* **88**, 317–332 (2011).
46. Vogler, C. *et al.* Microarray-based maps of copy-number variant regions in European and Sub-Saharan populations. *PLoS One* **5**, 1–8 (2010).
47. Redon, R. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).
48. McCarroll, S. A. *et al.* Common deletion polymorphisms in the human genome. *Nat. Genet.* **38**, 86–92 (2006).
49. Park, H. *et al.* Discovery of common Asian copy number variants using integrated high-resolution array CGH and massively parallel DNA sequencing. *Nat. Genet.* **42**, 400–405 (2010).
50. Ricke, R. M. & Van Deursen, J. M. Aneuploidy in health, disease, and aging. *J. Cell Biol.* **201**, 11–21 (2013).
51. Vockel, M., Riera-Escamilla, A., Tüttelmann, F. & Krausz, C. The X chromosome and male infertility. *Hum. Genet.* **1996**, (2019).
52. Boada, R., Janusz, J., Hutaff-Lee, C. & Tartaglia, N. The cognitive phenotype in Klinefelter syndrome: A review of the literature including genetic and hormonal factors. *Dev. Disabil. Res. Rev.* **15**, 284–294 (2009).
53. Paduch, D. A., Fine, R. G., Bolyakov, A. & Kiper, J. New concepts in Klinefelter syndrome. (2008) doi:10.1097/MOU.0b013e32831367c7.

54. Rafique, M., AlObaid, S. & Al-Jaroudi, D. 47, XXX syndrome with infertility, premature ovarian insufficiency, and streak ovaries. *Clin. Case Reports* **7**, 1238–1241 (2019).
55. Biradar, M. I., Chiang, K. M., Yang, H. C., Huang, Y. T. & Pan, W. H. The causal role of elevated uric acid and waist circumference on the risk of metabolic syndrome components. *Int. J. Obes.* **44**, 865–874 (2020).
56. Dennis, M. Y. *et al.* The evolution and population diversity of human-specific segmental duplications. *Nat. Ecol. Evol.* **1**, 1–10 (2017).
57. Arlt, M. F., Wilson, T. E. & Glover, T. W. Replication stress and mechanisms of CNV formation. *Curr. Opin. Genet. Dev.* **22**, 204–210 (2012).
58. Kulminski, A. M. *et al.* Pleiotropic Associations of Allelic Variants in a 2q22 Region with Risks of Major Human Diseases and Mortality. *PLoS Genet.* **12**, 1–20 (2016).
59. Marques, F. Z. *et al.* Measurement of absolute copy number variation reveals association with essential hypertension. *BMC Med. Genomics* **7**, 1–8 (2014).
60. Wang, Y. *et al.* Novel adipokine, FAM19A5, inhibits neointima formation after injury through sphingosine-1-phosphate receptor 2. *Circulation* **138**, 48–63 (2018).
61. Markan, K. R. *et al.* Adipose TBX1 regulates β -adrenergic sensitivity in subcutaneous adipose tissue and thermogenic capacity in vivo. *Mol. Metab.* **36**, 1–13 (2020).