



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE CIENCIAS
BIOLOGÍA EXPERIMENTAL

**Análisis de DNA mitocondrial en la población mestiza de la
Ciudad de México**

T E S I S

QUE PARA OPTAR EL GRADO DE:
MAESTRA EN CIENCIAS BIOLÓGICAS

PRESENTA:
BIÓL. CAROLINA LEÓN CAMPOS

TUTOR PRINCIPAL DE TESIS: DRA. AMÉRICA NITXIN CASTAÑEDA SORTIBRÁN
FACULTAD DE CIENCIAS, UNAM

COMITÉ TUTOR: DR. HÉCTOR RANGEL VILLALOBOS
CENTRO UNIVERSITARIO DE LA CIÉNEGA, UDG

COMITÉ TUTOR: DR. JOSÉ MARIANO GURADADO ESTRADA
FACULTAD DE MEDICINA, UNAM



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.



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE CIENCIAS
BIOLOGÍA EXPERIMENTAL

**Análisis de DNA mitocondrial en la población mestiza de la
Ciudad de México**

T E S I S

QUE PARA OPTAR EL GRADO DE:
MAESTRA EN CIENCIAS BIOLÓGICAS

PRESENTA:
BIÓL. CAROLINA LEÓN CAMPOS

TUTOR PRINCIPAL DE TESIS: DRA. AMÉRICA NITXIN CASTAÑEDA SORTIBRÁN
FACULTAD DE CIENCIAS, UNAM

COMITÉ TUTOR: DR. HÉCTOR RANGEL VILLALOBOS
CENTRO UNIVERSITARIO DE LA CIÉNEGA, UDG

COMITÉ TUTOR: DR. JOSÉ MARIANO GURADADO ESTRADA
FACULTAD DE MEDICINA, UNAM

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS

FACULTAD DE CIENCIAS

DIVISIÓN ACADÉMICA DE INVESTIGACIÓN Y POSGRADO

OFICIO FCIE/DAIP/0050/2020

ASUNTO: Oficio de Jurado

M. en C. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
Presente

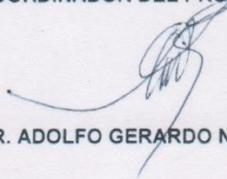
Me permito informar a usted que en la reunión ordinaria del Subcomité de Biología Experimental y Biomedicina del Posgrado en Ciencias Biológicas, celebrada el día **11 de noviembre de 2019** se aprobó el siguiente jurado para el examen de grado de **MAESTRA EN CIENCIAS BIOLÓGICAS** en el campo de conocimiento de **Biología Experimental** de la estudiante **LEÓN CAMPOS CAROLINA** con número de cuenta **309161572** con la tesis titulada "**Análisis de DNA mitocondrial en la población mestiza de la Ciudad de México**", realizada bajo la dirección de la **DRA. AMÉRICA NITXIN CASTAÑEDA SORTIBRÁN**, quedando integrado de la siguiente manera:

Presidente:	DRA. ELLA GLORIA VÁZQUEZ DOMÍNGUEZ
Vocal:	DR. LUIS ENRIQUE EGUIARTE FRUNS
Secretario:	DR. JOSÉ MARIANO GUARDADO ESTRADA
Suplente:	DR. DANIEL IGNACIO PIÑERO DALMAU
Suplente:	DRA. ANGÉLICA GONZÁLEZ OLIVER

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPÍRITU"
Ciudad Universitaria, Cd. Mx., a 15 de enero de 2020

COORDINADOR DEL PROGRAMA



DR. ADOLFO GERARDO NAVARRO SIGÜENZA



AGNS/MMVA/ASR/ipp

COORDINACIÓN DEL POSGRADO EN CIENCIAS BIOLÓGICAS
UNIDAD DE POSGRADO

Edificio D, 1º Piso. Circuito de Posgrados, Ciudad Universitaria
Alcaldía Coyoacán. C. P. 04510 CDMX
Tel. (+5255)5623 7002 <http://pcbiol.posgrado.unam.mx/>

AGRADECIMIENTOS INSTITUCIONALES

Quiero agradecer a mi amada casa de estudios la Universidad Nacional Autónoma de México, a mi querida Facultad de Ciencias y particularmente al Posgrado en Ciencias Biológicas por darme la oportunidad de crecer como profesionista permitiéndome estudiar un posgrado, para mí ha sido una experiencia llena de aprendizaje personal y profesional.

De igual manera agradezco el apoyo PAEP que me otorgó el Posgrado para poder tomar un curso en el extranjero, permitiéndome conocer y aprender algunas de las nuevas investigaciones y técnicas que se utilizan en el área de trabajo donde me desempeño.

Agradezco infinitamente a mi tutora la Dra. América Nitxin Castañeda Sortibrán, quien en todo momento me brindo su apoyo consejos y experiencia, los cuales resultaron vitales para el desarrollo de mi trabajo experimental realizado en el presente posgrado.

A los Doctores Héctor Rangel Villalobos y Mariano Guardado Estrada quienes formaron parte de mi comité tutorial, gracias por haber compartido un poco de su experiencia y consejos profesionales que tanto admiro y respeto, los cuales me ayudaron a perfeccionar mi trabajo de tesis.

“POR MI RAZA HABLARÁ EL ESPIRITÚ”

AGRADECIMIENTOS PERSONALES

Todas aquellas personas que estuvieron a mi lado durante esta travesía, saben lo que el posgrado significa para mí y una vez más, me resulta imposible pensar en todas aquellas personas que me ayudaron en algún punto del camino.

Mi más grande agradecimiento es y seguirá siendo a mis padres quienes estuvieron conmigo en todo momento, me brindaron su apoyo incondicional, siempre confiaron en mí y nunca permitieron que abandonara el camino.

A cada miembro de mi familia gracias por estar y por impulsarme de alguna manera a siempre seguir adelante.

A mi gran amigo y jefecito Mauro López Armenta, gracias por una vez más creer en mí, por permitirme hacer dos de las cosas que más amo en esta vida. Espero seguir aprendiendo de esa gran experiencia tuya.

A mi querida familia INCIFO, gracias por todos los buenos momentos que paso en el laboratorio con ustedes, por sus consejos y apoyo en todo momento.

A mi gran amigo de batallas David. He aquí otro camino recorrido juntos, gracias por cada una de las experiencias que compartimos durante el posgrado.

A mi amiguito querido Andrés. Gracias por acompañarme en las buenas, en las malas y en las peores. Eres una inspiración para mí.

A Hirais, porque al final del día, sin tu ayuda no lo hubiese logrado. Gracias pequeña.

A Alfonso, porque llegaste a darme el empujoncito final que necesitaba. De corazón gracias.

"La vida no es fácil, para ninguno de nosotros. Pero... ¡Qué importa! Hay que perseverar y, sobre todo, tener confianza en uno mismo. Hay que sentirse dotado para realizar alguna cosa y que esa cosa hay que alcanzarla, cueste lo que cueste"

Marie Curie

"En el fondo, los científicos somos gente con suerte: podemos jugar a lo que queramos durante toda la vida"

Lee Smolin

ÍNDICE GENERAL

RESUMEN	1
ABSTRACT	2
INTRODUCCIÓN	3
Globalización, migración y mestizaje.....	3
Genética de poblaciones humanas.....	4
OBJETIVOS	5
OBJETIVO GENERAL.....	5
OBJETIVOS PARTICULARES.....	5
ANTECEDENTES	6
Identificación humana	6
Genética Forense.....	6
DNA nuclear y su variabilidad en el genoma.....	7
Utilidad del mtDNA en casos de identificación.....	10
Polimorfismos del mtDNA	11
Origen y ancestría humana.....	13
Poblamiento de América	15
Evidencias del mtDNA.....	16
Diversidad de mtDNA en México.....	17
Evolución molecular	18
Estimación de flujo génico en una población	20
Relaciones filogenéticas en el mtDNA	21
Inferencia de aspectos históricos de las poblaciones.....	23
Determinación de la riqueza haplotípica de la población	23
JUSTIFICACIÓN	25
METODOLOGÍA	27
POBLACIÓN DE ESTUDIO.....	27
Determinación de la muestra poblacional	27
Obtención de las muestras biológicas.....	28
PROCESAMIENTO DEL mtDNA.....	29
Extracción de DNA.....	29

Amplificación de mtDNA	29
Separación y purificación de fragmentos de mtDNA	30
Marcaje con terminadores fluorescentes y purificación de los fragmentos de mtDNA.....	30
Secuenciación del mtDNA.....	31
ANÁLISIS DE SECUENCIAS	31
Obtención del haplotipo y haplogrupo de mtDNA.....	31
Determinación de los parámetros de diversidad genética	31
Determinación de estimadores de flujo génico	32
Estimación de las relaciones filogenéticas	32
Determinación de cambios en el tamaño poblacional.....	32
Determinación de la riqueza haplotípica de la población.....	32
RESULTADOS	33
Muestra poblacional.....	33
Caracterización de haplotipos y haplogrupos.....	34
Diversidad genética	36
Determinación de flujo génico dentro de la población.....	38
Relaciones filogenéticas de mtDNA.....	40
Inferencias sobre de cambios históricos en el tamaño poblacional	41
Determinación de riqueza haplotípica de la población.....	41
DISCUSIÓN	43
Lugar de origen de los ancestros maternos.....	43
Frecuencia de haplotipos y haplogrupos en población de la CDMX	43
Diversidad genética	46
Estimación del flujo génico en la población	47
Relaciones filogenéticas del mtDNA.....	49
Evaluación cualitativa de cambios en el tamaño poblacional.....	50
Riqueza haplotípica de la población de estudio	50
CONCLUSIONES	51
LITERATURA CITADA	53
ANEXO 1. FORMATOS DE TOMA DE MUESTRA	60
Formato de consentimiento informado.....	60
Cuestionario informativo para toma de muestra.....	62
ANEXO 2. ANCESTROS MATERNOS Y SU LUGAR DE ORIGEN	63

ANEXO 3. HAPLOTIPOS DE mtDNA	64
Haplogrupo A	65
Haplogrupo B.....	71
Haplogrupo C	73
Haplogrupo D.....	75
Haplogrupo H.....	76
Haplogrupo I.....	76
Haplogrupo J.....	76
Haplogrupo K.....	77
Haplogrupo L0.....	77
Haplogrupo L2.....	77
Haplogrupo L3.....	78
Haplogrupo T.....	78
Haplogrupo U.....	79
Haplogrupo W.....	79
ANEXO 4. ESTIMADORES DE FLUJO GÉNICO	80
ANEXO 5. REDES DE HAPLOTIPOS DE mtDNA	82
Haplotipos nativo americanos (Haplogrupo A)	82
Haplotipos nativo americanos (Haplogrupo B)	83
Haplotipos nativo americanos (Haplogrupo C)	84
Haplotipos nativo americanos (Haplogrupo D).....	85
Haplotipos europeos y africanos.....	86

ÍNDICE DE FIGURAS Y TABLAS

Fig. 1. Clasificación del genoma humano. Se muestra la clasificación del genoma humano de acuerdo a su estructura y función, así como el porcentaje de cada uno de ellos (Tomado y modificado de Goodwin et al., 2011).....	7
Fig. 2. DNA mitocondrial. Se muestra la estructura circular del genoma mitocondrial, así como la posición de la región control. Tomada de Rangel, 2019.....	9
Fig. 3. Herencia materna. Se muestra la herencia materna en un árbol genealógico de 18 individuos. Cada haplotipo mitocondrial se ejemplifica con una letra y color en distinto; los individuos en círculo ejemplifican mujeres y los cuadrados hombres. Tomado y modificado de Butler, 2011.	11
Fig. 4. Eva mitocondrial. Se ejemplifica la herencia mitocondrial; cada genotipo de mtDNA se muestra en un color distinto. Se observa cómo es que todos los genomas mitocondriales actuales provienen de uno denominado Eva mitocondrial.	14
Fig. 5. Filogenia de haplogrupos mitocondriales humanos y su geografía aproximada. (A) Haplogrupos y filogenia para grupos de DNA mitocondrial humano (mtDNA); MRCA, antepasado común más reciente. (B) Haplogrupos típicos en poblaciones humanas precoloniales por geografía aproximada. Basado en datos de MitoMAP (Lott et al., 2019) y referencias en ellos. Tomado de Royrvik et al., 2016.....	14
Fig. 6. Rutas de poblamiento del continente americano. Se ejemplifican las rutas de migración que siguieron los primeros pobladores del continente americano, a través del estrecho de Bering mediante las rutas propuestas: IFC, "Corredor Libre de Hielo" (Líneas rojas) y NPC "Costa Norte del Pacífico" (Líneas azules) hace aproximadamente 15,000 años. Tomada y modificada del portal https://www.thinglink.com/scene/846051939102752768	15
Fig. 7. Amplificados de la RC del mtDNA. Se muestran las bandas de amplificado de tres muestras con concentraciones diferentes de DNA: 006 [2.6 ng/μl], 007 [6.43 ng/μl] y 011 [25.91 ng/μl], las cuales fueron amplificadas a diferentes temperaturas de alineamiento (49°C, 51°C, 53°C, 55°C, 57°C y 59°C). El control negativo se amplificó con temperatura de alineamiento de 53°C.....	30
Fig. 8. Lugar de origen de los ancestros maternos. Se muestra el estado de la república mexicana o país de origen al que pertenecen los ancestros maternos, de las personas involucradas en el presente estudio.	33
Fig. 9. Ancestros maternos de acuerdo a su lugar de origen. Se muestra el porcentaje de ancestros maternos según su generación (madre, abuela, bisabuela o tatarabuela) de acuerdo a su lugar de origen (estado de la república mexicana o país), de las personas involucradas en el presente estudio.	34
Fig. 10. Presencia de los Haplogrupos de mtDNA en población de la CDMX. Se muestra la presencia (en %) de los haplogrupos de mtDNA de la muestra poblacional de la CDMX de acuerdo a su origen (1), y la presencia de cada haplogrupo dividido por su origen: amerindio (2), europeo (3) y africano (4).....	36
Fig. 11. Distancia geográfica vs. Distancia genética. Gráfica obtenida de la prueba de Mantel con la información de la distancia geográfica (expresada en kilómetros) y la distancia genética (F_{st} pareadas) por las alcaldías de la CDMX de la muestra poblacional.	39
Fig. 12. Red de Haplotipos de mtDNA para la población de la CDMX. Se muestra la red de haplotipos para los 208 haplotipos de mtDNA obtenidos de la muestra de 251 individuos de la CDMX. Cada haplotipo se encuentra representado con un nodo (círculo); la frecuencia de cada haplotipo se refleja en el tamaño del nodo correspondiente. Los puntos negros representan los vectores medios de la red.	40
Fig. 13. Distribucion mismatch de la población de la CDMX. Se muestra la gráfica de las diferencias encontradas entre cada par de secuencias de la región control del mtDNA vs. la frecuencia de cada una de las diferencias pareadas para la muestra poblacional de la CDMX.	41
Fig. 14. Acumulación de haplotipos (n=251). Se muestra la curva de acumulación de haplotipos (número de muestras vs. número de haplotipos) para el tamaño de muestra del presente estudio.	42

Fig. 15. Acumulación de haplotipos (n=6000). Se muestra la extrapolación de la curva de acumulación de haplotipos para el tamaño de muestra de 6000 individuos muestreados. 42

Fig. 16. Red media de haplotipos nativo americanos pertenecientes al haplogrupo A. Cada uno de 96 haplotipos pertenecientes al haplogrupo A, se representan con un nodo (círculo rojo), el tamaño del mismo es proporcional a la frecuencia del haplotipo en la población; los números representan la posición de las mutaciones que presenta cada haplotipo y los vectores medios se encuentran representados por los puntos negros..... 82

Fig. 17. Red media de haplotipos nativo americanos pertenecientes al haplogrupo B. Cada uno de 40 haplotipos pertenecientes al haplogrupo B, se representan con un nodo (círculo verde), el tamaño del mismo es proporcional a la frecuencia del haplotipo en la población; los números representan la posición de las mutaciones que presenta cada haplotipo y los vectores medios se encuentran representados por los puntos negros..... 83

Fig. 18. Red media de haplotipos nativo americanos pertenecientes al haplogrupo C. Cada uno de 35 haplotipos pertenecientes al haplogrupo C, se representan con un nodo (círculo lila), el tamaño del mismo es proporcional a la frecuencia del haplotipo en la población; los números representan la posición de las mutaciones que presenta cada haplotipo y los vectores medios se encuentran representados por los puntos negros..... 84

Fig. 19. Red media de haplotipos nativo americanos pertenecientes al haplogrupo D. Cada uno de 13 haplotipos pertenecientes al haplogrupo D, se representan con un nodo (círculo turquesa), el tamaño del mismo es proporcional a la frecuencia del haplotipo en la población; los números representan la posición de las mutaciones que presenta cada haplotipo y los vectores medios se encuentran representados por los puntos negros. 85

Fig. 20. Fig. V. Red media de haplotipos nativo americanos pertenecientes a los haplogrupos europeos y africanos. Cada uno de 26 haplotipos pertenecientes al haplogrupos europeos y africanos, se representan con un nodo (círculo de color de acuerdo al haplogrupo), el tamaño del mismo es proporcional a la frecuencia del haplotipo en la población; los números representan la posición de las mutaciones que presenta cada haplotipo y los vectores medios se encuentran representados por los puntos negros... 86

Tabla 1. Frecuencia de haplogrupos nativo americanos en población mexicana. Se muestra la frecuencia (%) de haplogrupos nativo americanos, europeos y africanos reportados en distintos trabajos para distintas poblaciones de México. (*) = Se muestran las frecuencias reportadas para población Mazahua y Otomí respectivamente separadas por una diagonal. (-) = Ausencia de datos. 18

Tabla 2. Muestra poblacional. Se muestran los datos por alcaldía del número de habitantes reportados en el 2015 por el INEGI, el porcentaje de habitantes por alcaldía que representa del total de la población y el número de muestras calculado para el muestreo del presente estudio. 28

Tabla 3. Haplogrupos de mtDNA en la CDMX. Se muestran los haplogrupos de mtDNA obtenidos, el origen al cual pertenecen, el número de muestras que presenta cada haplogrupo, así como el porcentaje del total de la muestra que representa cada uno de ellos. 35

Tabla 4. Sitios polimórficos. Se muestra el número de sitios polimórficos encontrados en la muestra de 251 secuencias, clasificados según sus características de número de variantes y si son informativos o no. 36

Tabla 5. Haplotipos según su frecuencia para la población de la CDMX. Se muestra el número de haplotipos que presenta cada haplogrupo de acuerdo a la frecuencia que presentan en la muestra poblacional de la CDMX, es decir, el número de haplotipos que son idénticos en 2, 3, 4, 5 o 6 personas diferentes. 37

Tabla 6. Valores de diversidad genética. Para cada haplogrupo (**HG**) se muestran los parámetros de diversidad genética: número de mutaciones encontradas (μ), número de sitios polimórficos informativos

(**Inf**) y no informativos (**No Inf**), promedio de diferencias nucleóticas de las secuencias, número de haplotipos (**h**), diversidad haplotípica (**Hd**), diversidad nucleotídica (π), total de sitios segregantes (θw) y finalmente el estadístico D de Tajima. *Significancia estadística: $P < 0.01$37

Tabla 7. Sitios polimórficos de acuerdo al haplogrupo. Se muestra el número de sitios polimórficos encontrados en la muestra de cada haplogrupo nativo americano, europeo y africano, clasificados según sus características de número de variantes y si son informativos o no.38

Tabla 8. Inserciones en las regiones de poli-citocinas de la región control del mtDNA. Se muestra el número de secuencias, y el porcentaje de las mismas en paréntesis, que presentan o no inserciones en las regiones de poli-citocinas de la región control del mtDNA.....38

Tabla 9. Matriz de valores p de los Fst y su significancia. Debajo de la diagonal, se muestran los valores p de los Fst entre las dieciséis alcaldías de la CDMX, sobre la diagonal se muestra si dicho valor obtenido es significativo (+) o no significativo (-). Los asteriscos (*) marcan el cruce entre la misma alcaldía, razón por la cual no se registra información. En rojo se muestran los valores significativos.80

Tabla 10. Matriz de valores de Fst pareadas y distancias geográficas. Debajo de la diagonal, se muestran las distancias genéticas (Fst pareadas) entre las dieciséis alcaldías de la CDMX, sobre la diagonal se muestra la matriz de las distancias geográficas entre las dieciséis alcaldías de la CDMX (km). Los asteriscos (*) marcan el cruce entre la misma alcaldía, razón por la cual se registra información.....81

RESUMEN

La migración en las poblaciones humanas favorece el flujo génico que a su vez genera cambios, reduciendo las diferencias entre las poblaciones. Es importante tener bases de datos actualizadas con frecuencias genéticas que representen una región geográfica de manera adecuada. Debido a su alta tasa de mutación, el genoma mitocondrial acumula una gran cantidad de mutaciones y genera una gran cantidad de haplotipos, razón por la cual los estudios que describan las frecuencias haplotípicas y nucleotídicas del mtDNA deben incluir muestras poblacionales suficientemente grandes, lo cual no ha sido realizado en México donde, el último estudio de este tipo en población de la CDMX se realizó en 2009, con sólo 100 individuos (Guardado *et al.*, 2009).

Con la finalidad de caracterizar las frecuencias de haplotipos y haplogrupos de la región control del mtDNA en población de la CDMX se analizó un fragmento de 917 pb de la región control del mtDNA en 251 individuos. Se estandarizó una metodología para el procesamiento del mtDNA a partir de muestras de células de descamación de mucosa oral desde la extracción de DNA hasta la secuenciación del mtDNA. Se realizó la caracterización de las mutaciones, el haplotipo y el haplogrupo para cada una de las secuencias. Se determinó parámetros de diversidad genética (variabilidad nucleotídica, número de sitios segregantes, variabilidad haplotípica y el *test D* de Tajima) para la población completa, así como para los grupos de haplogrupos de origen nativo americano, europeo y africano. Se estimó de flujo génico con *Fst* y prueba de Mantel y se obtuvieron redes medianas de haplotipos. Para determinar la historia poblacional (tamaño poblacional) de la CDMX se hizo un análisis de *mismatch distribution*. El 90% de los haplogrupos analizados son de origen amerindio o nativo americano, el 8% es de origen europeo y sólo el 2% es de origen africano. De los haplogrupos de origen amerindio, el 45% pertenece al haplogrupo A, el 18.7% al haplogrupo B, el 18.3% al haplogrupo C y el 7.6% al haplogrupo D. Se encontraron 7 haplotipos distintos de origen europeo y 3 de origen africano. Se caracterizaron 201 mutaciones, en 192 sitios polimórficos, un total de 208 haplotipos y una diversidad haplotípica del 99.76%. La mayor diversidad genética se encontró en el haplogrupo A, seguida de los haplogrupos B, C y D. Se observó flujo génico entre la mayoría de las muestras de la población estudiada. Se obtuvieron valores negativos del *test D* de Tajima, sugiriendo que la población se encuentra en expansión; situación que fue corroborada con un análisis de *mismatch distribution*. Las redes medianas de haplotipos obtenidas muestran patrones de estrella. La conformación de las redes muestra la presencia de 32 homoplasias, distribuidas en todos los haplotipos a excepción de los pertenecientes al haplogrupo D, siendo más frecuentes en los haplotipos de los haplogrupos A y B. Todos los resultados anteriormente mencionados sugieren que la población de la CDMX se encuentra en expansión.

ABSTRACT

Migration in human populations favors the gene flow that in turn generates changes, reducing differences between populations. It is important to have updated databases with genetic frequencies that adequately represent a geographic region. Due to its high mutation rate, the mitochondrial genome accumulates a large number of mutations and generates a large number of haplotypes, which is why studies describing the haplotypic and nucleotide frequencies of mtDNA must include sufficiently large population samples, which does not have been carried out in Mexico where, the last study of this type in the population of Mexico City was carried out in 2009, with only 100 individuals (Guardado *et al.*, 2009).

In order to characterize the frequencies of haplotypes and haplogroups of the mtDNA control region in the CDMX population, a 917 bp fragment of the mtDNA control region was analyzed in 251 individuals. A methodology for mtDNA processing was standardized from samples of oral mucosa desquamation cells from DNA extraction to mtDNA sequencing. The characterization of mutations, haplotype and haplogroup was performed for each of the sequences. Genetic diversity parameters (nucleotide variability, number of segregating sites, haplotypic variability and the Tajima D test) were determined for the entire population, as well as for haplogroup groups of Native American, European and African origin. Gene flow was estimated with F_{st} and Mantel test and medium haplotype networks were obtained. To determine the population history (population size) of the CDMX, an analysis of *mismatch distribution* was made. 90% of the haplogroups analyzed are of Native American or Native American origin, 8% are of European origin and only 2% are of African origin. Of the haplogroups of Native American origin, 45% belong to haplogroup A, 18.7% to haplogroup B, 18.3% to haplogroup C and 7.6% to haplogroup D. There were 7 different haplotypes of European origin and 3 of African origin. 201 mutations were characterized, in 192 polymorphic sites, a total of 208 haplotypes and a haplotypic diversity of 99.76%. The greatest genetic diversity was found in haplogroup A, followed by haplogroups B, C and D. Gene flow was observed among the majority of the samples in the population studied. Negative values were obtained from the Tajima D test, suggesting that the population is expanding; situation that was corroborated with an analysis of *mismatch distribution*. The medium networks of haplotypes obtained show star patterns. The conformation of the networks shows the presence of 32 homoplasms, distributed in all haplotypes except those belonging to haplogroup D, being more frequent in haplotypes of haplogroups A and B. All the results mentioned above suggest that the population of the CDMX is expanding.

INTRODUCCIÓN

Globalización, migración y mestizaje

La globalización es un proceso económico, social, político, tecnológico y cultural a nivel mundial, que se caracteriza por generar una mayor conexión, comunicación e interdependencia entre las naciones, llevando consigo una serie de cambios y tendencias hacia la pluralidad y la homogeneización. Consecuencia de la globalización es un alza en la tasa migratoria que genera mestizaje entre las poblaciones involucradas.

El término mestizaje hace referencia a la fusión de dos o más grupos humanos racialmente diferenciados, con orígenes y costumbres distintas que convergen tanto en su herencia biológica como en sus hábitos socioculturales (Ventura, 2010; Castañeda, 2015; López-Beltrán, C., y Deister, 2013). El mestizaje en México inició en el siglo XVI con la llegada de los españoles en 1511 y posteriormente con la conquista española en 1519. Aunque en su momento se trataron de evitar las mezclas raciales, durante el siglo XVII y XVIII el mestizaje en México comenzó a ser evidente, constituyéndose una población con individuos procedentes de tres grupos étnicos: indios (pobladores originales de México), blancos (españoles que llegaron a conquistar) y negros (esclavos de origen africano) (Bernabéu, y Varela, 2010).

En la actualidad, sin importar el origen racial de las poblaciones humanas, todas ellas se encuentran sujetas a mestizaje, debido a los constantes movimientos poblacionales entre una nación y otra. Según datos del INEGI, en 2010 habitaban en México 961,121 personas que habían nacido en otro país y en 2018, según la Encuesta Nacional de la Dinámica Demográfica, se estimó que el 18% de la población mexicana (22.5 millones de personas) son migrantes, de los cuales 1,074,000 personas nacieron en otro país.

La Ciudad de México cuenta con una población cercana a los 9 millones de habitantes y se encuentra sujeta a múltiples migraciones de personas que habitan otros estados de la república en busca de oportunidades laborales, así como de personas de otros países. Según el INEGI, en el 2010 llegaron a vivir a dicha ciudad un total de 239,125 personas procedentes de distintos estados de la república como el Estado de México, Puebla, Veracruz, Oaxaca y Guerrero, y se estimó que 71,691 habitantes residentes de la CDMX habían nacido en otro país.

Todos esos flujos migratorios, generan poblaciones dinámicas sometidas a constantes cambios en las frecuencias alélicas de los genes; los cuales son proporcionales a las diferencias entre las frecuencias de las poblaciones involucradas y a la tasa de migración que existe entre ambas naciones, homogeneizando cada vez más la información genética de ellas.

Genética de poblaciones humanas

La genética de poblaciones estudia la variabilidad genética de los diferentes grupos humanos desde una perspectiva evolutiva, así como las causas que determinan dicha variabilidad a través del análisis de los procesos micro-evolutivos que operan sobre la estructura genética de una población a lo largo de las generaciones. Aunque los procesos micro-evolutivos son iguales para todas las especies, en las poblaciones humanas, además de los factores biológicos y ecológicos deben contemplarse los culturales. Un ejemplo es que los cruzamientos no se producen de forma aleatoria por las tendencias selectivas de índole sociocultural en las poblaciones humanas (Fontdevila y Moya, 2007).

Otro aspecto a considerar en las poblaciones humanas son los procesos migratorios actuales que además de evitar el aislamiento geográfico-reproductivo, evitan la delimitación genética de los grupos humanos, generando grupos híbridos en áreas geográficas muy alejadas de sus zonas de origen. Debido a esta dispersión y mezcla de las variantes genéticas, es necesaria la reconstrucción de las rutas migratorias con la finalidad de dilucidar los patrones evolutivos (Herrera, 2013).

Las diferencias genéticas observadas entre las poblaciones, son la base fundamental de disciplinas como la Antropología Genética, la Genética Forense y recientemente la Medicina Genómica. Para estas disciplinas, es indispensable generar bases de datos poblacionales que recopilen la información de las variantes genéticas mediante el estudio de muestras representativas. Por su parte, la genética forense requiere de bases de datos de referencia con frecuencias alélicas para los marcadores o polimorfismos genéticos analizados, ya que éstas son necesarias para establecer la probabilidad de coincidencia al azar de un perfil genético en una determinada población. Estas bases de datos involucran diversos marcadores como pueden ser microsatélites de repetición tipo STR, polimorfismos de inserción/delección (InDels), polimorfismos de un solo nucleótido (SNP) tanto en DNA nuclear o en mtDNA. Es de suma importancia en dichas bases la actualización constante de las mismas, ya que como anteriormente se mencionó, la variación genética de las poblaciones humanas se encuentra en constante cambio, tanto por los fenómenos naturales que las generan (mutación y herencia) como por los procesos migratorios, consecuencia de la globalización actual en la que vivimos.

OBJETIVOS

OBJETIVO GENERAL

Caracterizar las frecuencias haplotípicas de la región control del mtDNA en una muestra de la población mestiza de la Ciudad de México.

OBJETIVOS PARTICULARES

- Establecer y estandarizar un protocolo metodológico desde la extracción hasta el análisis electroforético de mtDNA a partir de muestras de epitelio bucal.
- Describir y reportar los haplotipos y haplogrupos de la población de la Ciudad de México.
- Determinar parámetros de diversidad genética destacando variabilidad nucleotídica, haplotípica.
- Generar un sistema general de referencia útil para estudios de identificación con mtDNA.

ANTECEDENTES

Identificación humana

Se llama identidad a aquellas características y rasgos particulares de un individuo que lo diferencian del resto y lo hace único; se trata de un derecho humano universal que se encuentra reconocido por distintos tratados internacionales como lo son: la Declaración Universal de Derechos Humanos, el Pacto Internacional sobre Derechos Civiles y Políticos, la Convención Americana sobre Derechos Humanos y la Convención sobre los derechos del Niño (INEGI y UNICEF, 2018). Es un derecho que se tiene desde el nacimiento y se conserva aún después de la muerte; a pesar de ello, en muchas ocasiones tras la muerte dicha identidad se pierde, motivo por el cual es necesario llevar a cabo un proceso de identificación que le devuelva la identidad al cadáver en cuestión.

La identificación humana en el contexto forense es una tarea multidisciplinaria en la que trabajan de manera conjunta disciplinas como dactiloscopia, odontología y genética, las cuales buscan encontrar las características fenotípicas y genotípicas del cadáver que lo caracterizan para devolverle su identidad. En la mayoría de las ocasiones resulta sencilla la identificación mediante el uso de disciplinas rutinarias como lo es la dactiloscopia y la odontología. La primera disciplina analiza los patrones de las crestas papilares, también conocidos como huella dactilar. Debido a que el origen de dichas crestas, es el plegamiento de la piel desde la gestación, estos patrones resultan diferentes para cada persona aun cuando se trate de gemelos monocigóticos, lo que permite identificar las personas de manera inequívoca (Thompson y Black, 2007).

Otra disciplina que permite la identificación de personas de manera certera es la odontología, la cual estudia la forma, disposición y posibles heridas de los dientes, mandíbula y otros tejidos orales, así como también los trabajos dentales de ortodoncias, endodoncias, placas dentales parciales o totales, etc. Estas características, al tratarse de un conjunto de señas particulares, si se cuenta con expedientes dentales *ante mortem* de la persona (radiografías, impresiones o moldes dentales) es posible la identificación de personas que se encuentran en calidad de desconocidas (Thompson y Black, 2007).

Genética Forense

La genética a diferencia de las disciplinas anteriores, es capaz de identificar a las personas por su genotipo (características genéticas) y no por su fenotipo (características físicas). Esta tarea se logra mediante la obtención de perfiles genéticos y su posterior confronta con aquellos familiares en línea directa, lo que permite identificar las relaciones de parentesco biológico según sea el caso y determinar la identidad de las personas en cuestión.

El genoma humano: Características generales del DNA.

El material de estudio de la genética es el ácido desoxirribonucleico o DNA (por sus siglas en inglés). El DNA es la molécula en donde se almacena toda la información genética (genoma) de un individuo (Butler, 2005). En 1990 dio inicio el proyecto del Genoma Humano, el cual buscaba la decodificación del genoma humano a través de la técnica "Shotgun sequencing", llevada a cabo de manera simultánea en distintos centros involucrados en el proyecto. En 2001, se obtuvo una porción consenso del genoma humano de 2,910,000,000 pb, a partir de la cual se pudo determinar que únicamente el 1.1% del genoma son exones, 24% intrones y 75% del genoma se trata de DNA intergénico (Venter *et al.*, 2001). En la Fig. 1 se muestra la clasificación de los tipos de DNA en los que se divide el genoma humano de acuerdo a su estructura y función.

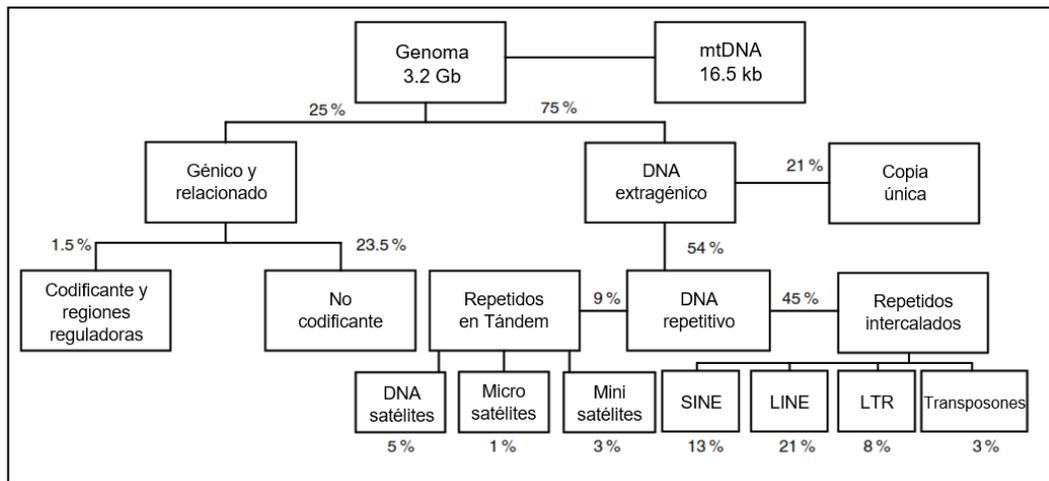


Fig. 1. Clasificación del genoma humano. Se muestra la clasificación del genoma humano de acuerdo a su estructura y función, así como el porcentaje de cada uno de ellos (Tomado y modificado de Goodwin *et al.*, 2011).

DNA nuclear y su variabilidad en el genoma

La mayor parte del DNA contenido en el genoma humano consiste de fragmentos repetidos no codificantes (54%), de los cuales el 9% son repetidos en bloque o *tándem*, a partir de los cuales podemos identificar tres categorías: DNA satélite, minisatélites y microsatélites (Fig. 2). Lo importante de dichos repetidos es que en ellos se encuentra la mayor parte de la variabilidad humana; se sabe que todos los humanos compartimos alrededor del 99% de la información genética, mientras que el 1% restante es distinto entre cada individuo (excepto gemelos monocigóticos), lo que permite individualizar e identificar a las personas si se analizan esas variantes o polimorfismos del genoma (Goodwin *et al.*, 2011; Thompson y Black, 2007).

La identificación humana mediante polimorfismos del DNA comenzó en 1984 cuando Alec Jeffreys descubrió los minisatélites; regiones del DNA que se repetían consecutivamente una y otra vez. También descubrió que el número de repetidos que tenían las personas podía ser

diferente entre ellas, lo que le permitió realizar las primeras pruebas de identificación humana (Jobling y Gill, 2004).

Los minisatélites tienen una longitud de 104 a 134 repetidos (Goodwin *et al.*, 2011), convirtiéndolos en fragmentos largos y hasta cierto punto difíciles de obtener si se cuenta con material biológico degradado o en pequeñas cantidades; es por ello que la identificación humana a partir del estudio de los microsátélites (los cuales constan de unidades de repetición de 2 a 6 pb) se ha convertido en la mayor herramienta de identificación humana en los últimos años (von Wurmb-Schwark *et al.*, 2009). Los microsátélites mayormente utilizados son los repetidos cortos en tándem o STR (por sus siglas en inglés); se trata de repetidos más pequeños que los minisatélites y son los *loci* más polimórficos del genoma humano (Willems *et al.*, 2014, Butler, 2005).

Los polimorfismos de un solo nucleótido o SNPs (por sus siglas en inglés) se tratan del polimorfismo más sencillo que existe, el cual consiste en la diferencia de la secuencia de DNA en una sola base. Estos cambios en la secuencia se deben a errores que ocurren en la replicación del DNA (Goodwin *et al.*, 2011). Los SNPs son abundantes en todo el genoma humano y se han utilizado para estudios de vinculación de enfermedades, así como para identificación humana (Butler, 2009). Los polimorfismos de un solo nucleótido, además de encontrarse presentes en el DNA nuclear también lo hacen en el DNA mitocondrial, situación que permite estudios de identificación humana, así como estudios de ancestría.

DNA mitocondrial

El DNA mitocondrial (mtDNA) es un tipo de DNA eucarionte, que como su nombre lo indica, corresponde al genoma particular de los organelos denominados mitocondrias. Es una molécula circular de 16,569 pb que codifica 37 genes: 22 tRNAs, 13 proteínas y 2 rRNAs (12S y 16S). En su mayoría se trata de DNA codificante, sin embargo, existe una región no codificante de 1,122 pb conocida como D-loop, región hipervariable o región control, en la cual se encuentra el sitio de inicio de la transcripción y que acumula un gran número de mutaciones tipo SNP características de este tipo de DNA (Fig. 2) (Butler, 2011).

El número de copias de mtDNA que puede tener una mitocondria varía de 1 a 15, teniendo en promedio de 4 a 5; debido a que una célula contiene cientos de mitocondrias, por célula se tienen de cientos a miles de copias de mtDNA (de acuerdo al tipo de célula que se trate) y, en promedio se estiman 500 copias de mtDNA por célula (Butler, 2011). De esta manera, se tiene que el genoma mitocondrial comprende alrededor del 0.25% del total de DNA en una célula, pero únicamente el 0.00006% del total de la información genética humana (Stoneking, 1994).

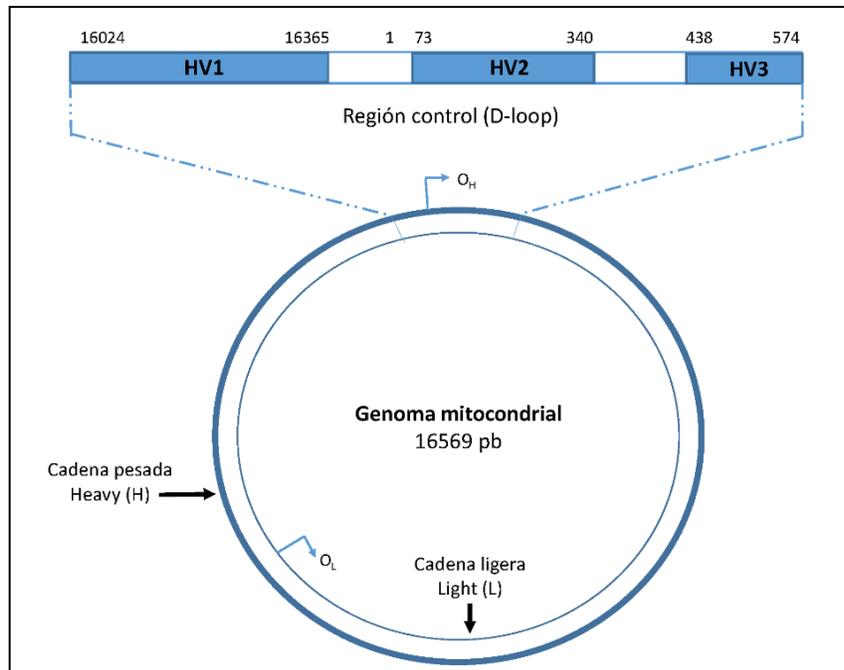


Fig. 2. DNA mitocondrial. Se muestra la estructura circular del genoma mitocondrial, así como la posición de la región control. Tomada de Rangel, 2019.

Una de las características de mayor relevancia del mtDNA es que no lleva a cabo recombinación genética de una generación a otra, como sucede con el DNA nuclear, sino que se hereda en bloque únicamente en vía materna. Al momento de la fecundación, únicamente ingresa al óvulo la cabeza del espermatozoide (que contiene el núcleo celular), dejando fuera la parte intermedia y con ello las mitocondrias. De esta manera, el material genético resultado de la fecundación consiste en dos copias de DNA nuclear (materno y paterno), así como una copia de mtDNA (materno).

Sin embargo, en algunas ocasiones al momento de la fertilización, alcanza a ingresar al óvulo una fracción muy pequeña de la parte intermedia de los espermatozoides, por lo que se activan mecanismos selectivos de eliminación de las mitocondrias paternas después de su entrada al óvulo (Chinnery y Hudson, 2013). Ejemplo de ello es mediante la enzima CPS-6 (endonucleasa G). Posterior a la fertilización, las mitocondrias paternas sufren despolarización y sus membranas internas pierden su integridad generando una marca autofágica; además de que el rompimiento de la membrana interna permite la entrada de la CPS-6 a la matriz de las mitocondrias paternas para degradar su DNA, de esta manera se preserva únicamente el mtDNA materno (Zhou, 2016).

Otro mecanismo que permite la eliminación del mtDNA paterno es mediante marcas de ubiquitinación. Las marcas de ubiquitinación en las mitocondrias de los espermatozoides se adquieren durante la diferenciación de las espermátidas en los testículos y se enmascaran mientras el espermatozoide se encuentra en el epidídimo. Posterior a la fertilización, el

espermatozoide se encuentra expuesto a péptidos como el glutatión y sinergistas del mismo que se encuentran presentes en el ovocito, dichos péptidos favorecen la exposición de la marca de ubiquitinación que contiene el espermatozoide. Posterior a la exposición de la marca, las mitocondrias del espermatozoide son detectadas y eliminadas del ovocito (Sutovsky, 2003).

El mtDNA paterno que ingresa al óvulo también se puede eliminar por efecto de dilución; los espermatozoides únicamente contienen 100 copias de mtDNA, comparado con las 100, 000 copias de un óvulo no fertilizado que, sumado al efecto de cuello de botella (pérdida de variabilidad en aquellas mutaciones que se encuentran en menores cantidades dentro del genoma) al cual se encuentra sometido el mtDNA, es que se favorece la herencia del mtDNA materno sobre el paterno. Los mecanismos por los cuales ocurren los cuellos de botella se deben a una segregación desigual ya sea de los nucleótidos del mtDNA durante la división celular o bien de todo genoma mitocondrial, así como a la replicación selectiva de una subpoblación específica de mtDNA (Chinnery y Hudson, 2013).

Lo anterior justifica que el genoma mitocondrial de un individuo es heredado en su totalidad vía materna. Sin embargo, un estudio realizado en tres familias no relacionadas entre sí, demostró que existe herencia biparental de mtDNA entre algunos de sus individuos. Una de las particularidades que presentan los genotipos de mtDNA con herencia biparental es una alta tasa de heteroplasmias y haplogrupos misceláneos. Aunque en dicho estudio se demuestra la herencia biparental en el mtDNA, la herencia paterna consiste en el 40% o menos del total del mtDNA y no se observa en todos los casos (Luo *et al.*, 2018).

Lo anterior genera duda sobre los mecanismos de herencia del mtDNA, si bien es cierto que aún se desconocen las razones por las cuales puede existir herencia biparental, una posible explicación es que sea debida a mutaciones en las endonucleasas G (CPS-6) que se encargan de eliminar el mtDNA paterno que ingresa a la matriz del óvulo posterior a la fertilización. A pesar de tal hallazgo en la herencia del mtDNA, es claro ver que la herencia materna continúa siendo dominante en la escala evolutiva y en determinadas ocasiones han existido eventos de herencia paterna que parecen no tener efectos detectables en el registro genético humano (Luo *et al.*, 2018).

Utilidad del mtDNA en casos de identificación

La identificación humana con sistemas de DNA nuclear (STR) en determinadas ocasiones no aporta los resultados esperados, o sencillamente no genera resultado alguno, debido al alto grado de degradación del DNA, debido a mínimas cantidades o porque se trata de DNA antiguo. Sin embargo, en esas muestras la recuperación de DNA no se encuentra totalmente perdida ya que, aunque la recuperación de DNA nuclear haya fallado, aún es posible la obtención de mtDNA. En escenarios de desastres masivos que involucren terremotos, incendios, derrumbes, etc. que hayan dejado un gran número de víctimas y cuerpos reducidos a escombros o cenizas, es donde el mtDNA resulta útil. Si bien es cierto que el poder de

significancia en un *match* es mucho menor comparado con el DNA nuclear, en cualquier escenario puede resultar lo suficientemente informativo para la identificación de una persona.

En la Fig. 3 se ejemplifica la herencia del mtDNA de 18 personas, de las cuales ocho de ellas presentan el mismo haplotipo; si bien es cierto que esto no permite individualizar una persona porque su genoma mitocondrial se comparte con todo su linaje materno, sí permite la inclusión o exclusión de parentescos biológicos debido a la diferenciación de sus haplotipos debido a la presencia o ausencia de mutaciones particulares de cada uno de ellos (Butler, 2011).

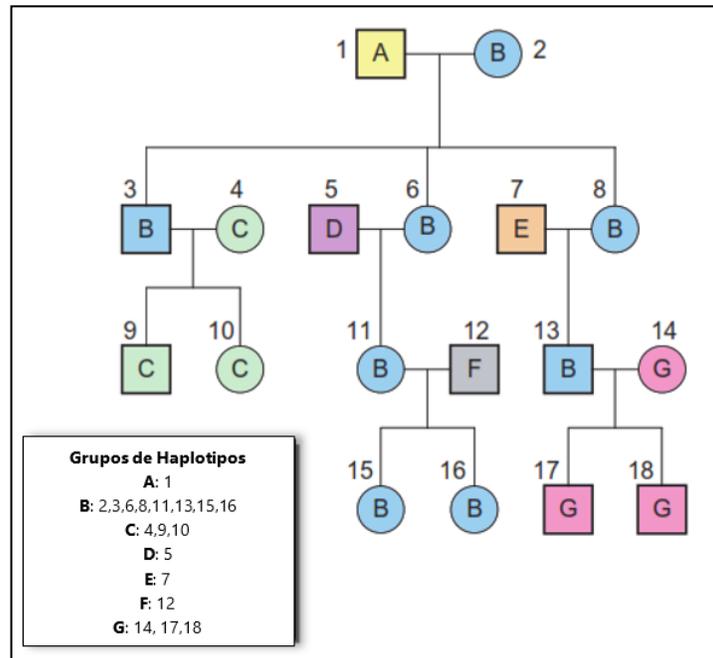


Fig. 3. Herencia materna. Se muestra la herencia materna en un árbol genealógico de 18 individuos. Cada haplotipo mitocondrial se ejemplifica con una letra y color en distinto; los individuos en círculo ejemplifican mujeres y los cuadrados hombres. Tomado y modificado de Butler, 2011.

Polimorfismos del mtDNA

El mtDNA tiene una tasa de mutación de 5 a 10 veces más alta que el DNA nuclear (Butler, 2011). Se estima que es de 0.017×10^{-6} sustituciones por sitio/año para el genoma mitocondrial en general, sin considerar a la región control (Ingman *et al.*, 2000). Esa gran tasa de mutación se debe a la exposición del mtDNA ante especies reactivas de oxígeno producto de la fosforilación oxidativa. El sistema respiratorio mitocondrial es la principal fuente intracelular de las especies reactivas de oxígeno (ROS) y los radicales libres, que se generan como subproductos durante la transferencia de electrones de NADH o FADH₂ al oxígeno molecular en condiciones fisiológicas normales (Wei, 1998). Lo anterior genera que las mutaciones se

acumulen rápidamente pasando de una generación a otra. Se estima que dichas mutaciones se transmiten una vez en cada 30 eventos, siendo la mayoría de las veces mutaciones en el cambio de un nucleótido (Goodwin *et al.*, 2011).

Dentro de la región control existen bloques de DNA altamente conservados y dos bloques altamente polimórficos a los que se les conoce como regiones hipervariables 1 y 2. En ellas se encuentran los niveles más altos de variación del mtDNA humano (Goodwin *et al.*, 2011; Thompson y Black, 2007); se estima que la tasa de mutación para dichas regiones es del $0.075-0.165 \times 10^{-6}$ sustituciones por sitio/año (Pakendorf y Stoneking 2005).

Cada célula contiene miles de moléculas de mtDNA y en la mayoría sus secuencias son idénticas, a esto se le conoce como **homoplasmia**. Sin embargo, se generan mutaciones frecuentes por la continua exposición del mtDNA a un ambiente oxidativo, ineficiente reparación del mtDNA y una constante replicación sumado a la naturaleza poliploide del mtDNA; dichas mutaciones pueden coexistir con el genotipo original en diferentes proporciones, a lo que se le conoce como **heteroplasmia** (Chinnery y Hudson, 2013). Un estudio estimó que el 14% de la población presenta un segundo tipo de mtDNA que constituye al menos el 1% de su genoma mitocondrial (Tully *et al.*, 2000).

A la suma de los polimorfismos de mtDNA que tiene un individuo, se le conoce como **haplotipo**. La herencia jerárquica sin recombinación de cada haplotipo, se le conoce como **haplogrupo**. Debido a que la herencia es uniparental y no existe recombinación genética, los haplogrupos de mtDNA son susceptibles a sufrir deriva génica, y con ello la eliminación o disminución de la variabilidad genética, situación que ha generado patrones de diferenciación en distintas áreas geográficas, especialmente a escala continental (Royrvik *et al.*, 2016). La caracterización del mtDNA en primer lugar conlleva la obtención de la secuencia nucleotídica, se recomienda que sea tanto la secuencia sentido como la anti-sentido (forward y reverse), para poder comparar entre ambas y corroborar los resultados. Posteriormente es necesaria la comparación de dicha secuencia con una secuencia de referencia.

La primera vez que se secuenció el genoma mitocondrial humano fue en 1981, en el laboratorio de Frederick Sanger en Cambridge, Inglaterra (Anderson *et al.*, 1981). A dicha secuencia se le nombró Secuencia de Referencia de Cambridge (CRS por sus siglas en inglés). Poco tiempo después fue sometida a una revisión para confirmar los resultados de la secuencia y poder utilizarla como la secuencia de referencia universal. En esa revisión se corrigieron 11 errores en la secuencia, aunque ninguno de ellos se encontraba en la región control y, a partir de dicha revisión se le nombró como Secuencia de Referencia de Cambridge Revisada (rCRS por sus siglas en inglés) (Andrews *et al.*, 1999). Es importante destacar que la rCRS no es una secuencia consenso del mtDNA, sino que se trata de una secuencia de Haplotipo Europeo (haplogrupo H2a2) que se utiliza como estándar para el reporte de los polimorfismos del mtDNA desde hace más de 30 años (Bandelt *et al.*, 2013).

Para determinar las mutaciones que tiene la secuencia a caracterizar, se debe alinear con la sCRS y aquellas posiciones en las que el nucleótido sea diferente respecto a la referencia, son consideradas mutaciones. Existen diferentes guías para tipificar las mutaciones del mtDNA, una de ellas es la de la Comisión de DNA de la Sociedad Internacional de Genética Forense (Parson *et al.*, 2014) y la del Grupo Europeo de Perfilación de DNA (EDNAP por sus siglas en inglés) (Tully *et al.*, 2001). El presente trabajo utilizará las recomendaciones de la segunda guía para la interpretación de las secuencias de mtDNA que se analicen.

Una vez que se tienen descritas todas las mutaciones de la secuencia en cuestión (haplotipo), es posible determinar el haplogrupo al cual pertenece. La caracterización de los haplogrupos de mtDNA se basa en la presencia de ciertas mutaciones que en conjunto se denominan "Marcadores de Haplogrupo". Las secuencias que presenten cierto conjunto de mutaciones pertenecerán a un determinado haplogrupo. Sin embargo, esto no significa que sean las únicas mutaciones que presente la secuencia, sino que adicional a ellas su haplotipo puede estar compuesto por otras mutaciones. Los marcadores de cada haplogrupo pueden ser consultados en la Base de Datos del Genoma Mitocondrial Humano MITOMAP (Lott, 2019).

Origen y ancestría humana

Conocer nuestro origen evolutivo y ancestros es un tema al cual se le ha buscado dar respuesta mediante distintos estudios, de carácter paleontológico, antropológico, arqueológicos y recientemente molecular. Estudios antropológicos y arqueológicos sitúan a los yacimientos más antiguos de los humanos al este de África. A partir del estudio de las mutaciones que se encuentran en la región control del mtDNA surge la hipótesis denominada "Eva mitocondrial" (Cann *et al.*, 1987), en la cual se propone que todas las mitocondrias humanas provienen de un mismo ancestro en común que vivió aproximadamente hace 200,000 años en África y que posteriormente migró hacia el resto del mundo (Stoneking, 1994).

Lo anterior no significa que únicamente existió una mujer en el inicio de los tiempos, y que a partir de su mtDNA deriva toda la descendencia humana actual, sino que fue sólo la descendencia de una mujer la que sobrevivió a lo largo de las generaciones hasta la actualidad. En contraparte, el resto de los genomas de mtDNA proveniente de otras mujeres sufrió procesos de extinción aleatorios (Fig. 4).

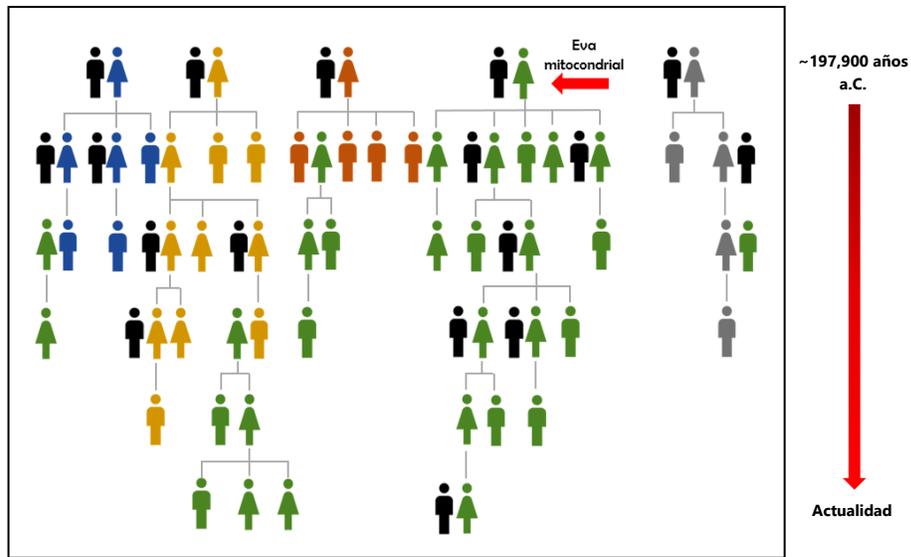


Fig. 4. Eva mitocondrial. Se ejemplifica la herencia mitocondrial; cada genotipo de mtDNA se muestra en un color distinto. Se observa cómo es que todos los genomas mitocondriales actuales provienen de uno denominado Eva mitocondrial.

Gracias al estudio y caracterización de los haplotipos de mtDNA en distintas poblaciones del mundo, se han clasificado haplogrupos que representan genomas que comparten mutaciones en común y que además muestran especificidad regional. En la Fig. 5 "A" se muestra la filogenia de los haplogrupos de mtDNA hasta ahora descritos; en "B" se representan las áreas geográficas donde se encuentran mayormente representados los haplogrupos. Se observa que los haplogrupos L (L0, L1, L2, L3, L4, L5 y L6) se encuentran restringidos a África. Por su parte, el macrohaplogrupo L3 se expandió desde África hacia Europa y Asia mediante los macrohaplogrupos M y N. Los haplogrupos H, I, J, N1b, T, U, V, W y X son característicos de Europa. En Asia los haplogrupos pertenecientes a los macrohaplogrupos M y N se encuentran representados, con una ligera mayoría de A, B, C y D; los haplogrupos G, Y y Z se encuentran predominantemente en Asia central y Siberia. En el continente americano, la variación se encuentra en los haplogrupos A, B, C, D y X (Pakendorf y Stoneking, 2005; Mishmar *et al.*, 2003; Reidla *et al.*, 2003).

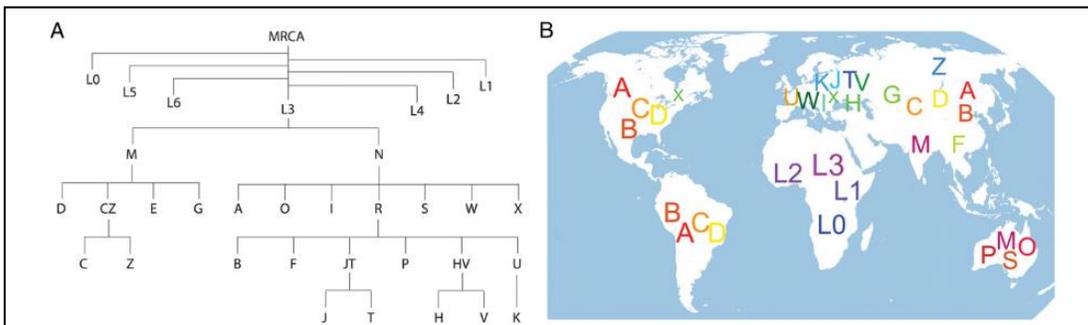


Fig. 5. Filogenia de haplogrupos mitocondriales humanos y su geografía aproximada. (A) Haplogrupos y filogenia para grupos de DNA mitocondrial humano (mtDNA); MRCA, antepasado común más reciente. (B) Haplogrupos típicos en poblaciones humanas precoloniales por geografía aproximada. Basado en datos de MitoMAP (Lott *et al.*, 2019) y referencias en ellos. Tomado de Royrvik *et al.*, 2016.

Poblamiento de América

Como anteriormente se mencionó, el origen del hombre moderno (el género *Homo*) sucedió en África, en donde evolucionó hace 150,000 años y comenzó a migrar al resto del mundo hace 100,000 años, principalmente hacia Europa y Asia (Templeton, 2006). Debido al aislamiento geográfico que guarda el continente americano respecto a Europa, Asia y África, América fue el último continente en ser poblado por el hombre moderno. El entendimiento de cómo sucedió tal situación, ha generado décadas de investigación paleontológica, arqueológica, genética, paleo ecológica y cultural, las cuales han permitido vislumbrar el modelo que explique la colonización de América (Potter *et al.*, 2018).

Durante el Pleistoceno, el extremo norte de Eurasia se encontraba unido con América territorio al que se le denominaba Beringia o Estrecho de Bering, que en la actualidad se encuentra sumergido bajo el mar. La teoría mayormente aceptada es que los primeros pobladores de América migraron a lo que hoy en día es Alaska a través del estrecho de Bering (Hace ~14,900 años) y posteriormente se fueron desplazando hacia el sur del continente hasta colonizarlo en su totalidad (Potter *et al.*, 2017).

Existen distintas teorías acerca de las rutas que siguieron para llegar al continente americano, una de ellas es la "Costa Norte del Pacífico" (NPC), la cual supone que las rutas de colonización sucedieron desde el sur y/o centro del Estrecho de Bering, mientras que la teoría del "Corredor libre de hielo" (IFC) plantea la posibilidad de múltiples rutas que ingresan al norte del continente por el oeste del Río Mackenzie entre las secciones del norte de Laurentida y la Cordillera de Hielo, o a través de las áreas de Liard y Peace River (Fig. 6) (Potter *et al.*, 2017; Potter *et al.*, 2018).

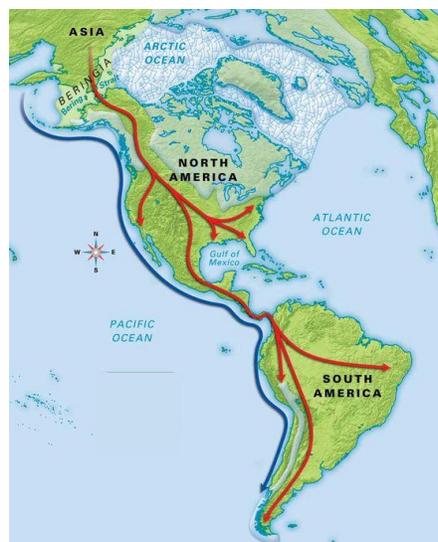


Fig. 6. Rutas de poblamiento del continente americano. Se ejemplifican las rutas de migración que siguieron los primeros pobladores del continente americano, a través del estrecho de Bering mediante las rutas propuestas: IFC, "Corredor Libre de Hielo" (Líneas rojas) y NPC "Costa Norte del Pacífico" (Líneas azules) hace aproximadamente 15,000 años. Tomada y modificada del portal <https://www.thinglink.com/scene/846051939102752768>.

Evidencias del mtDNA

Distintos análisis con secuenciación de genomas completos y regiones hipervariables del mtDNA indican que los ancestros de nativos americanos descienden de poblaciones de Siberia de hace 24,900 - 18,400 años de antigüedad, los cuales se diversificaron hace 16,000 - 13,000 años (Llamas *et al.*, 2016; Raghavan *et al.*, 2015). A pesar de que en Siberia existe una gran variabilidad de mtDNA, en el continente americano se perdió la gran mayoría de esa diversidad genética "original" debido al efecto fundador al que se vio sometida la población humana (Comas, 2014); es decir, la población del continente americano derivó a partir de un número pequeño de individuos provenientes de una gran población ancestral. El tamaño de la población de América fue muy pequeño por algunas generaciones, seguido de su posterior crecimiento, fenómeno conocido como cuello de botella (Templeton, 2006), en donde sumado a las constantes migraciones, se fueron generando gradientes genéticos de diversidad denominadas clinas genéticas. A medida que los individuos descendientes de una población avanzan y se desplazan a lo largo de las generaciones, se van diferenciando genéticamente de manera gradual, por el efecto de la deriva genética de la población actual; esto significa que se pueden encontrar cambios graduales en las frecuencias de las variantes genéticas de las poblaciones (Comas, 2014).

Un estudio con haplogrupos de mtDNA demuestra que existe una gran cantidad de variación en la distribución de frecuencias de haplogrupos entre regiones geográficas de América, lo que sugiere que los primeros colonizadores se diseminaron por todo el continente en pequeñas bandas que se vieron fuertemente afectadas por la deriva genética. Sin embargo, el hecho de que el 40% de los haplotipos compartidos entre los nativos americanos de América del Norte son compartidos entre individuos geográficamente distantes sugiere que los primeros habitantes de esta región experimentaron cantidades sustanciales de flujo genético (Malhi *et al.*, 2002).

Por los procesos de deriva génica anteriormente mencionados, en poblaciones nativas de América, se encuentran representados únicamente 5 haplogrupos de mtDNA que representan el 100% de la variabilidad de mtDNA. Los haplogrupos A, C y D (que representan el 58%), provienen de poblaciones del norte de Siberia, que llegaron a América a través del estrecho de Bering. El haplogrupo B debió haber llegado después ya que no se encuentra en poblaciones de Siberia; se encuentra representado en la costa Asiática del Pacífico, por lo que debió haber llegado a América por la ruta de la Costa Norte del Pacífico. Por su parte el haplogrupo X se encuentra concentrado en la parte central de América del Norte y es distante con el haplogrupo X que se encuentra en Europa; nuevamente, la posible ruta de llegada de éste haplogrupo al continente americano fue por la ruta del Norte (Mishmar *et al.*, 2003).

Diversidad de mtDNA en México

Las frecuencias observadas de los haplogrupos de mtDNA en México son: 89-97% de linajes de origen nativo americano, 2-5% con ancestría europea y 1-4% con ancestría africana (Green *et al.*, 2000 y González-Sobrino *et al.*, 2016). La mayoría de los haplotipos europeos que se encuentran en México pertenecen a los haplogrupos H, I, J, K, T, U, V, W y X; finalmente, los haplotipos africanos que se encuentran en México pertenecen principalmente al haplogrupo L3 (Green *et al.*, 2000).

Las frecuencias de los haplogrupos amerindios en población mexicana varían de acuerdo a la región geográfica, así como a las comunidades analizadas. En la Tabla 1 se muestran las frecuencias de los haplogrupos nativo americanos reportadas en los trabajos:

Green *et al.*, (2000). Estudio con 223 muestras de poblaciones de Juárez y Ojinaga, Chihuahua del norte del país.

Peñaloza *et al.*, (2007). Estudio con 513 muestras de individuos pertenecientes a 14 poblaciones indígenas (Mixteca-Alta, Mixteca-Baja, Otomí, Purépecha, Tzeltal, Tarahumara, Huichol, Nahua-Atocpan, Nahua-Xochimilco, Nahua-Zitlala, Nahua-Chilacachapa, Nahua-Ixhuatlancillo, Nahua-Necoxtla y Nahua-Coyolillo) distribuidas en principalmente en la zona centro y sur del país.

Sandoval *et al.*, (2009). Estudio con 477 muestras de 11 poblaciones indígenas (Triqui, Tarahumara, Purépecha, Otomí, Mixteca, Nahua Xochimilco, Nahua Zitlatla, Nahua Necoxtla, Maya y Pima.)

Guardado *et al.*, (2009). Estudio con 270 individuos de población de 18 estados (Ciudad de México, Veracruz, Michoacán, Guerrero, Oaxaca, Chiapas, Jalisco, San Luis Potosí, Zacatecas, Durango, Sinaloa, Estado de México, Hidalgo, Puebla, Tlaxcala, Morelos, Querétaro y Guanajuato) del país.

Gorostiza *et al.*, (2012). Estudio con 605 individuos provenientes de 8 grupos indígenas de Mesoamérica (Pima, Mayo, Huichol, Nahua, Otomí Valle, Otomí Sierra, Tepehua, Maya), en su mayoría pertenecientes al territorio mexicano.

Martínez *et al.*, (2013). Estudio con 742 individuos de población mestiza de 10 estados del país: Chihuahua, Sinaloa, Durango, Aguascalientes, Jalisco, Veracruz, Guerrero, Oaxaca, Chiapas y Yucatán.

González-Martín *et al.*, (2015). Estudio con 520 individuos pertenecientes a 8 grupos indígenas (Yucatán Maya, Quintana Roo Maya, Campeche Maya, Tojolabal, Tzotzil, Mazateco, Purepecha y Huichol).

González-Sobrino *et al.*, (2016). Estudio con 3,036 individuos pertenecientes a 10 estados de México (Sonora, Durango, Zacatecas, Guanajuato, Tamaulipas, Veracruz, Guerrero, Oaxaca,

Campeche y Yucatán) y 9 grupos indígenas (Tarahumara, Cora, Huichol, Southern Tepehuán, Otomí, Mixe, Mixteco, Zapoteco y Maya).

González-Oliver *et al.*, (2017). Estudio con 250 individuos de dos grupos indígenas del Estado de Méxicio, 149 de ellos Mazahua y 101 Otomíes.

Tabla 1. Frecuencia de haplogrupos nativo americanos en población mexicana. Se muestra la frecuencia (%) de haplogrupos nativo americanos, europeos y africanos reportados en distintos trabajos para distintas poblaciones de México. (*) = Se muestran las frecuencias reportadas para población Mazahua y Otomí respectivamente separadas por una diagonal. (-) = Ausencia de datos.

Trabajo reportado	Haplogrupo nativo americano				Haplogrupos Europeos	Haplogrupos Africanos
	A	B	C	D		
Green <i>et al.</i> , (2000)	33.6	26.5	23.3	5.8	5.4	4.5
Peñaloza <i>et al.</i> , (2007)	50.7	28.8	11.7	7.4	0.8	0.6
Sandoval <i>et al.</i> , (2009)	50.5	17.6	28.5	3.1	0.2	-
Guardado <i>et al.</i> , (2009)	51.1	17.8	18.5	5.9	5.9	0.8
Gorostiza <i>et al.</i> , (2012)	49.9	20.8	20.1	8.2	0.8	-
Martínez <i>et al.</i> , (2013)	47.0	23.7	15.9	6.2	5.3	1.9
González-Martín <i>et al.</i> , (2015)	49.4	24.4	19.4	6.7	-	-
González-Sobrino (2016)	49.8	24.3	15.4	7.4	1.8	1.2
*González-Oliver <i>et al.</i> , (2017)	31.5/57.4	35.0/20.8	13.4/14.8	20.1/7.0	-/-	-/-

El haplogrupo A es el más frecuente en población mexicana (31.5-57.4%). En segundo lugar, se encuentra el haplogrupo B (17.6-35%), destacando una mayor frecuencia en el norte y poblaciones indígenas del centro del país. El haplogrupo C tiene una frecuencia similar (11.7-28.5%), siendo mayor en poblaciones indígenas y población del norte del país. Finalmente, el haplogrupo D es el de menor frecuencia en México (3.1-20.1%). La frecuencia de los haplogrupos europeos, es menor al 1% en poblaciones indígenas, mientras que en poblaciones mestizas va de 1.8-5.9%; en cuanto a los haplogrupos africanos, aunque su frecuencia es baja México, es mayor en poblaciones mestizas (0.8-4.5%) respecto a lo reportado en poblaciones indígenas que va de 0 a 0.6% (Tabla 1).

Evolución molecular

La evolución molecular permite elucidar los procesos que generan cambios y adaptaciones constantes en las secuencias de DNA, RNA o las proteínas a lo largo del tiempo. El estudio de la adaptación molecular se ha realizado a partir del estudio de la distribución de polimorfismos en secuencias de DNA y mediante la determinación de sustituciones sinónimas y no sinónimas en secuencias codificantes de DNA, utilizando el codón como unidad mínima evolutiva (Eguiarte *et al.*, 2007). Ambas aproximaciones utilizan las predicciones de la teoría neutral, es por ello que es necesario conocer las bases de dicha teoría para el entendimiento de la evolución molecular.

La “Teoría neutral” propuesta por Motoo Kimura en 1968, pasa de una interpretación seleccionista de la evolución, a una en donde la mayor parte del cambio no es adaptativo.

Supone que la probabilidad de remplazo de un nucleótido por otro en una secuencia, sólo dependerá de su tasa de mutación (Perfectti *et al.*, 2009). De esta manera, propone que la mayoría de las mutaciones que surgen tienen una mínima o nula ventaja adaptativa, lo que las convierte en mutaciones "selectivamente neutrales". Bajo este supuesto, la selección natural sería una fuerza evolutiva que no actúa en la fijación o pérdida de dichas mutaciones, y se considera a la deriva génica como la principal fuerza evolutiva que lleva a cabo dicha tarea. El modelo de la teoría neutral permite la predicción de la cantidad de **polimorfismos** dentro de una población de una especie y el grado o la tasa de **divergencia** entre secuencias muestreadas de especies diferentes (Hamilton, 2009).

Para la caracterización de los de polimorfismos en la secuencia del DNA en una especie se utilizan los parámetros θ y π :

Número de sitios segregantes (S). Un sitio segregante es cualquier *locus* del DNA que tiene dos o más nucleótidos presentes dentro de la población; la suma total de los sitios segregantes se le denomina **S**. Teta (θ) es un valor calculado a partir de la expresión $\theta = 4N_e\mu$ (Kimura, 1968), en donde N_e es el tamaño efectivo poblacional y μ es la tasa de mutación). Debido a que dichos factores son difíciles de determinar de manera exacta, θ también puede determinarse de manera indirecta bajo modelos de neutralidad utilizando el total de sitios segregantes en un grupo de secuencias a partir de la siguiente fórmula: $\theta_w = \frac{K}{a}$ (Watterson, 1975) (donde **K** es igual al número de sitios segregantes en una muestra de secuencias (**S**) y $a = \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n-1}$ (n =número de secuencias en la muestra) (Eguiarte *et al.*, 2007).

Diversidad nucleotídica ($\pi = pi$). Es el número de nucleótidos diferentes por sitio entre dos secuencias tomadas al azar, es decir, la proporción de diferencias nucleotídicas entre todos los pares de secuencias valoradas con la frecuencia de las secuencias (Eguiarte *et al.*, 2007).

El valor de π es sensible a la frecuencia de los alelos o polimorfismos de cada secuencia en una muestra e independiente del tamaño de la muestra, mientras que θ sí depende del tamaño muestral y de los alelos o polimorfismos poco frecuentes (Hamilton, 2009). La relación que existe entre estos dos estimadores permite determinar si las secuencias se encuentran bajo el modelo neutral o se desvían de él (Eguiarte *et al.*, 2007).

Para determinar si la frecuencia de los polimorfismos encontrados en un grupo de secuencias se explica mediante la teoría neutral, se utiliza el estadístico **D**. Fumio Tajima desarrolló un *test* basado en las diferencias de π (proporción de diferencias nucleotídicas entre las secuencias comparadas por pares) y θ (total de sitios polimórficos en el grupo de secuencias) (Perfectti *et al.*, 2009); el cual se basa en la siguiente fórmula: $D = \frac{\pi - \theta}{\sqrt{V(\pi - \theta)}}$ en donde V es una estimación de la varianza muestral de la diferencia entre π y θ que Tajima obtuvo bajo el modelo de sitios infinitos (Templeton, 2006). Si π y θ son iguales, **D** será igual a cero, indicando que los polimorfismos observados se encuentran bajo equilibrio neutral. Si θ es mayor a π , **D** será negativa, lo que indica mutaciones deletéreas y si π es mayor a θ , **D** será positiva, sinónimo de

que algunos polimorfismos se encuentran bajo selección positiva incrementando sus frecuencias (Eguiarte *et al.*, 2007).

Cuando no existe recombinación en las secuencias (tal es el caso del mtDNA) un polimorfismo nuevo, genera la fijación del mismo y de todos los otros sitios; a esto se le llama barrido selectivo. Posterior a ese barrido selectivo, cada nueva mutación crea un sitio de segregación bajo el modelo de sitios infinitos e inicialmente esas mutaciones serán raras y contribuirán poco a la heterocigosidad media; una D negativa de un *Test* de Tajima estaría describiendo dicha situación. Por el contrario, si la selección natural se encuentra favoreciendo a determinados polimorfismos, existirán haplotipos con mutaciones de frecuencias intermedias más allá de las expectativas neutrales, los cuales contribuirán a la heterocigosidad media pero no significarán nuevos sitios de segregación por lo que el resultado de un *Test* de Tajima será positivo (Eguiarte *et al.*, 2007; Templeton, 2006). Es importante destacar que el valor D de Tajima es sensible a cambios en el tamaño de las poblaciones a lo largo del tiempo, a la estructura poblacional y a la selección natural (Hamilton, 2009).

Otro parámetro que se utiliza para determinar variabilidad genética en estudios poblacionales es la **diversidad haplotípica (H_d)**. La diversidad haplotípica es la probabilidad de que dos haplotipos tomados al azar en una población sean diferentes, mide la singularidad de un haplotipo determinado en una población dada en relación con el tamaño muestral y está dada por la fórmula: $H_d = 1 - \sum_{i=1}^H p_i^2$ (donde H es el número de haplotipos y p_i es la frecuencia del haplotipo) (Depaulis, 1998).

Estimación de flujo génico en una población

Todos los mecanismos que generen movimiento de genes de una población a otra es a lo que se le denomina flujo génico; el cual se presenta en mayor o menor grado debido al movimiento de los individuos o los gametos entre las poblaciones, eventos de extinción, colonización o recolonización de las poblaciones. El estudio del flujo génico permite conocer la estructura de las poblaciones para así determinar si poblaciones que se encuentran fragmentadas se comportan como metapoblaciones. Para mediar las tasas de flujo génico entre las poblaciones se utilizan métodos que se basan en datos moleculares, los cuales observan la distribución espacial de alelos o polimorfismos en las poblaciones, para después hacer inferencias de los niveles o patrones de flujo génico entre las poblaciones (Eguiarte *et al.*, 2007).

Un modelo utilizado para estimar el flujo génico es el modelo de islas infinitas, el cual considera subpoblaciones de igual tamaño que intercambian migrantes en una tasa constante. Debido a que las subpoblaciones son iguales, el modelo puede ser caracterizado con el tamaño poblacional (N) y la tasa de migración (m). Con la finalidad de dividir el coeficiente de endogamia en una población subdividida, en 1951 Wright introdujo un coeficiente de correlación entre la endogamia y la subdivisión entre las poblaciones (F_{st}). El F_{st} mide la variación de las frecuencias alélicas entre poblaciones, es decir, la diferenciación genética entre

ellas; el cual está dado por la fórmula $F_{st} = \frac{H_t - H_s}{H_t}$ donde H_t es la heterocigocidad esperada y H_s es la frecuencia observada de los heterocigotos en el total de la población (Eguiarte *et al.*, 2007; Templeton, 2006).

El estadístico F_{st} proporciona un estimado del balance de variabilidad genética en una población y entre poblaciones; es un estimador imparcial de la divergencia genética entre pares de poblaciones bajo un modelo de isla en el que todas las poblaciones divergieron al mismo tiempo y están unidas por similares tasas de migración. Dicho coeficiente tiene un rango de 0 a 1, donde 0 significa que comparten la información genética y 1 que no comparten la información genética (Templeton, 2006). Cuando la diferencia en las F_{st} es significativa, se dice que existe estructura entre las poblaciones analizadas, es decir, no existe flujo genético entre ellas, situación que genera poblaciones diferenciadas genéticamente.

La tasa de fijación de un alelo en una determinada población es inversamente proporcional al tamaño efectivo de la población (N), esta tasa puede ser contrarrestada por el flujo génico de otras poblaciones con una tasa m . A partir del valor de F_{st} se puede hacer una estimación indirecta de la tasa de migración (Nm) (flujo génico) necesaria para contrarrestar la diferenciación genética debida a la deriva génica mediante la ecuación $F_{st} = \frac{1}{2Nm+1}$ (para el caso de genomas diploides), despejando la ecuación se tiene que $Nm = \frac{1}{2} \left(\frac{1}{F_{st}} - 1 \right)$. Si el valor de Nm es mucho mayor que 1, significa que el flujo génico sobrepasa los efectos de la deriva genética, mientras que, si es menor de 1, la deriva actúa de forma independiente en la población (Eguiarte *et al.*, 2007; Slatkin, 1994).

Otro modelo para determinar si existe flujo génico dentro de una población es el de aislamiento por distancia, el cual busca determinar el grado de diferenciación genética entre dos poblaciones situadas en diferentes puntos geográficos. Para ello se utilizan las distancias genéticas entre las poblaciones (diferenciación genética entre las poblaciones). Uno de los modelos para medir la distancia genética entre las poblaciones mayormente utilizadas es el de F_{st} pareadas (cálculo de la F_{st} aplicado únicamente a dos poblaciones; en donde la población total que ahora se usa para calcular H_t se refiere sólo a las dos poblaciones de comparadas). A partir del cálculo de las F_{st} pareadas, se obtiene una matriz de distancias genéticas, la cual debe ser correlacionada con una matriz de distancias geográficas entre los mismos pares de poblaciones (Templeton, 2006). Para determinar si existe una relación lineal significativa entre el flujo génico y la distancia geográfica entre las poblaciones, se utiliza la prueba de Mantel (Eguiarte *et al.*, 2007).

Relaciones filogenéticas en el mtDNA

El mtDNA se ha convertido en una herramienta importante en el estudio de la historia y estructura poblacional humana, sin embargo, su interpretación en la mayoría de los casos resulta compleja (Bandelt *et al.*, 1995) La región control presenta una gran variabilidad debida a la alta tasa de mutación del genoma mitocondrial, aunado a que se trata de una región no

codificante del genoma que permite la acumulación de un gran número de mutaciones. A pesar de que cualquier posición de la región control es susceptible a mutar, algunas de ellas son extremadamente variables entre los linajes, mientras que otras se mantienen constantes. Durante la tipificación del mtDNA es difícil determinar si las mutaciones que se observan se deben a un cambio ancestral que ha sido heredado en el linaje o porque se produjeron de manera independiente en una línea no relacionada (homoplasia) (Fernández, 2000).

La mayoría de los estudios filogenéticos realizados con mtDNA que utilizan polimorfismos tipo SNP o secuencias de la región control, fallan en la formación de conjuntos anidados de haplotipos debido a que encuentran incompatibilidad entre los pares de nucleótidos (Bandelt *et al.*, 1995). Los análisis filogenéticos como el de máxima parsimonia (MP) o el de diferencias en pares de nucleótidos, otorgan el mismo peso a todas las posiciones, por lo que van a agrupar secuencias que comparten posiciones entre sí creando grupos polifiléticos. Sin embargo, en el análisis del mtDNA se podrían estar separando variantes estrechamente relacionadas, ya que es imposible determinar el estatus ancestral del haplotipo, debido a las homoplasias (ya sea por eventos de mutación paralela o reversiones) sumado al bajo número de caracteres informativos que se analizan (Fernández, 2000).

Ejemplo de ello es el trabajo de Excoffier y Smouse (1994), con datos de Polimorfismos de Longitud en Fragmentos de Restricción (RFLP) con 56 haplotipos; en este se calculó el número de árboles igualmente parsimoniosos, teniendo como resultado más de mil millones de árboles. En la práctica presentar un único árbol resultado de un análisis de máxima parsimonia elegido al azar en el que los caracteres se han resuelto de manera arbitraria, puede resultar tan engañoso como formar un consenso estricto de dos o más árboles en los que las verdaderas politomías no se pueden distinguir de ambigüedades (Bandelt *et al.*, 1995).

Debido a esto, Bandelt *et al.* (1995) desarrollaron un método de reconstrucción filogenética mediante un enfoque de red estricto que distingue los conflictos de cada carácter generando una representación sencilla y entendible. A este método de análisis le denominaron "Redes Medianas" ("*Median Networks*"), las cuales se generan a partir de la división de los grupos de haplotipos carácter por carácter, garantizando así incluir a todos los árboles más parsimoniosos del conjunto de datos. A diferencia de los métodos de construcción de árboles, las redes de haplotipos permiten interpretar homoplasias (altas tasas de homoplasia pueden conducir a que un solo haplotipo derive independientemente a lo largo de diferentes rutas a partir de un mismo ancestro). Las redes medianas o de haplotipos, también permiten visualizar la frecuencia de los haplotipos y exhibir rutas evolutivas probables a través de la misma red (Bandelt *et al.*, 1995).

Cuatro años después, Bandelt y colaboradores generaron un método de análisis para las redes medianas denominado "median joining" (MJ), en el cual se busca construir redes que combinan las características del algoritmo Kruskal para encontrar árboles de expansión mínima al favorecer conexiones y el algoritmo heurístico de máxima parsimonia (MP), el cual agrega secuencialmente nuevos vértices llamados "vectores medios", teniendo como resultado una

red que es aplicable también a caracteres multi estado (secuencias de aminoácidos). En el método MJ, sólo los vectores medios que tienen una alta posibilidad de aparecer como nodos de ramificación en un árbol de máxima parsimonia son generados considerando tripletes de secuencias. Estos vectores medios se pueden interpretar biológicamente como secuencias no muestreadas o secuencias ancestrales extintas (Bandelt *et al.*, 1999).

Inferencia de aspectos históricos de las poblaciones

El parámetro π , además de utilizarse como un promedio para medir los polimorfismos, permite examinar la distribución de la comparación de todos los pares de secuencias; a esto se le conoce como ***mismatch distribution***. Se trata de la distribución de la frecuencia del número de diferencias nucleotídicas entre todos los pares de secuencias en una muestra. Dicho análisis resulta útil para la estimación de la historia de una población a partir de una muestra de secuencias de DNA, por ejemplo, para la inferencia en los tamaños efectivos de una población mediante el uso de secuencias de DNA o en aquellas poblaciones que han mantenido un tamaño constante, su distribución puede utilizarse para identificar la acción de la selección natural (Hamilton, 2009).

El resultado de un análisis de *mismatch distribution* permite la construcción de un histograma, el cual se compara con el resultado de un modelo de una población en expansión repentina; y con ayuda de una prueba de χ -cuadrada se compara si la diferencia entre ambos histogramas es significativa o no, determinando así, si existen o no crecimientos demográficos en la población analizada (Eguiarte *et al.*, 2007). Para la construcción del modelo de una población en expansión, se requiere de los parámetros: θ_0 (Theta inicial): valor de theta antes de la expansión o disminución poblacional y θ_1 (Theta final): valor de theta después de la expansión o disminución poblacional y τ (Tau): fecha de la expansión o disminución poblacional medida en unidades de tiempo mutacional ($\tau = 2\mu t$, donde μ es la tasa de mutación por secuencia y por generación y t es el tiempo en generaciones) (Rogers y Harpending, 1992).

Determinación de la riqueza haplotípica de la población

La relación entre el número de OTUs (unidades taxonómicas operativas) observados y el esfuerzo de muestreo en un estudio, proporcionan información sobre el total de la diversidad del OTU en la comunidad o población que ha sido muestreada; si se correlaciona el número acumulativo de OTUs observados contra el esfuerzo de muestreo, se obtiene una curva de acumulación (Eguiarte *et al.*, 2007). Estas curvas alcanzan una asíntota cuando la probabilidad de añadir un nuevo OTU en el estudio alcanza finalmente el cero; en este punto se asume que la asíntota equivale a la riqueza real, sujeta al error de muestreo (incertidumbre) derivado del ajuste a los datos observados. Sin embargo, no son asíntotas cuando dicha probabilidad nunca se hace nula (González-Oreja *et al.*, 2010).

En las curvas de acumulación, al principio del muestreo se colectan OTUs comunes, y la adición de OTUs se produce rápidamente, por lo que la pendiente de la curva comienza siendo elevada;

a medida que prosigue el muestreo, comienzan a aparecer los OTUs raros y la pendiente comienza a descender hasta llegar a cero. En ese momento, es cuando teóricamente se ha muestreado el total de OTUs de la comunidad o población estudiada (Jiménez y Hortal, 2003). También se pueden realizar extrapolaciones a partir de curvas de acumulación, las cuales modelan el proceso de observar nuevos OTUs si se aumenta el esfuerzo de muestreo (Eguiarte *et al.*, 2007). Estos últimos análisis son útiles cuando se desea estimar la riqueza del OTUs estudiado o cuan representativa puede llegar a ser la muestra con la que se realiza determinado estudio.

JUSTIFICACIÓN

La Genética Forense es una disciplina de las Ciencias Forenses que se ocupa del estudio de la herencia biológica aplicada a problemas de orden legal, mediante el análisis de polimorfismos o variabilidad genética contenida en el DNA. En primera instancia, se busca la obtención de perfiles a partir del DNA nuclear, ya que este material genético es heredado por ambos progenitores y se encuentra sujeto a la recombinación, situación por la cual es distinto en cada individuo (exceptuando gemelos monocigóticos) y facilita la individualización de las personas. Por su parte, el mtDNA aunque se encuentra en mayores cantidades que el DNA nuclear en la célula, se hereda únicamente vía materna, lo que imposibilita individualizar personas y permite la identificación únicamente de haplogrupos y haplotipos. Esto no significa que carezca de valor como prueba pericial, al contrario, resulta de suma importancia para muestras que se encuentran altamente degradadas en las que el análisis de marcadores nucleares falla.

Una de las mayores problemáticas que enfrentan los estudios de mtDNA para la impartición de justicia en la Ciudad de México es que no se cuenta con estudios poblacionales de las frecuencias de los haplotipos y haplogrupos de mtDNA específicos para la población previamente mencionada, lo que genera una limitante al momento de dar soporte estadístico a las pruebas que se generen a partir de mtDNA. Los estudios de mtDNA de poblaciones humanas en México son pocos, la mayoría vinculada a poblaciones culturalmente diferenciadas (Peñaloza *et al.*, 2007; Sandoval *et al.*, 2009; Gorostiza *et al.*, 2012; Martínez *et al.*, 2013; González-Martin *et al.*, 2015; Raghavan *et al.*, 2015; González-Sobrino *et al.*, 2016; González-Oliver *et al.*, 2017), y un sólo estudio en poblaciones mestizas mexicanas que involucra a la Ciudad de México (Guardado *et al.*, 2009). El número muestral de habitantes de la Ciudad de México de ésta última publicación es de 100 individuos, dichas muestras se obtuvieron en un hospital de la zona centro de la ciudad, a partir de mujeres asistentes a un programa de colposcopia como parte de un monitoreo de salud. En el estudio realizado no se reporta el lugar específico (zona centro, norte, sur, poniente u oriente de la ciudad; o bien la alcaldía) en el que habitan las donantes de muestra; razón por la cual se desconoce si dicho muestreo cubre de manera homogénea la densidad poblacional de la ciudad y por lo tanto resulte representativa para dicha ciudad; tampoco se conoce el lugar de nacimiento de sus ancestros maternos, con el cual se pueda hacer un seguimiento del origen de los haplotipos y haplogrupos encontrados en la población de estudio.

La Ciudad de México cuenta con 8,918,653 habitantes (INEGI, 2015), los cuales no se encuentran distribuidos homogéneamente en el territorio, por ejemplo, alcaldías como Iztapalapa o Gustavo A. Madero cuentan con una mayor densidad poblacional comparada con la alcaldía Milpa Alta; por esta razón es importante realizar los muestreos conforme a la densidad poblacional de cada alcaldía para muestrear lo más homogéneamente dicha población. De ahí, el interés de la generación de estudios poblacionales de variabilidad nucleotídica y haplotípica de mtDNA específicos y homogéneos para la población de la Ciudad

de México, que soporten estadísticamente aquellos trabajos de identificación humana realizados con mtDNA.

METODOLOGÍA

POBLACIÓN DE ESTUDIO

Determinación de la muestra poblacional

El tamaño muestral del presente estudio fue determinado por los parámetros mínimos aceptados por la revista *Forensic Science International: Genetics*, que reportan en la publicación *Revised guidelines for the publication of genetic population data* del año 2017; en la cual se determina que el número mínimo muestral para estudios poblacionales que analicen la región control del mtDNA es de 200 muestras.

Debido a que la Ciudad de México cuenta con 8,918,653 habitantes (INEGI, 2015), se tomó un valor mayor al mínimo aceptado ($n=200$), por la revista anteriormente mencionada, para trabajos de caracterización de la región control del mtDNA; teniendo así una n muestral de 250 individuos. Para evitar sesgos de muestreo por zonas, se determinó un número mínimo de muestras por alcaldía de acuerdo a la densidad poblacional de cada una de ellas. Para determinar el número de individuos a muestrear por alcaldía, se calculó que porcentaje de los habitantes totales es representado por alcaldía y posteriormente se calculó el número de muestras por alcaldía basado en el porcentaje de habitantes de cada una con respecto al total de la n muestral; en la Tabla 2 se muestran los valores anteriormente mencionados.

Los criterios de inclusión considerados para seleccionar a los individuos fueron los siguientes:

- Haber nacido en la Ciudad de México (antes Distrito Federal).
- Vivir actualmente en alguna de las alcaldías (antes delegaciones) de la Ciudad de México.
- No estar emparentado biológicamente vía materna con algún otro individuo que participe en el estudio.
- Firmar el formato de consentimiento informado (mayores de edad). En caso de ser menor de edad, un tutor legal (madre/padre) debió haber firmado el formato de consentimiento informado.

Tabla 2. Muestra poblacional. Se muestran los datos por alcaldía del número de habitantes reportados en el 2015 por el INEGI, el porcentaje de habitantes por alcaldía que representa del total de la población y el número de muestras calculado para el muestreo del presente estudio.

Alcaldía	Habitantes	Porcentaje (%)	No. de muestras
Azcapotzalco	400,161	4.48	11
Coyoacán	608,479	6.82	17
Cuajimalpa de Morelos	199,224	2.23	6
Gustavo A. Madero	1,164,477	13.05	33
Iztacalco	390,348	4.37	11
Iztapalapa	1,827,868	20.49	51
Magdalena Contreras	243,886	2.73	7
Milpa Alta	137,927	1.54	4
Álvaro Obregón	749,982	8.4	21
Tláhuac	361,593	4.05	10
Tlalpan	677,104	7.59	19
Xochimilco	415,933	4.66	12
Benito Juárez	417,416	4.68	12
Cuauhtémoc	532,553	5.97	15
Miguel Hidalgo	364,439	4.08	10
Venustiano Carranza	427,263	4.79	12
TOTAL	8,918,653	100%	251

Obtención de las muestras biológicas

Para llevar a cabo la toma de muestra se les solicitó a los participantes firmar un formato de consentimiento informado (Anexo 1) en el cual aceptaban donar de manera voluntaria y sin fines de lucro la muestra biológica, así como también se les hacía de su conocimiento la confidencialidad de la información que proporcionarían para el estudio. De igual manera se les solicitó que respondieran a un cuestionario sobre datos propios del donante como nombre, edad y alcaldía en la que habita, así como también datos sobre su ancestro materno más lejano del que tuvieran conocimiento (Anexo 1), esto con el propósito de poder correlacionar los resultados de ancestría del donante con los de su ancestro materno.

La toma de muestra biológica consistió en un raspado de células de descamación bucal obtenido con un hisopo nuevo y estéril; el cual fue rotulado del puño y letra del donante con su nombre, su firma y la fecha del día de la toma de muestra. La muestra se tomó por duplicado.

PROCESAMIENTO DEL mtDNA

Previo al procesamiento de las muestras, los hisopos fueron puestos a secar en un desecador con humedad relativa de ~10% por al menos 12 horas o hasta que el algodón del mismo se encontraba totalmente seco. Para conservar los datos de confidencialidad del donante, a cada muestra se le asignó un número de control interno con el cual fue reconocida durante todo el presente estudio. Dicho número fue consecutivo conforme se fueron obteniendo las muestras (001-251).

Extracción de DNA

La extracción de DNA se estandarizó a partir de la modificación al método de sílica empleando una solución de unión con tiocianato de guanidina (GuSCN) reportado para realizar extracción de huesos por Rohland & Hofreiter (2007). El procedimiento estandarizado fue el siguiente:

Para llevar a cabo la extracción del DNA se cortó media cabeza del hisopo de manera longitudinal y fue colocada en un tubo de microcentrifuga de 1.7 ml. Se agregó 400 µl de Buffer de unión (GuSCN 5M, NaCl 25mM, Tris 50mM, pH 8.0) (Rohland & Hofreiter, 2007), 10 µl de proteinasa K [10 mg/ ml] y 4 µl de DTT 1M; se realizó un lisado durante 30 minutos a 56°C a 1100 rpm.

Finalizada la lisis se retiró el algodón del hisopo y se adicionaron 10 µl de sílica gel (Rohland & Hofreiter, 2007); la muestra fue homogeneizada en vórtex y posteriormente centrifugada a 10,000 rpm durante un minuto. Se retiró el sobrenadante y se lavó el material genético con 400 µl de etanol al 70%, se homogeneizó y centrifugó la muestra bajo las condiciones anteriores. Se retiró el sobrenadante y se colocó en baño seco a 60°C por 5 min para eliminar por completo el etanol; una vez seco el botón de sílica, se adicionaron 100 µl de buffer PE (QUIAGEN). La muestra se homogeneizó con un vórtex y se incubó por 5 min. a 60°C para posteriormente centrifugar un minuto a 13,500 rpm. El sobrenadante (DNA purificado) fue transferido a un tubo nuevo y almacenado a 4°C para su posterior procesamiento.

Amplificación de mtDNA

La estandarización del protocolo de la amplificación de mtDNA, fue basada en las recomendaciones del "Protocolo de análisis de DNA mitocondrial" (FBI, 2008), así como en las especificaciones de la polimerasa Platinum™ SuperFi™ Green PCR Master Mix (Invitrogen, 2017). Para amplificar la región control (RC) del mtDNA se utilizaron los primers:

A1 (15997) 5' **CAC CAT TAG CAC CCA AAG CT** 3' (FBI,2008)

D1 (409) 5' **CTG TTA AAA GTG CAT ACC GCC** 3' (FBI,2008)

Se realizó un ensayo en el cual se pusieron a prueba distintas concentraciones de DNA molde inicial, así como distintas temperaturas de alineamiento, esto con la finalidad de encontrar la combinación de ambos factores que permitieran la generación de un amplicon mayor (Fig. 7).

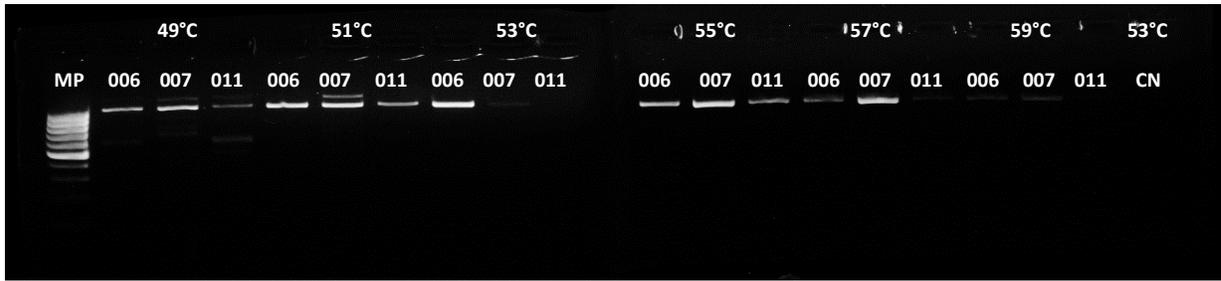
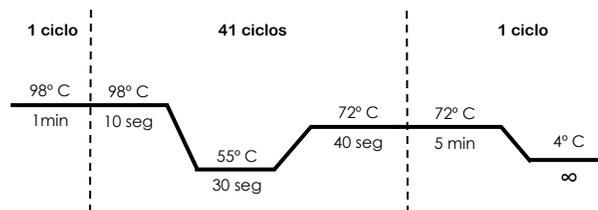


Fig. 7. Amplificados de la RC del mtDNA. Se muestran las bandas de amplificado de tres muestras con concentraciones diferentes de DNA: 006 [2.6 ng/ μ l], 007 [6.43 ng/ μ l] y 011 [25.91 ng/ μ l], las cuales fueron amplificadas a diferentes temperaturas de alineamiento (49°C, 51°C, 53°C, 55°C, 57°C y 59°C). El control negativo se amplificó con temperatura de alineamiento de 53°C.

Con base en el ensayo anterior, se establecieron los parámetros con los cuales se trabajaron todas las muestras: 6 μ l de Taq Platinum SuperFi Green PCR Master Mix, 2 μ l de primers [20pM/ μ l] y 4 μ l de templado de DNA [\sim 10ng/ μ l DNA nuclear], con las siguientes condiciones de termociclado:



Separación y purificación de fragmentos de mtDNA

Las muestras amplificadas se corrieron en geles de agarosa al 1.5%, la banda de amplificado de interés se cortó y purificó con el kit QIAquick Gel Extraction Kit (QIAGEN) siguiendo las especificaciones del “Protocol: QIAquick Gel Extraction Using a Microcentrifuge”. La cuantificación del material genético recuperado se llevó a cabo con espectrofotometría en el equipo NanoDrop™ 2000 (Thermo Scientific™).

Marcaje con terminadores fluorescentes y purificación de los fragmentos de mtDNA

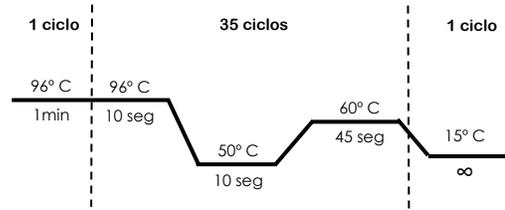
El protocolo de marcaje de las secuencias se estandarizó a partir de las recomendaciones de la Guía del usuario de BigDye™ Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, 2016) y del “Protocolo de análisis de DNA mitocondrial” (FBI, 2008). El procedimiento estandarizado fue el siguiente:

Los fragmentos purificados de la región control del mtDNA fueron marcados con terminadores fluorescentes mediante el kit BigDye™ Terminator v3.1 (Applied Biosystems) en la siguiente reacción de PCR: 1.8 μ l de BigDye Terminator v3.1 Ready Reaction Mix, 2.0 μ l de BigDye Terminator v1.1 & v3.1 5X Sequencing Buffer, 2.0 μ l de primer [5 pM/ μ l] y 40 ng de templado de mtDNA. El volumen final de la reacción fue de 12 μ l.

A continuación, se muestran los primers utilizados (FBI,2008) y las condiciones de termociclado utilizadas para el marcaje:

A1 (15997) 5' **CAC CAT TAG CAC CCA AAG CT** 3'
A2 (16159) 5' **TAC TTG ACC ACC TGT AGT AC** 3'
B1*(16391) 5' **GAG GAT GGT GGT CAA GGG AC** 3'
C1 (048) 5' **CTC ACG GGA GCT CTC CAT GC** 3'
D2* (285) 5' **GGG GTT TGG TGG AAA TTT TTT G** 3'
D1* (409) 5' **CTG TTA AAA GTG CAT ACC GCC** 3'

***Primer Reverso**



Los fragmentos de DNA marcados se purificaron mediante el Kit DyeEx® 2.0 Spin Kit, siguiendo las instrucciones del Quick-Start Protocol del mismo.

Secuenciación del mtDNA

Una vez marcados y purificados los fragmentos de la región control del mtDNA, se resuspendieron en 12 µl de Hi-Di™ Formamide. Las muestras fueron secuenciadas en el Analizador Genético 3500 (Applied Biosystems) con POP-7 y capilar de 50 cm.

ANÁLISIS DE SECUENCIAS

Obtención del haplotipo y haplogrupo de mtDNA

Una vez obtenidas las secuencias, se analizaron con el software BioEdit Sequence Alignment Editor (versión 7.2.6.1). Con la finalidad de determinar las mutaciones que presentaba cada muestra, las secuencias se compararon con la Secuencia de Referencia de Cambridge (NCBI Sequence Reference NC_012920.1) revisada en 2014. Con base en las mutaciones (haplotipo) que se obtuvieron para cada muestra, se determinó el haplogrupo de mtDNA al que pertenecían de acuerdo a los "Marcadores de Haplogrupo" reportados en la Base de Datos del Genoma Mitocondrial Humano (MITOMAP) (Lott, 2019) y las estimaciones del portal mtDNAMANAGER (Young *et al.*, 2007).

Determinación de los parámetros de diversidad genética

Para la obtención de los parámetros de diversidad genética como el número de mutaciones, sitios polimórficos, diversidad nucleotídica y haplotípica, así como el test *D* de Tajima se utilizó el programa DnaSP v.6 (Rozas *et al.*, 2017).

Determinación de estimadores de flujo génico

Para la obtención de los estimadores de flujo génico dentro de la población, se calcularon los valores de F_{st} para cada delegación, F_{st} pareadas entre las alcaldías y distancias genéticas entre las poblaciones con el software Arlequin 3.1, el valor de F_{st} general y N_m se obtuvo en DnaSP v.6. (Rozas *et al.*, 2017).

Se realizó una prueba de Mantel con la finalidad de determinar si entre las alcaldías existía aislamiento por distancia. La matriz de los valores de F_{st} pareados (distancia genética), se obtuvo en el software Arlequin 3.1; mientras que la matriz de distancias geográficas se obtuvo mediante el software Geographic Distance Matrix Calculator v. 1.2.3. Las coordenadas utilizadas para cada una de las alcaldías, fueron las reportadas como "Geo Point" en el "Portal de datos de la Ciudad de México" (Gobierno de la Ciudad de México, 2019). La prueba de Mantel fue realizada con el software XLSTAT (Addinsoft, 2019).

Estimación de las relaciones filogenéticas

Para estimar las relaciones filogenéticas de la región control de mtDNA se realizaron redes de haplotipos o "Median Networks" con el software Network 5.0.1.1 utilizando el algoritmo median joining (MJ).

Determinación de cambios en el tamaño poblacional

Se realizó un análisis de *mismatch distribution* para determinar el comportamiento de la población estudiada respecto a los cambios de tamaño, para ello se utilizó el software Arlequin 3.1 en el cual se obtuvieron los valores de Tau, Theta 0, Theta 1 así como la prueba de χ -cuadrada, el gráfico de dicho análisis se obtuvo mediante el software DnaSP v.6. (Rozas *et al.*, 2017).

Determinación de la riqueza haplotípica de la población

Debido a que el presente trabajo busca ser considerado como una base de datos de frecuencias de mtDNA para trabajos de identificación forense, se realizó una curva de acumulación de haplotipos con el número de muestras del presente estudio (251), así como una extrapolación de la curva con un tamaño de muestra más grande (6000) para determinar con qué tamaño de muestra se cubriría la totalidad o la gran mayoría de los haplotipos estimados para dicha población. Para ello se utilizó el software EstimateS versión 9.1.0.

RESULTADOS

Muestra poblacional

Se obtuvo una n total de 251 individuos habitantes de la CDMX muestreados de acuerdo a la n estimada para cada una de las alcaldías de dicha ciudad. Participaron un total de 166 mujeres y 85 hombres con un rango de edad de entre 1 y 76 años de edad. Del total de los individuos muestreados, el 93% dieron información correcta sobre el lugar de nacimiento de su ancestro materno más lejano del que tuvieran conocimiento (ANEXO 2).

En la Fig. 8 se muestra el lugar de origen de los ancestros maternos de los individuos participantes en el estudio, así como el número de individuos que pertenecen a cada lugar; mientras que en la Fig. 9 se muestra el total de los ancestros de acuerdo a las generaciones anteriores (madre, abuela, bisabuela, etc.) respecto al donante de muestra, de acuerdo al lugar de origen de cada uno de ellos.

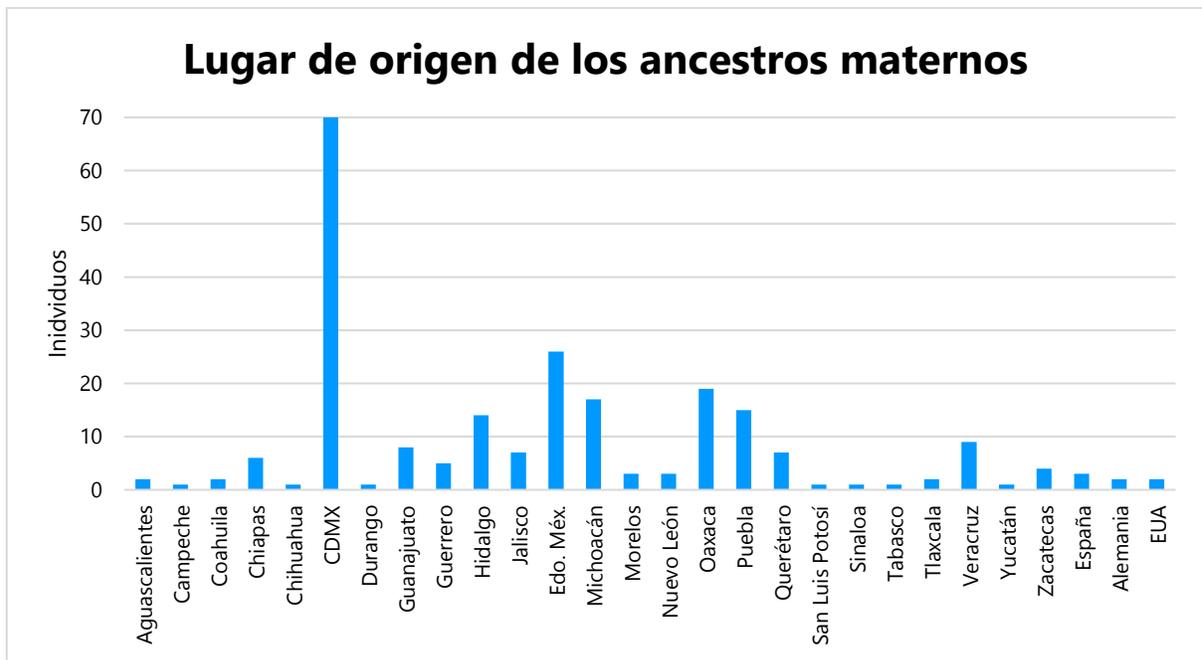


Fig. 8. Lugar de origen de los ancestros maternos. Se muestra el estado de la república mexicana o país de origen al que pertenecen los ancestros maternos, de las personas involucradas en el presente estudio.

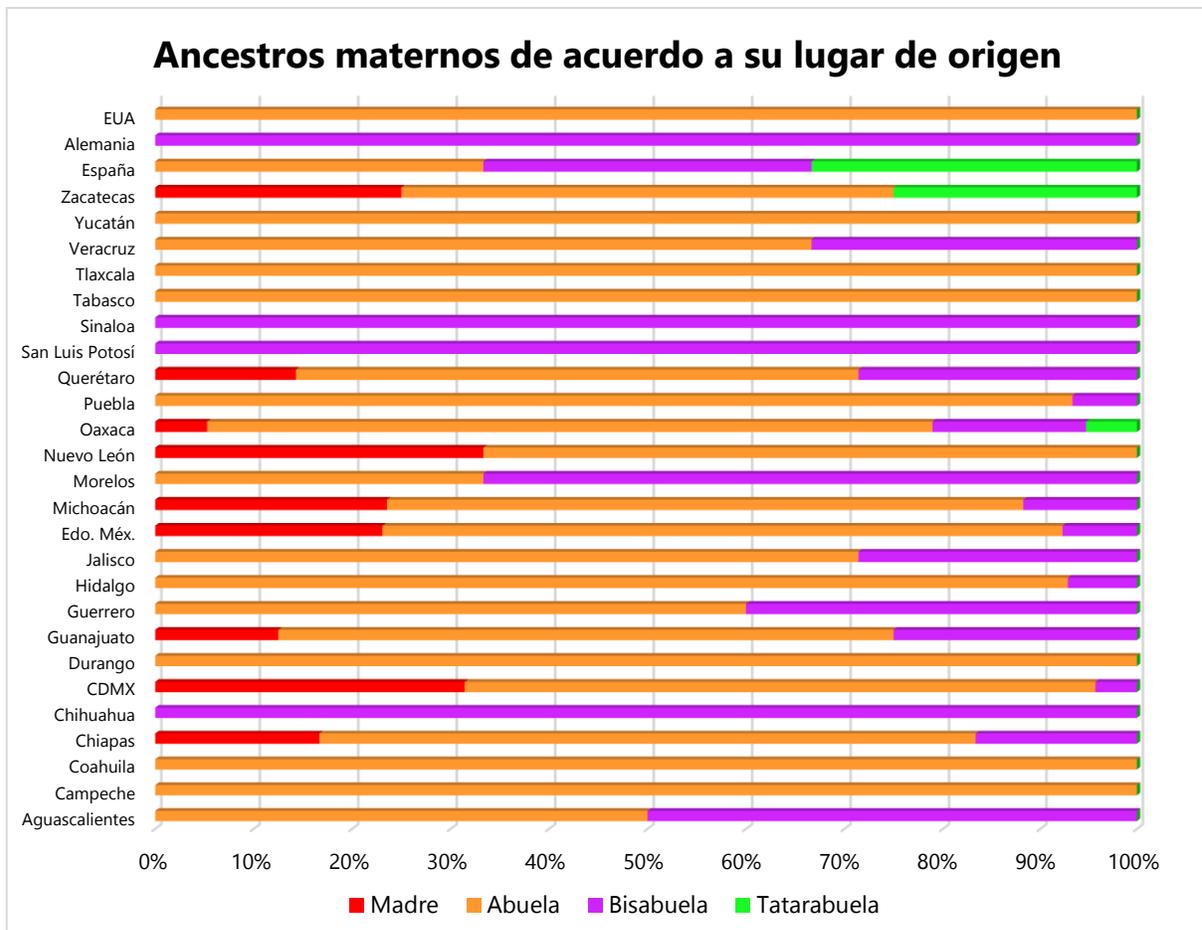


Fig. 9. Ancestros maternos de acuerdo a su lugar de origen. Se muestra el porcentaje de ancestros maternos según su generación (madre, abuela, bisabuela o tatarabuela) de acuerdo a su lugar de origen (estado de la república mexicana o país), de las personas involucradas en el presente estudio.

Caracterización de haplotipos y haplogrupos

Se realizó la caracterización del haplotipo y haplogrupo de un fragmento de 917 pb de la región control del mtDNA de las 251 muestras del presente estudio.

Se obtuvieron 208 haplotipos, los cuales se encuentran en el Anexo 3; en la Tabla 3 se muestran los haplogrupos obtenidos en la población de la CDMX, el número de muestras pertenecientes a cada uno de ellos y el porcentaje del total de la muestra que representa cada haplogrupo. En la Fig. 10 se encuentra graficada la distribución de los haplogrupos en la CDMX.

Tabla 3. Haplogrupos de mtDNA en la CDMX. Se muestran los haplogrupos de mtDNA obtenidos, el origen al cual pertenecen, el número de muestras que presenta cada haplogrupo, así como el porcentaje del total de la muestra que representa cada uno de ellos.

Ancestría	Haplogrupo	No. de muestras	Porcentaje (%)
Amerindio	A	113	45.0
	B	47	18.7
	C	46	18.3
	D	19	7.6
Europeo	H	5	2.0
	I	1	0.4
	J	1	0.4
	K	2	0.8
	T	7	2.8
	U	4	1.6
	W	1	0.4
Africano	L0	1	0.4
	L2	1	0.4
	L3	3	1.2
	TOTAL	251	100

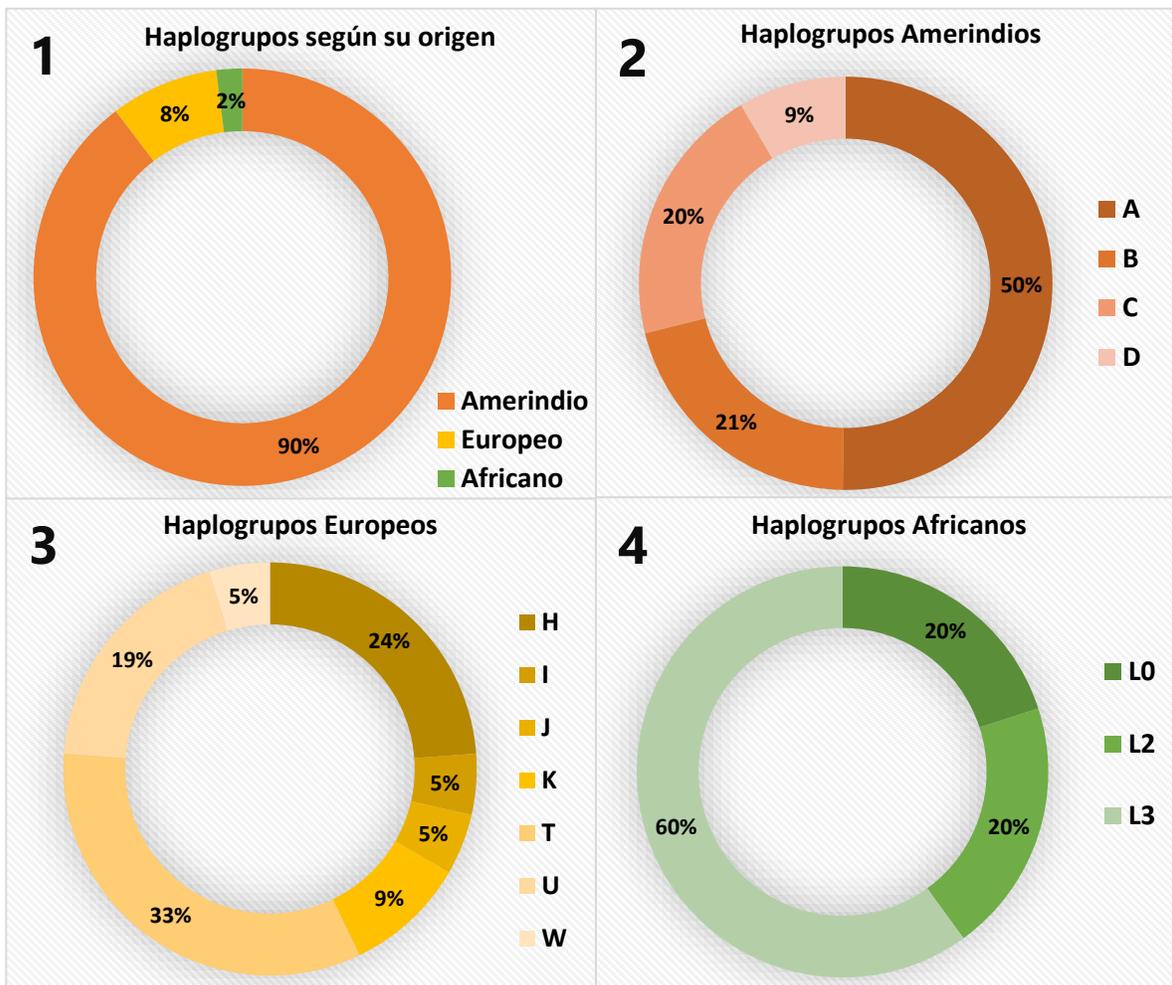


Fig. 10. Presencia de los Haplogrupos de mtDNA en población de la CDMX. Se muestra la presencia (en %) de los haplogrupos de mtDNA de la muestra poblacional de la CDMX de acuerdo a su origen (1), y la presencia de cada haplogrupo dividido por su origen: amerindio (2), europeo (3) y africano (4).

Diversidad genética

En las 251 secuencias, se encontraron un total de 201 mutaciones en 192 sitios polimórficos, de los cuales 110 son informativos, mientras que 82 no lo son (Tabla 4). Se encontraron de 3 a 22 diferencias nucleotídicas en las secuencias, teniendo como promedio 11.2077.

Tabla 4. Sitios polimórficos. Se muestra el número de sitios polimórficos encontrados en la muestra de 251 secuencias, clasificados según sus características de número de variantes y si son informativos o no.

Sitios polimórficos	Dos variantes	Tres variantes
Informativos	102	8
No informativos	81	1

Se obtuvieron un total de 208 **haplotipos (h)** distintos, de los cuales 182 fueron únicos (presentes en un solo individuo), mientras que 26 tuvieron una frecuencia mayor a 1 (presentes en dos o más personas) (Tabla 5). La **diversidad haplotípica (Hd)** fue de 0.9976. En cuanto a la **diversidad nucleotídica (π)** se obtuvo un valor de 0.01234 y un valor θ_w de 0.03466. El valor obtenido de la prueba D de Tajima fue de -2.05329*. **Significancia estadística: $P < 0.01$.*

Tabla 5. Haplotipos según su frecuencia para la población de la CDMX. Se muestra el número de haplotipos que presenta cada haplogrupo de acuerdo a la frecuencia que presentan en la muestra poblacional de la CDMX, es decir, el número de haplotipos que son idénticos en 2, 3, 4, 5 o 6 personas diferentes.

Frecuencia del haplotipo	HAPLOGRUPO														
	A	B	C	D	H	I	J	K	L0	L2	L3	T	U	W	TOTAL
1 persona	84	35	30	11	3	1	1	2	1	1	3	5	4	1	182
2 personas	10	3	2	1	1	-	-	-	-	-	-	1	-	-	18
3 personas	-	2	2	-	-	-	-	-	-	-	-	-	-	-	4
4 personas	1	-	-	-	-	-	-	-	-	-	-	-	-	-	1
5 personas	1	-	-	-	-	-	-	-	-	-	-	-	-	-	1
6 personas	-	-	1	1	-	-	-	-	-	-	-	-	-	-	2
TOTAL	96	40	35	13	4	1	1	2	1	1	3	6	4	1	208

Posterior al análisis poblacional de las 251 muestras de la Ciudad de México, se calcularon los valores de variabilidad genética por haplogrupos A, B, C y D, así como la variabilidad entre los haplotipos de origen europeo y africano (Tabla 6). También se obtuvo el número de sitios polimórficos por haplogrupo, así como el número de variantes que tiene cada uno de ellos (Tabla 7).

Tabla 6. Valores de diversidad genética. Para cada haplogrupo (HG) se muestran los parámetros de diversidad genética: número de mutaciones encontradas (μ), número de sitios polimórficos informativos (Inf) y no informativos (No Inf), promedio de diferencias nucleotídicas de las secuencias, número de haplotipos (h), diversidad haplotípica (Hd), diversidad nucleotídica (π), total de sitios segregantes (θ_w) y finalmente el estadístico D de Tajima. **Significancia estadística: $P < 0.01$.*

HG	μ	Sitios polimórficos		Diferencias nucleotídicas	h	Hd	π	θ_w	Tajima*
		Inf.	No Inf.						
A	119	56	60	5.54314	96	0.9959	0.00608	0.02402	-2.46486
B	57	26	29	5.07493	40	0.9917	0.00556	0.01364	-2.12683
C	48	22	26	4.79807	35	0.9778	0.00524	0.01192	-1.95302
D	23	8	15	3.57895	13	0.9064	0.00389	0.00716	-1.77544
Europeo	60	24	36	10.03810	19	0.9905	0.01092	0.01815	-1.59682
Africano	36	8	28	16	5	1	0.01741	0.01880	-0.55597

Tabla 7. Sitios polimórficos de acuerdo al haplogrupo. Se muestra el número de sitios polimórficos encontrados en la muestra de cada haplogrupo nativo americano, europeo y africano, clasificados según sus características de número de variantes y si son informativos o no.

Sitios polimórficos		HAPLOGRUPO				Haplogrupos europeos	Haplogrupos africanos
		A	B	C	D		
Informativos	Dos variantes	54	24	22	8	24	8
	Tres variantes	2	2	0	0	0	0
No informativos	Dos variantes	59	29	26	15	36	28
	Tres variantes	1	0	0	0	0	0

Además de analizar los sitios polimórficos, se analizaron las regiones de poli-citocinas, ubicadas en las regiones 303 a 309 y 311 a 315 de la región control del mtDNA, en las cuales es común encontrar inserciones debido a su naturaleza repetitiva; dichos resultados se encuentran en la Tabla 8.

Tabla 8. Inserciones en las regiones de poli-citocinas de la región control del mtDNA. Se muestra el número de secuencias, y el porcentaje de las mismas en paréntesis, que presentan o no inserciones en las regiones de poli-citocinas de la región control del mtDNA.

Región	Número de inserciones			
	0	1	2	3
303-309	69 (27.5%)	125 (49.8%)	49 (19.5%)	8 (3.2%)
311-315	4 (1.6%)	247 (98.4%)	0	0

Determinación de flujo génico dentro de la población

Como complemento a los análisis anteriormente realizados, se analizaron las muestras por alcaldía, con la finalidad de determinar si la muestra poblacional del presente estudio se comporta como una sola población o se encuentra estructurada por alcaldías. Para ello se calcularon los valores p de los F_{st} por alcaldía (ANEXO 4 Tabla I) para determinar si existe flujo génico entre los pares de las alcaldías y así determinar si hay estructura o no entre ellas. La F_{st} general para la población estudiada fue de -0.01239 (considerada cero), a partir de la cual se calculó N_m obteniendo un valor de -40.87 (considerado cero).

Se realizó una matriz de las F_{st} pareadas (distancia genética) y una matriz de las distancias geográficas entre las dieciséis alcaldías de la CDMX (ANEXO 4 Tabla II), las cuales se utilizaron para llevar a cabo una prueba de Mantel (Fig. 11); en donde se observan alcaldías cercanas entre sí con F_{st} bajas y viceversa, alcaldías cercanas entre sí con F_{st} altas. En dicha prueba se obtuvo un valor de correlación $r= 0.191$ y un valor de $p= 0.149$ (a partir de 10,000 permutaciones).

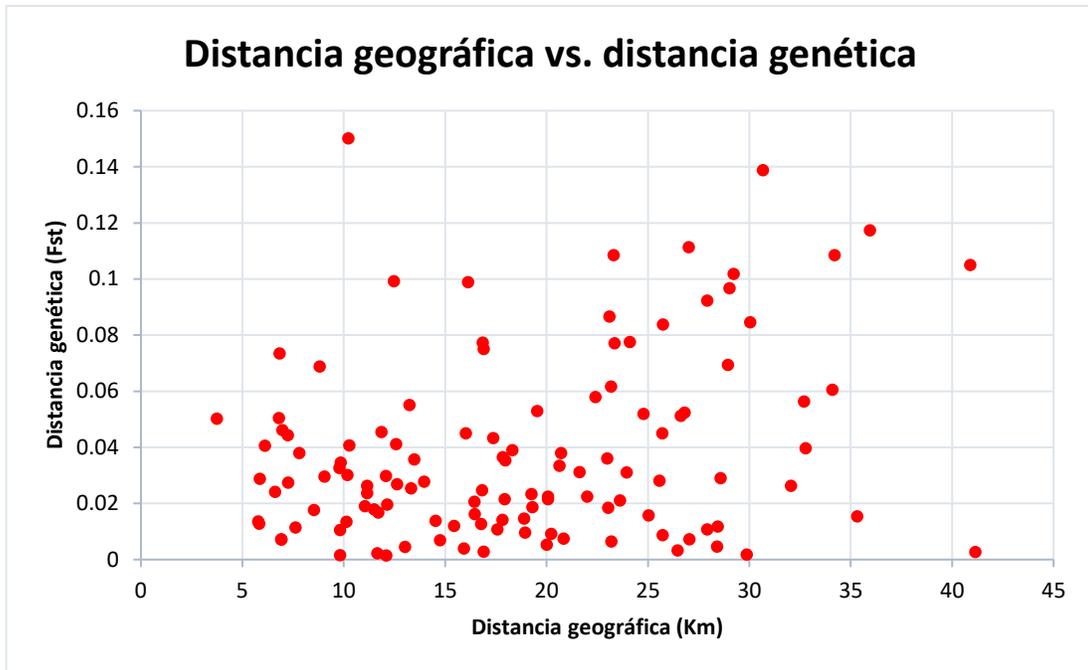


Fig. 11. Distacia geográfica vs. Distancia genética. Gráfica obtenida de la prueba de Mantel con la información de la distancia geográfica (expresada en kilómetros) y la distancia genética (F_{st} pareadas) por las alcaldías de la CDMX de la muestra poblacional.

Relaciones filogenéticas de mtDNA

Con la finalidad de vislumbrar las relaciones filogenéticas que guarda la población de la CDMX, basadas en los polimorfismos de la región control del mtDNA, se generó la red de los 208 haplotipos obtenidos a partir de la muestra de 251 individuos, la cual se muestra en la Fig. 12. Para un mayor detalle, se generó la red por haplogrupos (A, B, C y D), así como otra red por los haplotipos de origen europeo y africano, dichas redes se encuentran en el Anexo 5.

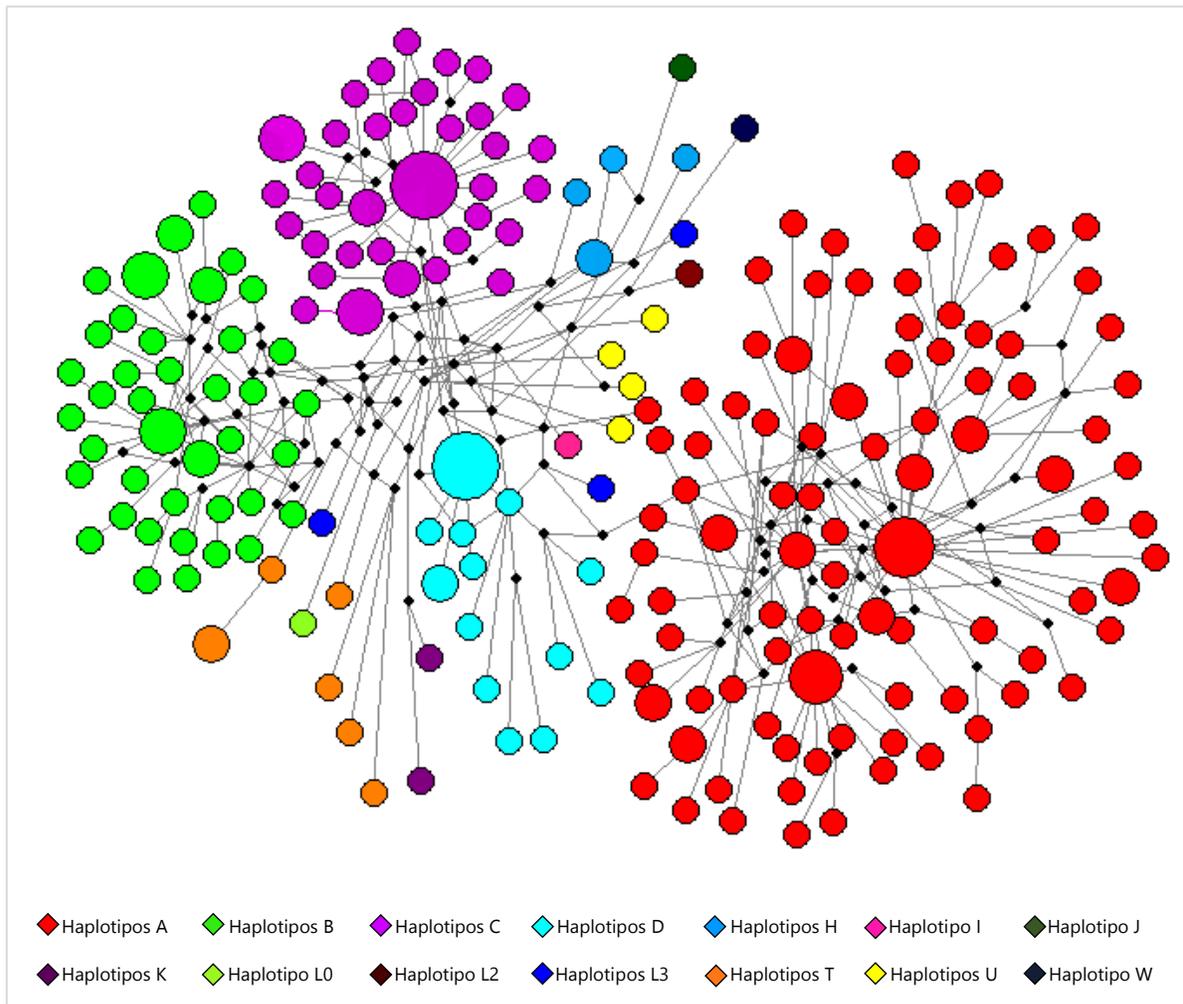


Fig. 12. Red de Haplotipos de mtDNA para la población de la CDMX. Se muestra la red de haplotipos para los 208 haplotipos de mtDNA obtenidos de la muestra de 251 individuos de la CDMX. Cada haplotipo se encuentra representado con un nodo (círculo); la frecuencia de cada haplotipo se refleja en el tamaño del nodo correspondiente. Los puntos negros representan los vectores medios de la red.

Inferencias sobre de cambios históricos en el tamaño poblacional

El análisis de *mismatch distribution* para las 251 secuencias de la muestra realizado en Arlequin generó los siguientes valores para los estimadores $\tau=14.613$, $\theta_0 = 1.493$ y $\theta_1 = 18857.80$ y un valor de $p= 0.57456$. Dichos valores fueron utilizados para generar una gráfica de *mismatch distribution* en el programa DnaSP con un modelo de crecimiento-disminución de la población para la estimación de los valores esperados (Fig. 13).

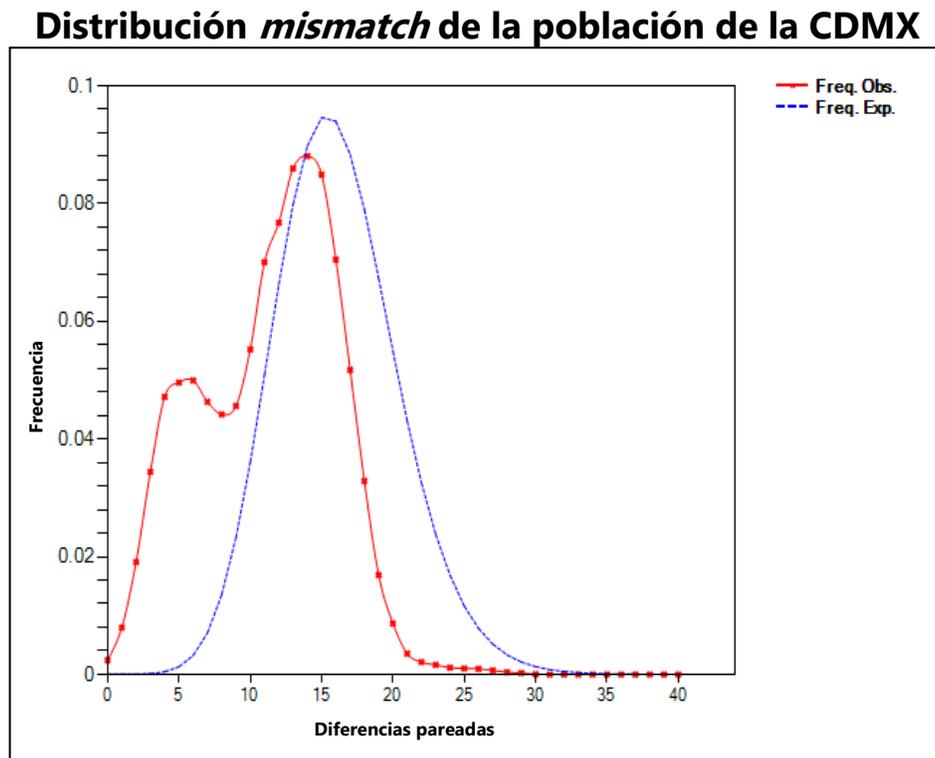


Fig. 13. Distribución *mismatch* de la población de la CDMX. Se muestra la gráfica de las diferencias encontradas entre cada par de secuencias de la región control del mtDNA vs. la frecuencia de cada una de las diferencias pareadas para la muestra poblacional de la CDMX.

Determinación de riqueza haplotípica de la población

Se realizó una curva de acumulación de haplotipos con el tamaño de muestra del presente trabajo ($n=251$) (Fig. 14), así como una extrapolación a un tamaño de muestra de 6000 individuos (Fig. 15), para determinar cuál es el número mínimo de muestras con el cual se alcanzaría a muestrear la mayoría de haplotipos de mtDNA presentes en la población. Se observa que la asíntota de la curva comienza en un número muestral superior a los 4000 individuos.

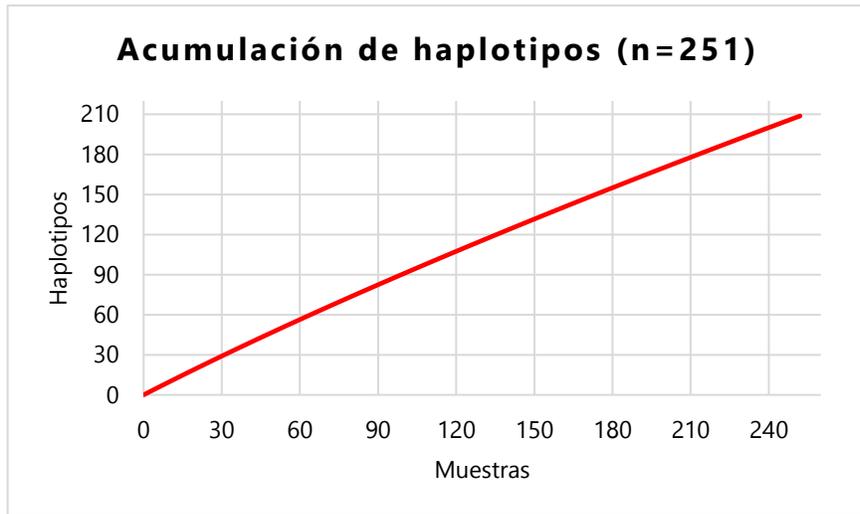


Fig. 14. Acumulación de haplotipos (n=251). Se muestra la curva de acumulación de haplotipos (número de muestras vs. número de haplotipos) para el tamaño de muestra del presente estudio.

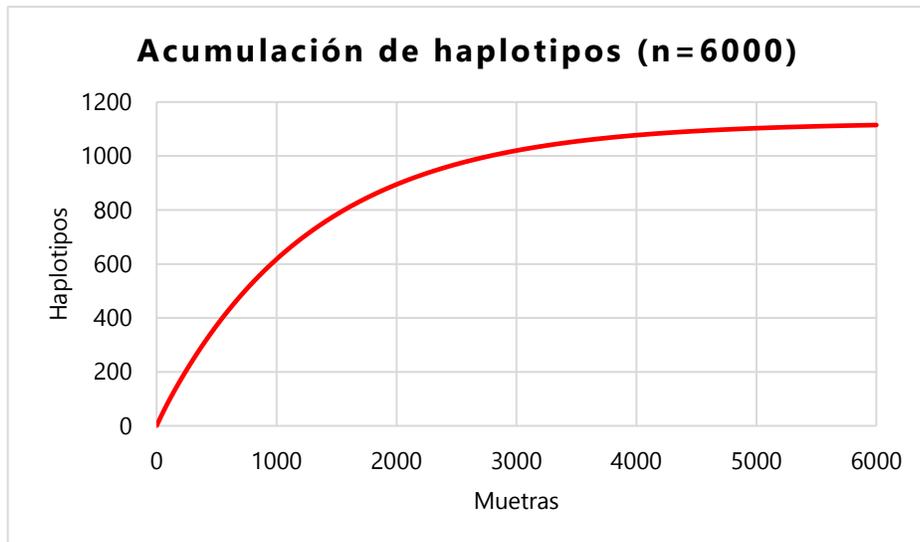


Fig. 15. Acumulación de haplotipos (n=6000). Se muestra la extrapolación de la curva de acumulación de haplotipos para el tamaño de muestra de 6000 individuos muestreados.

DISCUSIÓN

Lugar de origen de los ancestros maternos

El 68.2% de los donantes de muestra proporcionó información de su abuela materna, seguido del 16.3% de quienes sólo dieron información de sus madres, el 14.2% pertenece a información de sus bisabuelas y únicamente el 1.3% reportó información de sus tatarabuelas (ANEXO 2).

Tomando en consideración la información de los ancestros matrilineales aportada por los donantes de muestra, el 70% de ellos nació fuera de la CDMX y sólo el 30% son originarios de la misma. Aquellos ancestros que tienen un lugar de origen distinto al de la CDMX, pertenecen a 24 de los 31 estados de la república mexicana; así como también se encontraron ancestros pertenecientes a otros países (Alemania, España y Estados Unidos de América). Los estados que presentan un mayor número de ancestros son principalmente estados del centro del país como el Estado de México, Oaxaca, Michoacán, Puebla e Hidalgo, con 26, 19, 17, 15 y 14 individuos respectivamente; lo anterior es consistente con los datos de migraciones hacia la CDMX por parte de otros estados del país reportadas por el INEGI en 2010.

De las 70 personas (30%) que tienen ancestros con lugar de origen en la CDMX, sólo 3 de ellos han sido habitantes de dicha ciudad por al menos cuatro generaciones, 45 de ellos han sido habitantes por al menos tres generaciones y 22 de ellos únicamente por al menos dos generaciones. Lo anterior, comprueba que la Ciudad de México ha sido y sigue siendo blanco de múltiples migraciones de la mayoría de los estados de la república desde al menos cuatro generaciones atrás respecto a los individuos muestreados en el presente estudio.

Frecuencia de haplotipos y haplogrupos en población de la CDMX

En el presente trabajo, se encontró una frecuencia del 90% para haplogrupos nativo americanos, cifra similar a la obtenida en otros trabajos de población mestiza mexicana (Tabla 1) (Green *et al.*, 2000; Guardado *et al.*, 2009 y Martínez *et al.*, 2013). Por su parte, cuando se trata de poblaciones indígenas que por usos y costumbres no permiten procesos migratorios que favorezcan el mestizaje, la frecuencia de los haplogrupos amerindios representa alrededor del 99% (González-Sobrino *et al.*, 2016; Peñaloza *et al.*, 2007 y Gorostiza *et al.*, 2012).

En cuanto a los haplogrupos europeos, en el presente trabajo se obtuvo una presencia del 8%, que comparado con los trabajos de Green *et al.* (2000); Guardado *et al.* (2009) y Martínez *et al.* (2013) representa un aumento de alrededor del 3% y comparado con el trabajo de González-Sobrino *et al.* (2016), representa un aumento de poco más del 6% en la frecuencia de estos haplogrupos. Finalmente, en la frecuencia de los haplogrupos africanos se obtuvo un 2%, que resulta menor al reportado por Green *et al.*, (2000), pero ligeramente mayor al reportado por Guardado *et al.*, (2009), Martínez *et al.*, (2014) y González-Sobrino *et al.*, (2016).

Aunque de manera general las frecuencias de haplogrupos corresponden a la ancestría esperada para población mexicana, de acuerdo a otros trabajos reportados; se observa un claro aumento en las frecuencias de los haplotipos europeos, situación que refleja los fenómenos migratorios a los que se encuentra sometida la CDMX por tratarse de una ciudad mestiza cosmopolita, así como al muestreo relativamente homogéneo que tiene el presente trabajo para la población de estudio.

Las frecuencias de haplogrupos nativo americanos obtenidas en el presente trabajo, muestran lo siguiente:

1) El haplogrupo A (45%) es menor en la población de la CDMX, comparado con los trabajos anteriormente mencionados (Tabla 1) que muestrearon estados de la zona sur, oeste y centro del país; únicamente en poblaciones del norte y en poblaciones indígenas del Estado de México (población Mazahua), se encuentra una menor frecuencia a la obtenida en el presente trabajo (Green *et al.*, 2000; González-Oliver *et al.*, 2017).

2) Haplogrupos B (18.7%), su frecuencia resultó ser menor a la reportada en siete de los nueve trabajos anteriormente mencionados (Tabla 1), siendo mayor únicamente en los trabajos de Sandoval *et al.*, (2009) y Guardado *et al.*, (2009).

3) Haplogrupo C (18.3%), las frecuencias reportadas para este haplogrupos en población mexicana son las más variables (11.7-28.5%) y no se aprecia una relación entre poblaciones mestizas, indígenas o de acuerdo a alguna región del país, por lo que la frecuencia reportada en este trabajo se encuentra en un valor intermedio quizá esperado para población de la CDMX.

4) Haplogrupo D (7.6%), su frecuencia fue mayor respecto a los trabajos anteriormente mencionados (Tabla 1), excepto para población mazahua del Estado de México, la cual tiene una frecuencia de 20.1% (González-Oliver *et al.*, 2017), así como para algunas poblaciones indígenas del país (Gorostiza *et al.*, 2012).

Una vez realizada la caracterización del haplogrupo, se correlacionó con la información proporcionada por el donante del lugar de origen de su ancestro materno. Se encontró que el 99% de los individuos con haplogrupo A tienen un ancestro con lugar de origen en algún estado de la República Mexicana, mientras que el 1% restante su lugar de origen es Estados Unidos de América. Por su parte, todos los ancestros de los individuos con haplogrupo B y C tienen su lugar de origen en algún estado de la República Mexicana. En cuanto a los ancestros de los individuos con haplogrupo D, el 89% de ellos nacieron en algún estado de la República Mexicana, mientras que el 11% restante tiene origen en España. De los ancestros de los individuos con haplogrupo europeo el 76% nació en México, el 12% en Alemania, un 6% en España y otro porcentaje igual en Estados Unidos de América.

La caracterización de los haplogrupos nativo americanos, se llevó a cabo mediante la presencia de "mutaciones tipo" (MITOMAP) (Lott, 2019), las cuales en algunos casos son específicas para

cada haplogrupo. Tal es el caso de la mutación 16111T del haplogrupo A, la mutación 16189C del haplogrupo B o la 16298C del haplogrupo C. Sin embargo, a pesar de ser específicas para un haplogrupo, se observa que dichas mutaciones comienzan a aparecer en otros haplogrupos como el B (en el caso de las mutaciones 16111T y 16298C), o en múltiples haplotipos como A, C, T y L (en el caso de la mutación 16189C) y aunque si bien es cierto, dicha aparición sucede en una frecuencia muy baja. Lo anterior es indicativo de que al paso de las generaciones los haplotipos han ido acumulando cada vez más mutaciones (homoplasias), dando pie a la diversificación de los haplogrupos y a tener una mayor variabilidad entre los haplotipos. Por el contrario, se encontraron tres mutaciones específicas para un único haplogrupo (16327T, 249d, 290d y 291d), todas ellas pertenecientes al haplogrupo C. Este tipo de mutaciones “específicas” para un haplogrupo que se presentan en más de un haplogrupo se han reportado en trabajos como el de Guardado *et al.*, (2009) y Malhi *et al.*, (2002).

Dos mutaciones (73G y 263G) fueron consideradas como “ancestrales” (Perego *et al.*, 2010), debido a que se encuentran presentes en el 96% y 99% de las muestras, sin importar el haplogrupo o la ancestría del mismo, por lo que ninguna de ellas fue utilizada para la caracterización de los haplogrupos.

De los 208 haplotipos obtenidos, el 87.5% de ellos (182 haplotipos; Tabla 5) son únicos (sólo se muestrearon una vez) en la población de la CDMX estudiada, el 8.6% (18 haplotipos) se repite una vez, el 1.9% (4 haplotipos) lo presentan tres individuos, mientras que el 0.96% (2 haplotipos) se repite en 6 personas y finalmente el 0.48% (un haplotipo) se encuentra en 4 y 5 personas de la población. El haplogrupo A, al ser el más frecuente en la población, presenta un mayor número de haplotipos que se repiten en la población. Por su parte, cada uno de los haplogrupos C y D, presentan los haplotipos más frecuentes de la población con 6 individuos cada uno de ellos, mostrando que las mutaciones de dichos haplotipos se encuentran conservadas y no existe tanta variabilidad como en el haplogrupo A. Los haplotipos más frecuentes pueden tratarse de haplotipos que presentan las “mutaciones tipo” del haplogrupo y no acumulan otras mutaciones que los hayan diferenciado. Este tipo de haplotipos son los que por lo regular se encuentran en los nodos centrales de las redes de haplotipos, porque es a partir de ellos que surgen haplotipos nuevos.

En comparación con el trabajo de Guardado *et al.*, (2009), un trabajo aplicable para población de la CDMX debido a que su *n* contempla población de dicha ciudad, se puede observar que, en un tiempo de 10 años, donde el crecimiento de la población de la CDMX fue de 8.8 a poco más de 8.9 millones de habitantes (INEGI, 2015), las frecuencias de haplotipos europeos y africanos han aumentado, contrario a lo sucedido con las frecuencias nativo americanas. Estos cambios en las frecuencias haplotípicas de mtDNA pueden ser debidas a las migraciones a las que se encuentra sometida la CDMX, pero primordialmente son debidas al tipo de muestreo que se realizó en ambos trabajos, siendo el del presente trabajo el más homogéneo y representativo para dicha ciudad considerando la densidad poblacional de la misma.

Diversidad genética

La variabilidad genética observada es mayor a la reportada en cualquier otro trabajo previo de población mexicana (datos no mostrados). En 251 secuencias analizadas se encontraron 201 mutaciones en 192 sitios polimórficos. Del total de los sitios polimórficos más de la mitad (57.3%) son informativos; es decir, que el sitio presenta una misma mutación en al menos dos secuencias. Lo anterior es indicativo de que esa mutación probablemente se ha fijado en la población, al heredarse de un haplotipo a otro, o bien que sean sitios informativos por su contenido de mutaciones con características de homoplasias. El porcentaje restante (42.7%) se tratan de mutaciones no informativas, que más allá de resultar no informativas, pueden tratarse de polimorfismos únicos en la muestra, resultado de mutaciones recientes.

Aunque las inserciones que se encuentran en las regiones de poli-citocinas de la región control de mtDNA no son tomadas en cuenta para la caracterización de los haplotipos, es importante destacar que dichas inserciones se tratan de un polimorfismo muy frecuente en el mtDNA que vale la pena analizar. En la primera región de poli-citocinas (303-309) se observa que es posible tener no sólo una inserción, sino hasta tres de ellas. Cabe mencionar que, conforme el número de inserciones aumente, se volverá cada vez menos recurrente. Sin embargo, lo más frecuente es una inserción, ya que prácticamente la mitad de las secuencias analizadas lo presenta (48.9%). En cuanto a la segunda región (311-315), únicamente se observa una inserción en prácticamente todas las secuencias analizadas (98.4%); esto indica que dicha inserción está cercana a fijarse en la población.

La diversidad haplotípica obtenida es del 99.76%, con un total de 208 haplotipos, ambos valores son mayores a los presentados por Guardado *et al.*, (2009), trabajo que incluso tiene una muestra poblacional total mayor a la aquí analizada. Lo anterior sugiere que, en un lapso relativamente corto de 10 años, los haplotipos en la CDMX muestran una mayor cantidad de mutaciones, favoreciendo a la diferenciación y diversificación de los haplotipos, situación que se traduce en un aumento de la diversidad haplotípica.

Analizando la variabilidad genética por haplogrupos, es posible ver que el haplogrupo A es el que presenta una mayor variabilidad con 96 haplotipos distintos y una diversidad haplotípica del 99.59%. El segundo lugar de variabilidad genética se encuentra en el haplogrupo B, el cual cuenta 40 haplotipos con una diversidad haplotípica mayor al 99% (99.59%). En tercer lugar, se encuentra el haplogrupo C, que cuenta con 35 haplotipos con una variabilidad haplotípica del 97.78%. Finalmente, el haplogrupo D es en donde se presenta una menor diversidad genética en 13 haplotipos teniendo una variabilidad haplotípica del 90.64%. Para los dos primeros haplogrupos (A y B), los valores obtenidos de diversidad genética son mayores a los reportados por Guardado *et al.*, (2009); contrario a lo que sucede en los haplogrupos C y D, donde se observa una notable disminución en dichos valores.

Sobre la diversidad de los haplogrupos europeos, se analizaron siete haplogrupos con 60 sitios polimórficos con una mutación cada uno, y un total de 19 haplotipos con una variabilidad

nucleotídica del 99.05%. Para los haplogrupos africanos, se analizaron tres haplogrupos y se encontraron 36 mutaciones en el mismo número de sitios polimórficos, en 5 haplotipos con una diversidad haplotípica del 100%. A pesar de que el número de muestras obtenidas para haplotipos no nativo americanos es reducido (26 secuencias), se tiene una gran diversidad para ambos grupos. Sin embargo, en los haplogrupos europeos dicha variabilidad disminuyó respecto a lo reportado por Guardado *et al.* (2009), debido a que en el presente trabajo se encontró un haplotipo europeo en dos individuos distintos, situación que no sucedió en el trabajo anteriormente mencionado. Contrario a lo que sucedió con los haplotipos africanos en donde se observó una mayor diversidad genética.

En cuanto a los estimadores de variación genética para la población de la CDMX, se obtuvieron los valores $\pi = 0.01234$ y $\theta_w = 0.03466$; y debido a que ambos estimadores guardan la relación $\theta_w > \pi$, el valor D del test de Tajima es negativo (-2.05329). Dicha relación indica que existe un exceso de polimorfismos de baja frecuencia o mutaciones raras, lo que describe a una población que se encuentra en expansión y/o la presencia de selección direccional.

El exceso de polimorfismos de baja frecuencia puede explicarse por la alta tasa de mutación que presenta el mtDNA y a que dichas mutaciones han ocurrido recientemente, razón por la cual se presentan en muy pocos haplotipos o incluso lo hacen de manera particular para alguno de ellos; así como también porque se traten de mutaciones introducidas a la población debidas a efectos migratorios. Por su parte, el valor negativo del test D de Tajima indica que los haplotipos ancestrales (haplotipos con polimorfismos de alta frecuencia o mutaciones comunes), al paso del tiempo acumulan polimorfismos de baja frecuencia, haciendo cada vez más diferentes y diversos entre sí a los haplotipos; situación que genera un aumento en la variabilidad tanto nucleotídica como haplotípica en la población. Los resultados del *test* de Tajima, de acuerdo a la bibliografía (Hamilton, 2009), suelen ser valores negativos cuando se analizan polimorfismos de poblaciones humanas debido a un nivel de estructura poblacional muy bajo, así como un crecimiento poblacional rápido en el pasado reciente.

Adicional al análisis anteriormente descrito de los estimadores de variación genética, así como el *test* de Tajima para la muestra poblacional de la CDMX, dichos valores se obtuvieron para cada uno de los haplogrupos nativo americanos, así como para los haplogrupos de población europea y africana. Los resultados muestran valores similares, a los obtenidos en el análisis de la muestra poblacional completa, por lo que el comportamiento de esos grupos de haplogrupos es igual al de la muestra poblacional completa de la CDMX, previamente ya descrito.

Estimación del flujo génico en la población

Para determinar si dentro de la muestra poblacional analizada en el presente estudio, existía estructura o diferenciación entre las alcaldías, se llevó a cabo el análisis de estimadores de flujo génico como el cálculo de los valores p de las F_{st} entre las delegaciones, los cuales muestran

que en 111 pares de alcaldías no existen diferencias significativas entre sus variaciones genéticas, indicando que sí existe flujo génico entre ellas, lo que las lleva a homogenizarse y comportarse como una misma población. Sólo en nueve pares de alcaldías se observaron diferencias significativas; en cinco de esas nueve comparaciones se encuentra involucrada la alcaldía Gustavo A. Madero y en las cuatro restantes lo hace la alcaldía Tlalpan. Las diferencias que tienen estas alcaldías es con Cuauhtémoc, Tláhuac, Iztapalapa, Magdalena Contreras y Xochimilco. La diferenciación entre estas alcaldías puede deberse a que cada una de esas alcaldías se inició a partir de poblaciones originales distintas; y aún no se ha llegado a un equilibrio entre la deriva y flujo génico, que las homogenice genéticamente, por lo que aún tienen una estructuración genética distinta entre ellas.

A pesar de que las diferencias significativas entre los pares de alcaldías fueron pocas, el cálculo de la F_{st} general resulta ser negativo (considerado cero), que más que considerarse sinónimo de estructuración genética puede ser debido a la gran variabilidad encontrada en la región control del mtDNA. Debido a esa gran variación es difícil que la población presente las mismas mutaciones y sea homogénea. El valor de Nm obtenido fue negativo (considerado cero), lo cual indica que no existe equilibrio entre el flujo génico y la deriva genética en la población analizada, en donde la deriva se encuentra actuando de forma independiente en cada población favoreciendo la diferenciación genética entre las alcaldías. Lo anterior indica que el tamaño efectivo de migración necesario para contrarrestar los efectos de la deriva génica en la región control del mtDNA para la población de la CDMX es prácticamente infinito o no definido.

Por su parte, la prueba de Mantel para estimar el flujo genético mediante el aislamiento por distancia, con una $r=0.191$ y $p=0.149$ indica que no existe una correlación entre la distancia geográfica y la distancia genética (F_{st} pareadas) de las dieciséis alcaldías de la CDMX; es decir, la diferenciación genética entre los pares de alcaldías no se debe a la distancia geográfica entre ellas.

Los resultados anteriormente mencionados, muestran que, en la muestra poblacional para la CDMX del presente estudio, de manera general existe flujo genético, aunque este no evita que dicha población sea heterogénea. El aislamiento por distancia entre las alcaldías no es significativo considerando la capacidad de movilidad que tienen los humanos para desplazarse de un lugar a otro, además de que se tratan de distancias relativamente cortas, no mayores a los 50 kilómetros. Si bien es cierto que en las alcaldías Gustavo A. Madero y Tlalpan se presenta una diferenciación genética con hasta cinco alcaldías más, no se puede considerar que su estructura genética es completamente diferente a la del resto de la población de la CDMX, porque no se diferencian con todas las alcaldías de dicha ciudad y, con aquellas en las que se diferencian, resultan no tener diferencias significativas con el resto de la población. Quizá el poco flujo génico observado se deba a situaciones socio económicas o socio políticas que limiten en cierta medida el entrecruzamiento y por tanto diferenciación genética entre los

habitantes de dichas alcaldías. Sin embargo, para probar dicha hipótesis se deben llevar a cabo más análisis que no son objetivo del presente trabajo.

Relaciones filogenéticas del mtDNA

Debido a que cualquier región no codificante del mtDNA tiene las mismas posibilidades de mutación, es común encontrar homoplasias dentro de los polimorfismos del mtDNA. Es por ello que la construcción de árboles filogenéticos no resultan ser el mejor método de análisis, ya que puede agrupar o desagrupar erróneamente distintas secuencias. El método de análisis ideal para este tipo de secuencias resultan ser las redes de haplotipos (Bandelt *et al.*, 1995), las cuales sí permiten darle una interpretación a las homoplasias, permiten visualizar la frecuencia de los haplotipos y ver la variabilidad y diversificación de los mismos.

En la red de haplotipos generada para la población de la CDMX se pueden diferenciar con claridad cuatro agrupaciones de haplotipos, los cuales pertenecen a los haplogrupos de origen nativo americano: A, B, C y D. También se pueden apreciar que se agrupan algunos haplogrupos europeos como el H, K, T y U, a pesar de tener muy pocos haplotipos cada uno. Para poder analizar con mayor detalle la red, se realizó una red para cada haplogrupo de origen nativo americano y otra para los haplogrupos de origen europeo y africano; en ellas se puede apreciar lo siguiente:

Claramente el haplogrupo A es el más diverso genéticamente; su red de haplotipos presenta un patrón radial, es decir, a partir de un haplotipo en común surgen otros que se vuelven cada vez más distintos y diversos entre sí. También se puede apreciar que entre los haplotipos y los vectores medios se forman polígonos (cuadros), los cuales representan las homoplasias; dentro de este haplogrupo se presentan 14 polimorfismos (en las posiciones: 64, 136, 152, 153, 16111, 16172, 16209, 16223, 16239, 16300, 16311, 16372, 16391 y 16519) productos de homoplasias, algunas de ellas presentes en dos o más haplotipos.

En la red del haplogrupo B, de igual manera se observa un patrón radial y nuevamente se observan patrones en la red que indican homoplasias, esta vez 13 polimorfismos (en las posiciones: 114, 146, 152, 16104, 16092, 16140, 16182, 16183, 16295, 16311, 16357, 16362 y 16467). Por otra parte, en la red del haplogrupo C, se pueden observar haplotipos con mayor frecuencia (nodos más grandes), sinónimo de que existe una menor variabilidad dentro de los haplotipos, en este haplogrupo, únicamente se encontraron 7 polimorfismos producto de homoplasias (en las posiciones: 57, 152, 195, 16181, 16298, 16311 y 16519). En cuanto a la red del haplogrupo D se puede apreciar un patrón radial más claro, un haplotipo con una frecuencia visiblemente mayor a la del resto de los haplotipos B y lo más notorio de dicha red es que no presenta patrones de homoplasia.

Por su parte, en la red de haplogrupos europeos y africanos se aprecian haplotipos únicos, a excepción de un par pertenecientes a los haplogrupos H y T. Al igual que en las redes

anteriores, se aprecian patrones que indican homoplasias, en este caso, únicamente de 5 polimorfismos (en las posiciones: 73, 152, 16222, 16294 y 16519). El hecho de que prácticamente todos los haplotipos de esta red sean únicos, se debe a la baja frecuencia (2%) de haplotipos no nativos americanos en la población de la CDMX.

Los patrones de estrella o radiales que se aprecian en cada una de las redes de haplotipos generadas, son indicativos de una expansión poblacional (Forster, 2015), resultado acorde a los valores obtenidos en las pruebas de variabilidad genética obtenidas en el presente trabajo.

Evaluación cualitativa de cambios en el tamaño poblacional

Para corroborar el supuesto de expansión poblacional observado en los patrones de las redes de haplotipos y en el test D de Tajima, se realizó una gráfica de *mismatch distribution* en la cual se observa un patrón muy cercano a una distribución unimodal (Fig. 12), la cual se acerca mucho a las frecuencias esperadas obtenidas bajo el supuesto de una población en crecimiento-disminución. El valor de p (0.57456) indica que no hay diferencias significativas entre las frecuencias esperadas de las diferencias pareadas y las frecuencias observadas para las secuencias de la región control del mtDNA de la población de la CDMX, lo cual indica que la población de estudio se encuentra en crecimiento poblacional, resultado que corrobora los patrones radiales observados en las redes de haplotipos de mtDNA y el resultado del test D de Tajima que de igual manera indican expansión poblacional.

Riqueza haplotípica de la población de estudio

En la curva de acumulación de haplotipos obtenida con las 251 muestras del presente estudio se observa un patrón casi lineal, sinónimo de que con ese tamaño de muestra aún no se ha muestreado la mayoría de los haplotipos estimados para la población; dicho de otra forma, cada nueva muestra que se incluya y se analice en el estudio, representará un nuevo haplotipo.

Con la finalidad de determinar el tamaño de muestra necesario para describir la gran mayoría de los haplotipos, se realizó la extrapolación de la curva de acumulación de haplotipos con 6000 muestras, la cual muestra una riqueza aproximada a 1100 haplotipos, esto significa que para muestrear y describir el 80% de ellos, se deben analizar alrededor de 2000 muestras, mientras que para cubrir el 90% de dicha riqueza, sería necesario el análisis de 2700 muestras.

Con lo anterior, se determina que el presente estudio describe cerca del 19% de la riqueza haplotípica estimada para la población de la CDMX. A pesar de que el presente trabajo no describe la mayor parte de la riqueza haplotípica, es probable que los haplotipos más comunes ya se hayan muestreado, quedando únicamente los haplotipos raros. Además, es importante considerar que la riqueza aquí descrita, va conforme la densidad poblacional de la CDMX, es decir, no presenta sesgos de muestreo en cuanto a las poblaciones mayormente o menormente pobladas de dicha ciudad.

CONCLUSIONES

- Se estandarizó una metodología que va desde la extracción hasta la secuenciación y análisis de mtDNA, la cual resultó funcional y aplicable para laboratorios que deseen llevar a cabo estudios con mtDNA.
- Sólo el 30% de los individuos presenta ancestría en la CDMX de al menos dos generaciones, el 70% restante de los individuos tiene ancestros con lugar de origen distinto al de la ciudad estudiada, reforzando que la ciudad de estudio se encuentra sujeta a múltiples migraciones, debidas a distintos factores socioculturales.
- La frecuencia de ancestría materna definida por haplogrupos en la población de la CDMX se compone por: 90% de ancestría nativo americano, 8% europea y 2% africana.
- De los haplogrupos obtenidos, el 45% corresponde al haplogrupo A, 18.7% al haplogrupo B, 18.3% al haplogrupo C y el 7.6% al haplogrupo D.
- Respecto a los haplogrupos con ancestría europea y africana se observó una mayor frecuencia respecto a estudios previos realizados en CDMX, aumentando su variabilidad genética.
- Se encontraron 201 mutaciones en 192 sitios polimórficos, 208 haplotipos y una diversidad genética del 99.76%.
- La alta tasa de mutación del mtDNA genera que, al paso de las generaciones y los años, los haplotipos acumulen un mayor número de mutaciones, las cuales pueden no tener el mismo origen (homoplasias) y presentarse en haplogrupos distintos con diferentes historias evolutivas.
- Se encontraron 32 homoplasias, la mayoría de ellas entre los haplogrupos A y B, los de mayor variabilidad en la población de CDMX.
- El aumento en el número de mutaciones, sitios polimórficos, número de haplotipos distintos, diversidad haplotípica y nucleotídica, así como los valores negativos para el *test D* de Tajima demuestran que la población de la CDMX se encuentra en expansión o crecimiento.
- La alta frecuencia en los polimorfismos de baja frecuencia demuestra que la mayoría de las mutaciones encontradas son recientes debido a crecimientos poblacionales; favoreciendo una diferenciación haplotípica y un aumento en la diversidad genética.

- En la muestra de la población de la CDMX, existe flujo génico entre la mayoría de las alcaldías, aunque no hay un equilibrio del mismo con la deriva genética, situación que provoca que la población es heterogénea debido a la gran variabilidad que presenta la región control del mtDNA.
- Los patrones radiales de las redes de haplotipos en conjunto con el análisis de *mismatch distribution* corroboran que la población estudiada se encuentra en crecimiento o expansión.
- Las frecuencias de los haplotipos y haplogrupos generadas en este trabajo sobre los polimorfismos de la región control del mtDNA representan un primer esfuerzo para generar una base de datos útil con fines de identificación humana para población de la CDMX.

LITERATURA CITADA

Addinsoft (2019). XLSTAT statistical and data analysis solution. New York, USA. Sitio Web: <https://www.xlstat.com>

Anderson, S., Bankier, A., Barrell, B., Bruijn, M., Coulson, A., Drouin, J., Eperon, I., Nierlich, D., Roe, B., Sanger, F., Schreier, P., Smith, A., Staden, R. y Young, I. (1981). Sequence and organization of the human mitochondrial genome. *Nature* Vol. 290: 457-465.

Andrews, R., Kubacka, I., Chinnery, P., Lightowlers, R., Turnbull, D. y Howell, N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genetics* Vol. 23: 147.

Applied Biosystems (2016). BigDye™ Terminator v3.1 Cycle Sequencing Kit. User Guide. Life Technologies Corporation, USA. Publication Number: 4337035.

Bandelt, H., Forster, P., Sykes, B. y Richards, M. (1995). Mitochondrial Portraits of Human Populations Using Median Networks. *Genetics* 141: 743-753.

Bandelt, H., Forster, P. y Röhl, A. (1999). Median-Joining Networks for Inferring Intraespecific Phylogenies. *Mol. Biol. Evol.* 16(1): 37-48.

Bandelt, H., Kloss, A., Richards, M., Yong, Y. y Logan, I. (2013). The case for the continuing use of the revised Cambridge Reference Sequence (rCRS) and the standardization of notation in human mitochondrial DNA studies. *Journal of Human Genetics*: 1-12.

Bernabéu, S. y Varela, C. Editores (2010). *La ciudad americana: mitos, espacios y control social*. España: Doce calles.

Butler, J. (2011). *Advanced Topics in Forensic DNA Typing: Methodology*. California, USA: Elsevier Inc.

Butler, J. (2005). *Forensic DNA Typing. Biology, Technology and Genetic STR Markers*. California, USA: Elsevier Academic Press.

Butler, J. (2009). *Fundamentals of Forensic DNA Typing*. California, USA: Elsevier Academic Press.

Cann, R., Stoneking, M. y Wilson, A. (1987). Mitochondrial DNA and human evolution. *Nature* 325: 31-36.

Castañeda, M. (2015). La identidad étnica y cultural de la literatura indígena mexicana en la modernidad. *Ulúa. Revista de Historia, Sociedad y Cultura* Num. 25: 143-155.

Chinnery, P. y Hudson, G. (2013). Mitochondrial genetics. *British Medical Bulletin* 106: 135-159.

- Comas, D. (2014). La genética de las migraciones humanas. Siguiendo el rastro de las migraciones a través de nuestro genoma. *Science Studies Journal* 4: 1-7.
- Depaulis, F. y Veuille, M. (1998). Neutrality Test Base on the Distribution of Haplotypes Under an Infinite-Site Model. *Molecular Biology and Evolution* 15: 1788-1790.
- Editorial (2017). Revised guidelines for the publication of genetic population data. *Forensic Science International: Genetics* 30: 160-163.
- Eguiarte, L., Souza, V. y Aguirre, X. (2007). *Ecología molecular*. México: SEMARNAT, Instituto Nacional de Ecología y Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO).
- Excoffier, L. y Smouse, P. (1994). Using Allele Frequencies and Geographic Subdivision to Reconstruct Gene Trees Within a Species: Molecular Variance Parsimony. *Genetics* 136: 343-359.
- Federal Bureau of Investigation (2008). *Mitochondrial DNA Analysis Protocol*. Revision 2.
- Fernández, E. (2000). *Polimorfismos de DNA Mitocondrial en Poblaciones Antiguas de la Cuenca Mediterránea* (tesis doctoral). Universidad de Barcelona, España.
- Fontdevila, A., Moya, A. (2007). *Introducción a la genética de poblaciones*. Madrid, España: Editorial Síntesis.
- Forster, M. (2015). *Network 5.0.0.0 User Guide*. Version date: 24 december 2015. Fluxus Technology Ltd.
- Gobierno de la Ciudad de México (2019) Portal de datos de la Ciudad de México. Recuperado de: <https://datos.cdmx.gob.mx/explore/dataset/alcaldias/table/?location=9,19.32072,-99.15261>.
- Goodwin, W., Linacre, A. y Hadi, S. (2011). *An Introduction to Forensic Genetics*. Hoboken, USA: Wiley-Blackwell.
- González-Martín, A., Gorostiza, A., Regalado, L., Arroyo, Tirado, S., Nuño, I., Rubi, R., Sandoval, K., Coble, M., Rangel, H. (2015). Demographic History of Indigenous Populations in Mesoamerica Based on mtDNA Sequence Data. *PLOS ONE* 10 (8): e0131791.
- González-Oliver, A., Garfias, E., Glenn, D. y Quinto, M. (2017). Mitochondrial DNA Analysis of Mazahua and Otomi Indigenous Populations from Estado de México Suggests a Distant Common Ancestry. *Human Biology* 89 (3): 195-216.
- González-Oreja, J., de la Fuente, A., Hernández, L., Buzo, D. y Bonache, C. (2010). Evaluación de estimadores no paramétricos de la riqueza de especies. Un ejemplo con aves en áreas verdes de la ciudad de Puebla, México. *Animal Biodiversity and Conservation* 33.1
- González-Sobrino, B., Pintado, A., Sebastián, L., Morales, F., Contreras, A., Aguilar, Y., Chávez, J., Carrillo, A., Silva, I. y Medrano, L. (2016). Genetic Diversity and Differentiation in Urban and

Indigenous Populations of México: Patterns of Mitochondrial DNA and Y-Chromosome Lineages. *Biodemography and Social Biology* 62 (1): 53-72.

Gorostiza, A., Acunha, V., Regalado, L., Tirado, S., Granados, J., Sámano, D., Rangel, H. y González, A. (2012). Reconstructing the History of Mesoamerican Populations through the Study of the Mitochondrial DNA Control Region. *PLOS ONE* Vol. 7(9): e44666.

Green, L., Derr, J. y Knight, A. (2000). mtDNA Affinities of the Peoples of North-Central México. *Am. J. Hum. Genet.* Vol. 66: 989-998.

Guardado, M., Juárez, E., Medina, I., Wegier, A., Macías, A., Gómez, G., Cruz, F., Román, E., Piñero, D., Kofman, S. y Berumen, J. (2009). A great diversity of Amerindian mitochondrial DNA ancestry is present in the Mexican mestizo population. *Journal of Human Genetics* 54: 695-705.

Hamilton, M. (2009). *Population genetics*. Hoboken, USA: Wiley-BlackWell.

Herrera, E. (2013). La Genética de Poblaciones y el Origen de la Diversidad Humana. *Rev. Med. Hondur.* Vol. 8(1): 40-45.

Ingman, M., Kaessmann, H., Pääbo, S. y Gyllensten, U. (2000). Mitochondrial genome variation and the origin of modern humans. *Nature* Vol. 408: 708-713.

Instituto Nacional de Estadística y Geografía (INEGI) (2015). Número de habitantes por Delegación. Agosto 2107, de INEGI Recuperado de: <http://cuentame.inegi.org.mx/monografias/informacion/df/poblacion/>

Instituto Nacional de Estadística y Geografía (INEGI) y Fondo de las Naciones Unidas para la Infancia (UNICEF) (2018). *Derecho a la Identidad. Cobertura del registro de nacimiento en México*. México.

Invitrogen (2017). *Invitrogen™ Platinum™ SuperFi™ Green PCR Master Mix. User Guide*. Thermo Fisher Scientific. Pub. No. MAN0014885 Rev. B.

Jiménez, A. y Hortal, J. (2003). Las curvas de acumulación de especies y la necesidad de evaluar la calidad de los inventarios biológicos. *Revista Ibérica de Aracnología* Vol. 8: 151-161.

Jobling, M. y Gill, P. (2004). Encoded Evidence: DNA in Forensic Analysis. *Nature Reviews Genetics* Vol. 5: 739-751.

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature* 217 (129): 624-626.

Llamas, B., Fehren, L., Valverde, G., Soubrier, J., Mallick, S., Rohland, N., Nordenfelt, S., Valdiosera, C., Richards, S., Rohrlach, A., Barreto, M., Flores, I., Tomasto, E., Watson, L., Makowski, K., LeBoreiro, I., Mansilla, J., Ballivián, J., Rivera, M., Burger, R., Ceruti, M., Reinhard, J., Wells, S., Politis, G., Santoro, C., Standen, V., Smith, C. Reich, D., Ho, S., Cooper, A. y Haak, Wolfgang. (2016). Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Science Advances* Vol. 2 (4): e1501385.

López, C., García, V. (2013). Aproximaciones científicas al mestizo mexicano. *Historia, Ciências, Saúde-Manguinhos*, Rio de Janeiro Vol. 20(2): 391-410.

López-Beltrán, C., y Deister, V. G. (2013). Aproximaciones científicas al mestizo mexicano. *Historia, ciencias, saude-manguinhos* 20(2): 391-410.

Lott, M. MITOMAP: A Human Mitochondrial Genome Database. 2019. Sitio Web: <http://www.mitomap.org>

Luo, S., Valencia, A., Zhang, J., Lee, N-C., Slone, J., Gui, B., Wang, X., Li, Z., Dell, S., Brown, J., Chen, S., Chien, Y-H., Hwu, W-L., Fan, P-C., Wong, L., Atwal, P. y Huang, T. (2018). Biparental Inheritance of Mitochondrial DNA in Humans. *PNAS* 115 (51): 13039-13044.

Malhi, R., Eshleman, J., Greenberg, J., Weiss, D., Schultz, B., Kaestle, F., Lorenz, J., Kemp, B., Johnson, J. y Smith, D. (2002). The Structure of Diversity within New World Mitochondrial DNA Haplogroups: Implications for the Prehistory of North America. *Am. J. Hum. Genet.* Vol. 71: 905-919.

Martínez, G., Salazar, J., Haro, J., Rubi, R., Velarde, J., Muñoz, J., López, M., Carrillo, E., Canseco, L., Bravi, C., López, M. y Rangel, H. (2013). Maternal Admixture and Population Structure in Mexican-Mestizos based on mtDNA haplogroups. *American Journal of Physical Anthropology* Vol. 151(4): 526-537.

Mishmar, D., Ruiz, E., Golik, P., Macaulay, V., Clark, A., Hosseini, S., Brandon, M., Easley, K., Chen, E., Brown, M., Sukernik, R., Olckers, A. y Wallace, D. (2003). Natural selection shaped regional mtDNA variation in humans. *PNAS* Vol. 100(1): 171-176.

Pakendorf, B. y Stoneking, M. (2005). Mitochondrial DNA and Human Evolution. *Annual Reviews Genomics Human Genetics* 6: 165-83.

Parson, W., Gusmao, L., Hares, D., Irwin, J., Mayr, W., Morling, N., Pokorak, Prinz, M., Salas, A., Schneider, P., Parsons, T. (2014). DNA Commission of the International Society for Forensic Genetics: Revised and extended guidelines for mitochondrial DNA typing. *Forensic Science International: Genetics* 13: 134-142.

Peñaloza, R., Arenas, D., Cerda, R., Buentello, L., González, G., Torres, J., Álvarez, B., Mendoza, I., Flores, M., Sandoval, L., Loeza, F., Ramos, I., Muñoz, L. y Salamanca, F. (2007). Characterization of mtDNA Haplogroups in 14 Mexican Indigenous Populations. *Human Biology* Vol. 79(3): 313-320.

Perego, U., Angerhofer, N., Pala, M., Olivieri, A., Lancioni, H., Hooshiar, B., Carossa, V., Ekins, J., Gómez, A., Huber, G., Zimmermann, B., Corach, D., Babudri, N., Panara, F., Myres, N., Parson, W., Semino, O., Salas, A., Woodward, S., Achilli, A. y Torroni, A. (2010). The initial popling of the Americas: A growing number of founding mitochondrial genomes from Beringia. *Genome Research* Vol. 20: 1174-1179.

Potter, B., Reuther, J., Holliday, V., Holmes, C., Miller, S. y Schmuck, N. (2017). Early colonization of Beringia and Northern North America: Chronology, routes, and adaptive strategies. *Quaternary International* Vol. 444, Part B: 36-55.

Potter, B., Baichtal, J., Beaudoin, A., Fehren-Schmitz, L., Haynes, C., Holliday, V., Holmes, C., Ives, J., Kelly, R., Llamas, B., Malhi, R., Miller, S., Reich, D., Reuther, J., Schiffels, S. y Surovell, T. (2018). Current evidence allows multiple models for the peopling of the Americas. *Science Advances* Vol. 4 (8): eaat5473.

Raghavan, M., Steinrücken, M., Harris, K., Schiffels, S., Rasmussen, S., DeGiorgio, M., Albrechtsen, A., Valdiosera, C., Ávila, M., Malaspina, A., Eriksson, A., Moltke, I., Metspalu, M., Homburger, J., JeffWall, Cornejo, O., Moreno, V., Korneliusson, T., Pierre, T., Rasmussen, M., Campos, P., de Barros, P., Allentoft, M., Lindo, J., Metspalu, E., Rodríguez, R., Mansilla, J., Henrickson, C., Seguin, A., Malmström, H., Stafford Jr., T., Shringarpure, S., Moreno, A., Karmin, M., Tambets, K., Bergström, A., Xue, Y., Warmuth, V., Friend, A., Singarayer, J., Valdés, P., Balloux, F., Leborerio, I., Vera, J., Rangel, H., Pettener, D., Luiselli, D., Davis, L., Heyer, E., Zollikofer, C., Ponce, M., Smith, C., Grimes, V., Pike, K., Deal, M., Fuller, B., Arriaza, B., Standen, V., Luz, M., Ricaut, F., Guidon, N., Osipova, L., Voevoda, M., Posukh, O., Balanovsky, O., Lavryashina, M., Bogunov, Y., Khusnutdinova, E., Gubina, M., Balanovska, E., Fedorova, S., Litvinov, S., Malyarchuk, B., Derenko, M., Moshier, M., Archer, D., Cybulski, J., Petzelt, B., Mitchell, J., Worl, R., Norman, P., Parham, P., Kemp, B., Kivisild, T., Tyler, C., Sandhu, M., Crawford, M., Villems, R., Smith, D., Waters, M., Goebel, T., Johnson, J., Malhi, R., Jakobsson, M., Meltzer, D., Manica, A., Durbin, R., Bustamante, C., Song, Y., Mielsen, R. y Willerslev, E. (2015). Genomic Evidence for the Pleistocene and recent population history of Native Americans. *Science* Vol. 349(6250): aab3884.

Reidla, M., Kivisild, T., Metspalu, E., Kaldma, K., Tambets, K., Tolk, H., Parik, J., Loogväli, E., Derenko, M., Malyarchuk, B., Bermisheva, M., Zhadanov, S., Pennarun, E., Gubina, M., Golubenko, M., Damba, L., Fedorova, S., Gusar, V., Grechanina, E., Mikerezi, I., Moisan, J., Chaventré, A., Khusnutdinova, E., Osipova, L., Stepanov, V., Voevoda, M., Achilli, A., Rengo, C., Janicijevic, O., Rudan, P., Anagnou, N., Michalodimitrakis, E., Koziel, S., Usanga, E., Geberhiwot, T., Herranstand, C., Howell, N., Torroni, A. y Villems, R. (2003). Origin and Diffusion of mtDNA Haplogroup X. *Journal of Human Genetics* 73: 1178-1190.

Rogers, A. y Harpending, H. (1992). Population Growth Makes Waves in the Distribution of Pairwise Genetic Differences. *Mol. Biol. Evol.* 9(3): 552-569.

Rohland, N. y Hofreiter, M. (2007). Comparison and optimization of ancient DNA extraction. *BioTechniques* 42: 343-352.

Rozas, J., Ferrer, A., Sánchez, J., Guirao, S., Librado, P., Ramos, S. y Sánchez, A. (2017). DnaSP v6: DNA Sequence Polymorphism Analysis of Large Datasets. *Mol. Biol. Evol.* 34: 3299-3302.

- Royrvik, E., Burgstaller, J. y Johnston, I. (2016). mtDNA diversity in human populations highlights the merit of haplotype matching in gene therapies. *Molecular Human Reproduction* Vol. 22(11): 809-817.
- Sandoval, K., Buentello, L., Peñaloza, R., Avelino, H., Salas, A., Calafell, F. y Comas, D. (2009). Linguistic and maternal genetic diversity are not correlated in Native Mexicans. *Human Genetics* 126(4): 521-531.
- Slatkin, M. (1994). Gene flow and population structure. In *Ecological Genetics*. New Jersey, U.S.A. Princeton University Press.
- Stoneking, M (1994). Mitochondrial DNA and Human Evolution. *Journal of Bioenergetics and Biomembranes* Vol. 26(3): 251-259.
- Sutovsky, P. (2003). Ubiquitin-Dependent Proteolysis in Mammalian Spermatogenesis, Fertilization, and Sperm Quality Control: Killing Three Birds with One Stone. *Microscopy Research and Technique* 61: 88-102.
- Templeton, A. (2006). *Population Genetics and Microevolutionary Theory*. Hoboken, USA: John Wiley & Sons, Inc.
- Thompson, T y Black, S (2007). *Human Identification. An Introduction*. United States of America: CRC Press Tylor & Francis Group.
- Tully, G., Bär, W., Brinkmann, B., Carracedo, A., Gill, P., Morling, N., Parson, W., Schenider, P. (2001). Considerations by the European DNA profiling (EDNAP) group on the working practices, nomenclature and interpretation of mitochondrial DNA profiles. *Forensic Science International* 124: 83-91.
- Tully, L., Parsons, T., Steighner, R., Holland, M., Marino, M., Prenger, V. (2000). A Sensitive Denaturing Gradient-Gel Electrophoresis Assay Reveals a High Frequency of Heteroplasmy in Hypervariable Region 1 of the Human mtDNA Control Region. *American Journal of Human Genetics* 67(2): 432-443.
- Venter, C., *et al.* (2001). The sequence of the human genome. *Science* Vol. 291: 1304-1351.
- Ventura, M. (2010) *Fronteras y Mestizajes. Sistemas de clasificación en Europa, América y África*. Universitat Autònoma de Barcelona. Ballaterra, España: Servei de Publicacions.
- von Wurmb-Schwark, N., Preusse-Prange, A., Heinrich, A., Simeoni, E., Bosch, T., & Schwark, T. (2009). A new multiplex-PCR comprising autosomal and y-specific STRs and mitochondrial DNA to analyze highly degraded material. *Forensic Science International: Genetics* 3(2), 96-103.
- Watterson, G. (1975). On the Number of Segregating Sites in Genetical Models without Recombination. *Theoretical Population Biology* 7: 256-276.

Willems, T., Gymrek, M., Highnam, G., Mittelman, D., Erlich, Y., & 1000 Genomes Project Consortium. (2014). The landscape of human STR variation. *Genome research* 24(11), 1894-1904.

Young, H., Song, I., Ha, E., Cho, S., Yang, W. y Shin, K. (2007). mtDNAManager. 2019. Yonsei University. Sitio Web: <http://mtmanager.yonsei.ac.kr/index.php>

Zhou, Q., Li, H., Li, H., Nakagawa, A., Lin, J., Lee, E., Harry, B., Skeen-Gaar, R., Suehiro, Y., William, D., Mitani, S., Yuan, H., Kang, B. y Xue, D. (2016). Mitochondrial endonuclease G mediates breakdown of paternal mitochondria upon fertilization. *Science* Vol. 353(6297): 394-399.

ANEXO 1. FORMATOS DE TOMA DE MUESTRA

Formato de consentimiento informado

CARTA DE CONSENTIMIENTO INFORMADO PARA EMPLEO DE MUESTRAS BIOLÓGICAS EN ESTUDIOS DE DNA

Ciudad de México a _____ de 201_.

Yo, _____,
manifiesto donar voluntariamente y sin mediar trato comercial alguno, muestra biológica de saliva para estudios genéticos con finalidades de investigación académica. Asimismo, declaro haber recibido información suficiente y detallada de las finalidades del Proyecto al que mis muestras están destinadas, siendo consiente que podrán ser empleadas en otros proyectos de investigación académica vinculados a estudios de Genética de Poblaciones.

El Proyecto que por principio tendrá uso mí muestra biológica intitulado “Análisis de ADN mitocondrial en población mestiza de la Ciudad de México” tiene por objeto investigar cuales son las frecuencias de haplotipos y haplogrupos presentes en el ADN mitocondrial de los pobladores de la Ciudad de México. Me queda suficientemente claro que para establecer relaciones dichas frecuencias, los investigadores deben extraer y analizar el DNA de la muestra, y que la información obtenida servirá para conocer mejor la Genética de Poblaciones de la Ciudad de México además de brindar soporte estadístico para el ámbito forense.

También se me informó que el riesgo para mi salud es nulo puesto que únicamente se someterá a análisis la saliva donada, sin requerirse intervención alguna sobre mi organismo salvo el procedimiento necesario para la obtención de la muestra, a saber: exudado bucal con material nuevo, estéril y no tóxico.

En atención a lo anterior, autorizo el uso de mi muestra biológica donada exclusivamente para el tipo de análisis genéticos aquí expuestos y con las finalidades académicas y de investigación que me fueron planteadas, quedando claro a ambas partes:

1. Que el trato comercial con la muestra, con la información obtenida de ella y/o con mis datos personales, está prohibido por Ley, y por lo tanto no puede ser objeto de comercio por ninguna de las partes, ya sea entre sí o con un tercero. Asimismo, el Proyecto de investigación se compromete a guardar la confidencialidad de todos mis datos personales y/o familiares, evitando en todo momento identificarme directamente con la información genética obtenida de la muestra, misma a la que se le dará exclusivamente el uso académico que me fue explicado y que consta en el presente documento.
2. Que en todo momento guardo plena potestad sobre la muestra biológica proporcionada, razón por la cual, estoy en completa libertad de revocar este Consentimiento cuando así lo considere necesario de acuerdo a mis intereses y/o creencias, sin que ello implique explicación alguna al Proyecto en curso sobre mi decisión.
3. Que en todo momento debo ser notificado del estado que guarda mi muestra, así como de la información genética desprendida de ella y de las conclusiones a que llegue la investigación.
4. Que cualquier otra finalidad académica (diferente de la aquí expuesta) que se le quiera dar a la información obtenida de mi muestra, primeramente, deberá serme planteada y suficientemente explicada a fin de decidir libremente si deseo o no que mi muestra biológica tome parte en ello.

Leída la presente carta de consentimiento y quedando de conformidad, firma quien saben hacerlo y quien no imprime su huella digital.

El donante (o tutor)

Nombre y firma del donante

Por el Proyecto

“Hemos discutido el contenido de esta Carta con el arriba firmante, habiéndole explicado la naturaleza del estudio y los beneficios potenciales”

Responsables del Proyecto.

Dra. América Nitxin Castañeda Sortibran
M en C. Mauro López Armenta
Biól. Carolina León Campos

Investigador del Proyecto y responsable de toma de muestra.

Nombre y firma de quien toma muestra

Laboratorio Genética. Instituto de Ciencias Forense. TSJCDMX.
Niños Héroes No. 130, Col. Doctores, C.P.06720, Ciudad de México, D.F. Tel. 5134 1300. Ext. 2339.

Cuestionario informativo para toma de muestra

Datos del donante	
*No. de muestra:	
Fecha de toma de muestra:	
Nombre:	
Edad:	
Delegación en la que vive:	
Correo electrónico:	
**Número telefónico celular:	
**Número telefónico casa:	
*Coordenadas:	
Datos del ancestro conocido más lejano vía materna	
Parentesco biológico	
Nombre completo:	
Año de nacimiento:	
Si ya no vive año de fallecimiento:	
Localidad, ciudad y país de origen:	
*Coordenadas:	

* Datos que deberá llenar el investigador ** Datos opcionales

ANEXO 2. ANCESTROS MATERNOS Y SU LUGAR DE ORIGEN

Información de los ancestros maternos. Se muestra el número de ancestros maternos (clasificados en madre, abuela, bisabuela y tatarabuela) de acuerdo a su lugar de origen (República Mexicana o país), el total por estado o país, así como el total de ancestros por cada nivel de ancestría (madre, abuela, bisabuela y tatarabuela).

	Madre	Abuela	Bisabuela	Tatarabuela	Total
Aguascalientes	0	1	1	0	2
Campeche	0	1	0	0	1
Coahuila	0	2	0	0	2
Chiapas	1	4	1	0	6
Chihuahua	0	0	1	0	1
CDMX	22	45	3	0	70
Durango	0	1	0	0	1
Guanajuato	1	5	2	0	8
Guerrero	0	3	2	0	5
Hidalgo	0	13	1	0	14
Jalisco	0	5	2	0	7
Edo. Méx.	6	18	2	0	26
Michoacán	4	11	2	0	17
Morelos	0	1	2	0	3
Nuevo León	1	2	0	0	3
Oaxaca	1	14	3	1	19
Puebla	0	14	1	0	15
Querétaro	1	4	2	0	7
San Luis Potosí	0	0	1	0	1
Sinaloa	0	0	1	0	1
Tabasco	0	1	0	0	1
Tlaxcala	0	2	0	0	2
Veracruz	0	6	3	0	9
Yucatán	0	1	0	0	1
Zacatecas	1	2	0	1	4
España	0	1	1	1	3
Alemania	0	0	2	0	2
EUA	0	2	0	0	2
Total	38	159	33	3	233

ANEXO 3. HAPLOTIPOS DE mtDNA

- El haplotipo de cada muestra se presenta en orden de acuerdo al haplogrupo al que pertenece.
- El listado de haplogrupos se presenta en orden alfabético.
- En la primera columna se encuentra el número de muestra al cual corresponde dicho haplotipo.
- La posición de las mutaciones se encuentra en color morado en la primera fila. Las posiciones en rojo representan los "Marcadores de Haplogrupo" (en el caso de haplogrupos Amerindios), mutaciones mediante las cuales se caracterizaba el haplogrupo. Las posiciones en color verde representan mutaciones ancestrales. Las posiciones en color turquesa representan sitios de inserción de citocinas en regiones de Poli-C.
- Las mutaciones en color negro representan transiciones y las de color verde transversiones. En color rojo representan las heteroplasmias.

Haplogrupo A

	16,075	16,111	16,129	16,136	16,187	16,193	16,223	16,240	16,243	16,249	16,263	16,265	16,274	16,290	16,311	16,319	16,355	16,362	16,384	16,390	16,391	16,512	16,519	16,554	64	73	114	143	146	151	152	153	195	198	214	215	235	263	273	288	308	314					
1	T					T								T	A	C						C		G		G			C			G					G	G		G		.1C					
4	T					T	T							T	A	C				A						T	G			C			G					G	G			.1C	.1C				
8							T	G		C				T	A	C										T	G			C			G	C				G	G	T		.1C	.1C				
10	T						T							T	A	C							C		T	G			A	C	T		G				G	G			.1C	.1C					
11	T	A					T							T	A	C				A		C		T	G		G											G			.3C	.1C					
13	T						T							T	A	C							C		T	G				C			G					G	G			.1C	.1C				
15	T				T		T							T	A	C									T	G				C			G					G	G			.1C	.1C				
19	T						T				C			T	A	C											G			C			G						G	G			.2C	.1C			
21	T						T							T	C	A	C		G~ A		A		C				G			C			G			G				G	G		.2C	.1C			
23	T						T		C					T	A	C							C		T	G				C			G						G	G			.1C	.1C			
24	T						T							T	A	C									T	G				C			G						G	G			.1C				
28	T						T							T	A	C									T	G				C			G						G	G			.1C				
31	T						T						A	T	A	C									T	G				C			G						G	G			.1C	.1C			
32	T						T				G			T	C	A	C										G			C			G							G	G			.1C	.1C		
33	T						T							T	A	C								C		T	G				C			G						G	G			.1C	.1C		
37							T							T	A	C									T	G				C			G							G	G			.1C	.1C		
42							T				G			T	C	A	C										G				C			G							G	G			.1C		
46		T		C			T							T	A	T	C								T	G				C			G								G	G			.1C	.1C	
50	C	T					T							T	A	C									T	G				C											G	G			.1C	.1C	
51	T						T							T	A	C								C			G				C		C				T~ C					G	G	G		.2C	.1C

Haplogrupo A (Continuación)

	16,111	16,136	16,156	16,172	16,175	16,209	16,210	16,223	16,234	16,239	16,261	16,278	16,290	16,292	16,300	16,311	16,319	16,325	16,362	16,390	16,391	16,519	41	59	62	64	73	94	97	103	106	107	108	109	110	111	143	146	150	152	153	183	185	189	194	235	238	263	308	310	314	348
54	T				C	T							T			A	C	C			C					T	G								C	T		G			G		G		G		.2	.1				
57	T						T						T			A	C				C							G							C		G			G		G		.1	.1							
58	T		C		C	T	T	T					T			A	C									T	G							C	C				G		G		.1	.1								
59	T	C					T						T			A	C										T	G						C		G			G		G			.1	.1							
64	T						T					T	T			A	C										T	G						C		G	G			G		G		.1	.1							
65	T						T		T				T			A	C	A			T						T	G						C		G			G		G		.2	.1								
66	T	A					T						T			C	A	C									T	G						C		G			G		G		.1	.1								
68	T						T						T			A	C											G		A				C	C	G			G		G		.2	.1								
69	T			G			T						T	G		A	C											G						C					T	G		G		.1	.1							
70	T												T			A	C	A									T	G						C		G			G		G		.1	.1								
74	T						T						T			A	C										T	G						C		G			G		G			.1	.1							
75	T						T						T			A	C										T	G							A	C	C	G		A		G			.1	.1						
76							T						T			A	C										T	G							C		G			G	G		G		.1	.1						
77	T				C		T						T			A	C										T	G	A						C		G			G		G		.2	.1							
78	T	C					T						T	T		C	A	C										G							C		G			G		G			.1	.1						
81	T				C		T						T			A	C											G							C		G			G		G		.2	.1							
83	T												T			A	C	A									T	G							C		G			G		G		.1	.1							
84	T						T						T			A	C	A									T	G							C		G			G		G			.1	.1						
85	T						T						T			A	C										T		A	d	d	d	d	d	d			C		G			G		G			.1	.1			
86							T						T			C	A	C									T	G								C					G		G		.2	.1						

Haplogrupo A (Continuación)

	16,093	16,111	16,114	16,129	16,136	16,150	16,152	16,153	16,172	16,187	16,189	16,192	16,209	16,223	16,224	16,239	16,240	16,248	16,261	16,263	16,274	16,284	16,290	16,291	16,311	16,319	16,325	16,356	16,362	16,371	16,390	16,391	16,519	16,536	41	64	73	146	152	153	195	215	221	226	235	253	263	308	311	314	316	367
131													T									T			A		C			C			T	G	C					G		G		.1		.1						
132	T		A										T	C								T			A	C	C						T	G	C		G	C		G		G		.1		.1						
133	T												T									T	T		A		C			C			T	G	C		G			G		G		.1		.1						
135	T												T									T			A		C			C			T	G	C		G			G		G		.1		.1						
136	T			C				C					T									T		C	A		C			C			T	G	C	C	G			G		G		G		.1		.1				
138	T												T		T							T			A		C						T	G	C	C	G			G		C-T		G		.1		.1				
142	T												T					A				T			A		C						T	G	C		G			G		G				.1		.1				
144	T												T				T					T			A		C		A					G	C		G			G		G		C		.2		.1				
146	T												T									G		T		A		C					T	G	C	C	G			G		G		G		.2		.1				
147	T												T						C			T			A		C	G			C	T		G	C		G		G		G		G		.1		.1					
148	T								.1	C			T									T		C	A		C								G	C		G			C		G		G		.2		.1			
149	T												T					A				T			A		C								G	C		G			G		G		.1		.1					
152	C	T	T										T									T			A		C	C						T	G	C				G		G		.1		.1						
157	T												T									T			A		C			C				T	G	C		G			G		G		.1		.1					
159	T												T					A				T			A		C			C				T	G	C					G		G		.1		.1					
160	T											T	T									T			A		C		A	C				T	G	C		G			G		G				.1		.1			
165	T										C		T									T	T		A		C			C				T	G	C		G			G		G				.1		.1			
168					C	A						C	T				G					T			A		C							T	G	C		G			G		G				.1		.1			
173	T												T									T			A		C								G	C		G			G		G		.2		.1			.1		
174	T				T								T									T			A		C								G	C		G			C		G		G		.1		.1			A

Haplogrupo A (Continuación)

	16,092	16,093	16,104	16,111	16,114	16,124	16,136	16,156	16,172	16,189	16,192	16,223	16,240	16,241	16,243	16,248	16,259	16,265	16,272	16,274	16,290	16,291	16,294	16,300	16,311	16,319	16,356	16,362	16,391	16,399	16,400	16,463	16,468	16,469	16,519	16,527	64	73	103	146	152	153	189	194	200	235	263	308	314					
176	C			T								T								T	C					A		C	A	G						T	G												.1	C				
177		C		T	T							T									T					A	C	C																							.1	.1	C	C
182				T		C						T									T				C	A		C				G					T	G		C		G							.1	.1	C	C		
183				T								T		C						A	T					A		C							T		G		C		G								.1	C				
184				T								T									T		G			A		C																						.1	.1	C	C	
187				T								T	G		T						T								C								T	G		C		G							.1	C				
188				T								T									T					A		C																						.1	.1	C	C	
190				T						C											T					A		C																							.1	C		
191				T		C						T									T					A		C																							.1	C		
193				T								T							G		T					C	A		C																					.2	C			
194				T			A					T									T				C	A		C																						.2	C			
196												T	G								T					A		C																						.1	C			
201												T									T		T			A		C																						.1	C			
206				T		C		C				T									T					A		C																						.1	C			
208				T								T						A	G		T				C	A		C																					.2	C				
209				T								T											G			A		C																						.1	C			
211				T								T									T					A		C																						.1	C			
212			A	T							T	T									T					A		C																						.1	C			
214				T								T									T					A		C																						.1	C			
215				T								T				T					T					A		C																						.1	C			

Haplogrupo A (Continuación)

	16,111	16,124	16,136	16,153	16,177	16,210	16,223	16,240	16,288	16,290	16,311	16,319	16,335	16,348	16,362	16,519	16,526	64	72	73	103	143	146	151	153	195	198	199	204	235	236	263	308	314
216	T						T			T		A			C					G			C		G			C		G		G	.1C	.1C
218	T				G		T			T		A			C			T		G			C		G				G		G	.1C	.1C	
219	T		C				T			T	C	A			C			T		G			C		G		T		G		G	.1C	.1C	
221	T						T			T		A				A	T			G			C		G				G		G	.1C	.1C	
223							T			T										G								C		C	G		.1C	
226							T			T	C	A			C					G			C		G				G		G	.1C	.1C	
227	T					C	T			T		A			C					G	A		C		G				G		G	.1C	.1C	
230							T	G		T		A			C	C		T	C	G			C		G	C			C	G		G	.1C	.1C
237	T						T		C	T		A		T	C	C		T		G			C		G				G		G	.2C	.1C	
238				A			T	G		T		A			C			T		G			C						G		G	.1C	.1C	
243	T	C					T			T		A			C	C		T		G			C	T	G				G		G	.1C	.1C	
245	T						T			T		A	G				A			G		G~A	C		A~G				G		G	.1C	.1C	
249	T				G		T			T		A			C			T		G			C		G				G		G	.1C	.1C	

Haplogrupo B

	16,075	16,086	16,092	16,104	16,111	16,140	16,164	16,182	16,183	16,189	16,217	16,223	16,233	16,262	16,278	16,294	16,295	16,298	16,311	16,319	16,357	16,362	16,391	16,451	16,467	16,468	16,483	16,497	16,519	16,527	70	73	75	114	146	152	168	185	200	215	228	263	308	310	314			
3						G		C	C	C																		A		G												.1	.1					
14				A				C	C	C	C																		C		G												.2	.1				
16					T			C	C	C	C					T											A		C	T	G					C							.1	.1				
17									C	C	C				T														C		G												.3	.1				
20									C	C	C																		C		.1 G	G	A	G										.3	.1			
29								C	C	C	C						T											G		C		G												.1	.1			
30									C	C	C				T														C		G													.1	.1			
35									C	C	C							C											C		G													.1	.1			
38								C	C	C	C										A								C		G													.2	.1			
44								C	C	C	C																		C		G													.1	.1			
45									C	C	C																		C		G								G					.1	.1			
49								C	C	C	C		G																C		G														.1	.1		
52								C	C	C	C											C							C		G								G					.2	.1			
53							G		C	C	C																		A		G													.1	.1			
55								C	C	C	C				T	T													C		G													.2	.1			
60		c						C	C	C	C																		C		G													.2	.1			
73									C	C	C																	A	T		C		G												A	G	.1	.1
80								C	C	C	C	T																		C		G													A	G	.3	.1
82		c							C	C	C					T														C		G														.1	.1	
92									C	C	C											C								C		G													.1	.1		
94									C	C	C	C																		C		G													.3	.1		
101									C	C	C											C							T	C		G														d	.1	
110									C	C	C	C					T													G		C														.1	.1	
115									C	C	C	C										A								C		G														.2	.1	

Haplogrupo C

	16,051	16,129	16,172	16,181	16,188	16,189	16,204	16,218	16,223	16,266	16,274	16,298	16,311	16,325	16,327	16,336	16,354	16,362	16,381	16,399	16,519	16,526	16,527	57	58	60	73	146	152	182	195	198	199	204	207	215	241	247	249	252	263	288	290	291	308	314	333
2								T	A	C	C	T									C						G																	.1	.1		
7								T			C	C	T					C			C						G																				
9								T			C	C	T	A									C		.1	C	G																			.2	.1
12								T			C	C	T														G																		.1	.1	
18								T			C	C	T						C	G							G																		.1	.1	
22								T	T		C	C	T														G																		.1	.1	
34								T			C	C	T									C					G		C																		
39			C					T			C	C	T														G	C																	.1	.1	
40								T			C	C	T														G			T															.1	.1	
47	G				T		A	T				C	T					C			C	T					G		C																.1	.1	
56								T			C	C	T														G				T														.1	.1	
62								T			C	C	C	T													G																				
67					C		T	T			C	C	T														G																			.1	.1
71			C					T			C	C	C	T													G		C																	.1	.1
96		A	C					T			C	C	C	T								C					G	C	C																.1	.1	
100								T			C	C	T				T										G																		.1	.1	
106				G				T			C	C	T														G																		.2	.1	
107	G				T		A	T				C	T					C			C	T					G																		.1	.1	
118	G				T		A	T				C	T					C			C	T					G																		.1	.1	
120								T			C	C	C	T														G																		.1	.1
122								T			C	C	T															G																		.1	.1
123								T			C	C	T														G	C		T																.1	.1

Haplogrupo C (Continuación)

	16,051	16,129	16,144	16,153	16,172	16,181	16,188	16,189	16,192	16,204	16,209	16,223	16,298	16,311	16,319	16,325	16,327	16,336	16,362	16,519	16,527	57	58	73	127	146	152	189	195	198	214	215	241	246	249	263	290	291	308	314		
126												T	C			C	T							G	C						G				d	G	d	d	.2C	.1C		
129												T				C	T							G						T					d	G	d	d	.1C	.1C		
130												T	C			C	T							G											d	G	d	d		.1C		
137												T	C			C	T							G					T						d	G	d	d	.1C	.1C		
139		A			C							T	C	C		C	T		C				G		C	C								d	G	d	d	.2C	.1C			
145												T	C			C	T		C					G											d	G	d	d	.1C	.1C		
150												T	C			C	T							G									C		d	G	d	d	.2C	.1C		
153	G								T			T	C			C	T		C					G				G							d	G	d	d	.1C	.1C		
154	G							C				T	C	C		C	T		C					G					C						d	G	d	d	.1C	.1C		
156		A			C							T	C	C		C	T		C				G		C	C									d	G	d	d	.2C	.1C		
158				A								T	C			C	T		C					G												d	G	d	d	.1C	.1C	
162																C	T							G							G					d	G	d	d		.1C	
166												T	C		A	C	T							G												d	G	d	d	.1C	.1C	
169												T	C			C	T							G												d	G	d	d	.2C	.1C	
178												T	C			C	T		C					G												d	G	d	d	.2C	.1C	
185				C								T	C			C	T							G												d	G	d	d		.1C	
195	G									C		T	C			C	T							G												d	G	d	d	.1C	.1C	
197				A								T	C			C	T							G													d	G	d	d	.2C	.1C
199												T				C	T							G													d	G	d	d	.1C	.1C
203	G					T				A		T				C	T		C	C	T			G												d	G	d	d	.1C	.1C	
205												T	C			C	T							G													d	G	d	d		.1C
228												T	C			C	T							G													d	G	d	d	.1C	.1C
246					G							T	C			C	T	A				C	.1C	G					C			G				d	G	d	d	.2C	.1C	
250					G							T	C											G													d	G	d	d		.1C

Haplogrupo D

	16,051	16,093	16,104	16,155	16,223	16,256	16,274	16,278	16,291	16,292	16,325	16,357	16,362	16,368	16,390	16,465	16,519	73	140	146	150	204	207	211	214	263	267	308	314
25					T					T	C		C			T	C	G								G		.1C	.1C
26					T						C		C				C	G								G		.2C	.1C
27		C			T		A				C		C					G				C				G			.1C
36					T						C		C				C	G								G		.1C	.1C
41					T						C		C				C	G								G		.1C	.1C
48					T						C		C				C	G								G		.1C	.1C
61			T		T						C	C	C					G	T							G		.1C	.1C
91					T						C		C		A		C	G			T					G			.1C
105					T						C		C				C	G								G		.2C	.1C
114					T						C		C				C	G								G		.1C	.1C
151		C			T				T		C		C					G		C						G			.1C
167					T						C		C		A		C	G			T					G			.1C
180					T						C		C					G								G			.1C
204	G				T						C		C					G								G			.1C
213				A~G	T						C		C	C			C	G						G		G		.2C	.1C
224					T						C		C					G		C			A			G		.1C	.1C
231			T		T	T					C	C	C					G								G			.1C
242					T						C		C			T		G							G	G	C		.1C
251					T			T			C		C					G		C						G		.1C	.1C

Haplogrupo L3

	16,124	16,167	16,189	16,209	16,223	16,278	16,292	16,295	16,307	16,311	16,320	16,362	16,519	16,527	73	150	152	189	195	263	308	314
171	C				T	T						C		T	G					G	.1C	.1C
239			C	C			T	T		C			C		G			G		G		.1C
241		T			T				G		T		C		G	T	C		C	G		.1C

Haplogrupo T

	16,114	16,126	16,153	16,163	16,186	16,189	16,192	16,266	16,294	14,304	16,519	64	73	150	152	195	198	200	253	263	308	314
5		C							T		C	T	G						T	G	.2C	.1C
43		C						T	T	C	C		G				T	G		G	.1C	.1C
72	T	C	A				T		T		C		G	T						G	.1C	.1C
143		C		G	T	C			T	C	C		G		C	C				G		.1C
172		C		G	T	C			T		C		G		C	C				G	.2C	.1C
192		C		G	T	C			T	C	C		G		C	C				G		.1C
210																						

Haplogrupo U

	16,083	16,086	16,088	16,134	16,145	16,172	16,187	16,192	16,219	16,235	16,270	16,278	14,304	16,343	16,356	16,391	16,519	73	150	152	195	228	251	263	296	314
189			C		A	C			G	G		T					C	G						G		.1C
235		C						T			T		C					G	T			A	A	G		.1C
247				T											C		C	G		C	C			G	T	.1C
248	T						T							G		A		G	T	C				G		.1C

Haplogrupo W

	16,223	16,292	16,519	73	189	194	195	199	204	207	263	308	314
108	T	T	C	G	G	T	C	C	C	A	G	.1C	.1C

ANEXO 4. ESTIMADORES DE FLUJO GÉNICO

Tabla 9. Matriz de valores p de los Fst y su significancia. Debajo de la diagonal, se muestran los valores p de los Fst entre las dieciséis alcaldías de la CDMX, sobre la diagonal se muestra si dicho valor obtenido es significativo (+) o no significativo (-). Los asteriscos (*) marcan el cruce entre la misma alcaldía, razón por la cual no se registra información. En rojo se muestran los valores significativos.

	AZC	COY	CUJ	XOH	TLH	GAM	IZC	BJ	IZP	AO	VC	MC	MIL	TLP	CUH	MH
AZC	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
COY	0.9053	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CUJ	0.8828	0.9775	*	-	-	-	-	-	-	-	-	-	-	-	-	-
XOH	0.459	0.6318	0.8086	*	-	+	-	-	-	-	-	-	-	-	-	-
TLH	0.3799	0.5606	0.6445	0.7725	*	+	-	-	-	-	-	-	-	+	-	-
GAM	0.1289	0.1406	0.7246	0.0293	0.0127	*	-	-	+	-	-	+	-	-	+	-
IZC	0.8281	0.793	0.9424	0.3926	0.2588	0.7617	*	-	-	-	-	-	-	-	-	-
BJ	0.9775	0.9688	0.9697	0.7158	0.501	0.1943	0.9531	*	-	-	-	-	-	-	-	-
IZP	0.8164	0.7266	0.9141	0.3574	0.2461	0.0273	0.834	0.7598	*	-	-	-	-	+	-	-
AO	0.7275	0.9736	0.9951	0.6523	0.4414	0.2451	0.9873	0.9981	0.501	*	-	-	-	-	-	-
VC	0.7481	0.4043	0.7832	0.1641	0.1045	0.7529	0.9805	0.8496	0.3965	0.7822	*	-	-	-	-	-
MC	0.3135	0.207	0.2969	0.5215	0.4648	0.0107	0.1231	0.3145	0.169	0.1856	0.0518	*	-	+	-	-
MIL	0.9922	0.9473	0.7949	0.917	0.9121	0.3682	0.8828	0.9756	0.9092	0.9424	0.751	0.9824	*	-	-	-
TLP	0.165	0.2422	0.7451	0.0693	0.0273	0.8721	0.624	0.4707	0.0342	0.4639	0.5557	0.0059	0.2959	*	+	-
CUH	0.3174	0.4033	0.7969	0.9004	0.6631	0.0215	0.2959	0.6309	0.1055	0.5127	0.1377	0.2227	0.8643	0.0449	*	-
MH	0.957	0.7109	0.7334	0.6387	0.4629	0.083	0.7334	0.957	0.6045	0.7402	0.5811	0.499	0.9775	0.1582	0.5596	*

ABREVIATURAS: Azcapotzalco (AZC), Coyoacán (COY), Cuajimalpa (CUJ), Xochimilco (XOH), Tláhuac (TLH), Gustavo A. Madero (GAM), Iztacalco (IZC), Benito Juárez (BJ), Iztapalapa (IZP), Álvaro Obregón (AO), Venustiano Carranza (VC), Magdalena Contreras (MC), Milpa Alta (MIL), Tlalpan (TLP), Cuahutémoc (CUH) y Miguel Hidalgo (MH).

Tabla 10. Matriz de valores de Fst pareadas y distancias geográficas. Debajo de la diagonal, se muestran las distancias genéticas (Fst pareadas) entre las dieciséis alcaldías de la CDMX, sobre la diagonal se muestra la matriz de las distancias geográficas entre las dieciséis alcaldías de la CDMX (km). Los asteriscos (*) marcan el cruce entre la misma alcaldía, razón por la cual se registra información.

	AZC	COY	CUJ	XOH	TLH	GAM	IZC	BJ	IZP	AO	VC	MC	MIL	TLP	CUH	MH
AZC	*	17.97	22.41	28.42	29.87	7.26	13.48	11.86	20.07	17.94	11.16	25.73	40.89	32.05	6.94	6.8
COY	0.035	*	16.85	11.05	16.46	20.08	9.79	6.11	10.14	10.19	13.03	13.97	23.31	15.44	11.66	12.64
CUJ	0.058	0.0773	*	24.79	32.78	28.59	24.11	16.9	26.81	6.84	25.71	7.63	34.2	17.84	20.72	16.03
XOH	0.005	0.019	0.0519	*	9.86	28.95	16.9	16.82	12.1	19.31	20.64	18.9	12.47	13.25	21.63	23.63
TLH	0.002	0.0161	0.0397	0.0345	*	27.93	16.45	20.24	9.83	26.47	19.54	27.92	16.13	23.1	23.05	27.05
GAM	0.0274	0.02148	0.029	0.0694	0.0923	*	12.14	14.54	18.32	23.2	8.53	30.68	41.15	35.33	8.81	12.58
IZC	0.036	0.0326	0.0775	0.0028	0.0206	0.02	*	7.24	6.62	17.38	3.74	23.18	29.02	25.04	6.91	12.08
BJ	0.045	0.0405	0.0751	0.0246	0.009	0.0138	0.0443	*	11.5	10.27	9.05	16.78	29.23	20.84	5.79	6.98
IZP	0.022	0.0135	0.0524	0.0015	0.0105	0.039	0.0241	0.0179	*	20.01	9.82	23.96	23.35	22.99	13.33	17.83
AO	0.022	0.0302	0.0734	0.0187	0.003	0.0064	0.0432	0.0406	0.005	*	19.26	7.82	30.05	15.93	14.75	11.15
VC	0.026	0.0045	0.045	0.0334	0.0529	0.018	0.0502	0.0295	0.002	0.023	*	25.74	32.7	28.44	5.87	11.7
MC	0.0087	0.02777	0.0114	0.0146	0.011	0.1388	0.0616	0.0127	0.031	0.038	0.0837	*	27.01	10.23	22	18.94
MIL	0.105	0.1084	0.1085	0.0992	0.099	0.003	0.0967	0.1018	0.077	0.085	0.056	0.111	*	17.58	34.09	35.95
TLP	0.0263	0.01197	0.0364	0.0551	0.0866	0.015	0.0157	0.0075	0.036	0.004	0.012	0.150	0.0107	*	26.63	25.57
CUH	0.0071	0.00218	0.0379	0.0311	0.018	0.0688	0.00719	0.0135	0.0253	0.007	0.0288	0.0225	0.061	0.0512	*	5.84
MH	0.05	0.0268	0.045	0.021	0.007	0.0412	0.0298	0.0461	0.014	0.024	0.017	0.01	0.117	0.0281	0.0127	*

ABREVIATURAS: Azcapotzalco (AZC), Coyoacán (COY), Cuajimalpa (CUJ), Xochimilco (XOH), Tláhuac (TLH), Gustavo A. Madero (GAM), Iztacalco (IZC), Benito Juárez (BJ), Iztapalapa (IZP), Álvaro Obregón (AO), Venustiano Carranza (VC), Magdalena Contreras (MC), Milpa Alta (MIL), Tlalpan (TLP), Cuahutémoc (CUH) y Miguel Hidalgo (MH).

ANEXO 5. REDES DE HAPLOTIPOS DE mtDNA

Haplotipos nativo americanos (Haplogrupo A)

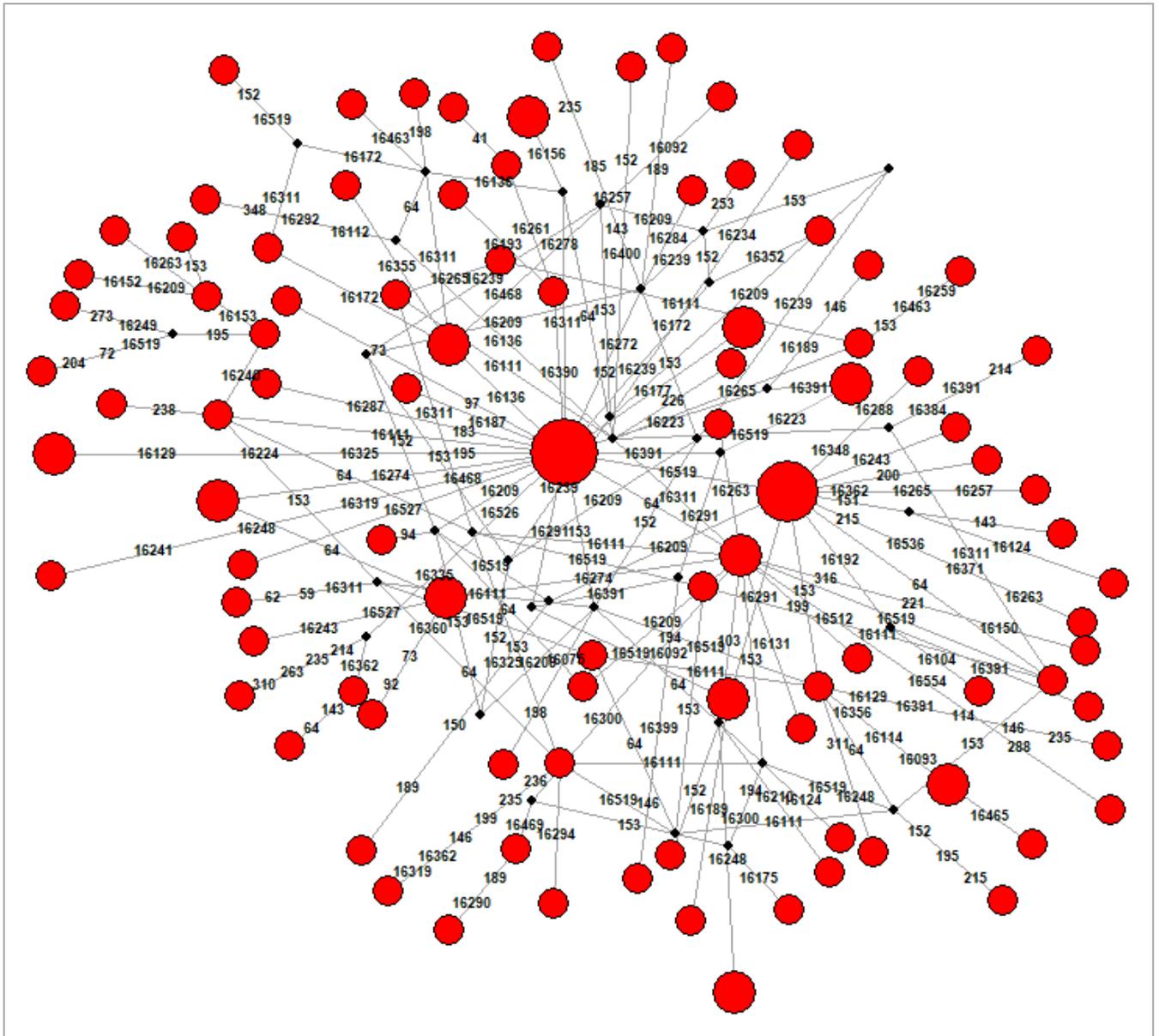


Fig. 16. Red media de haplotipos nativo americanos pertenecientes al haplogrupo A. Cada uno de 96 haplotipos pertenecientes al haplogrupo A, se representan con un nodo (círculo rojo), el tamaño del mismo es proporcional a la frecuencia del haplotipo en la población; los números representan la posición de las mutaciones que presenta cada haplotipo y los vectores medios se encuentran representados por los puntos negros.

Haplotipos nativo americanos (Haplogrupo B)

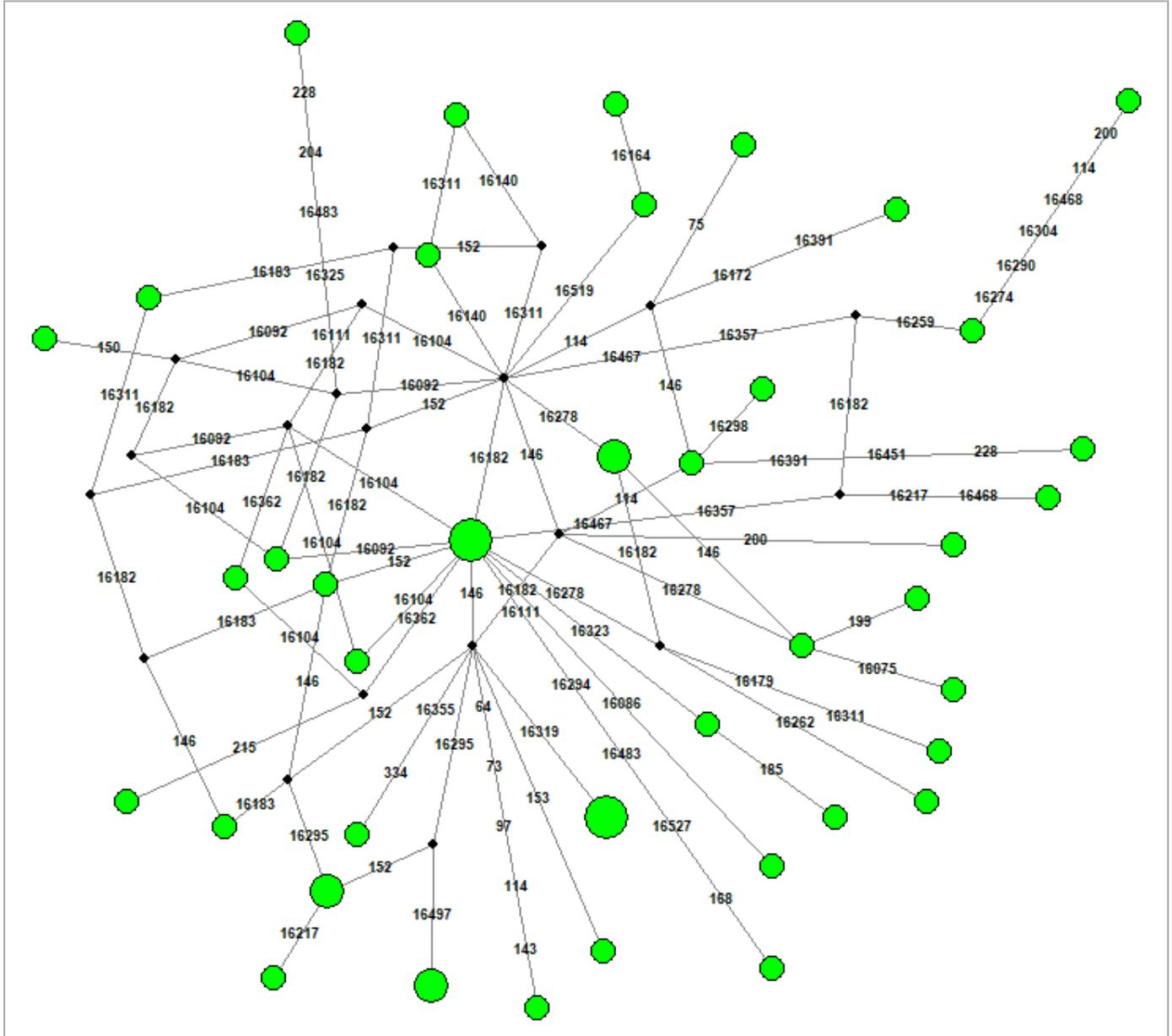


Fig. 17. Red media de haplotipos nativo americanos pertenecientes al haplogrupo B. Cada uno de 40 haplotipos pertenecientes al haplogrupo B, se representan con un nodo (círculo verde), el tamaño del mismo es proporcional a la frecuencia del haplotipo en la población; los números representan la posición de las mutaciones que presenta cada haplotipo y los vectores medios se encuentran representados por los puntos negros.

Haplotipos nativo americanos (Haplogrupo C)

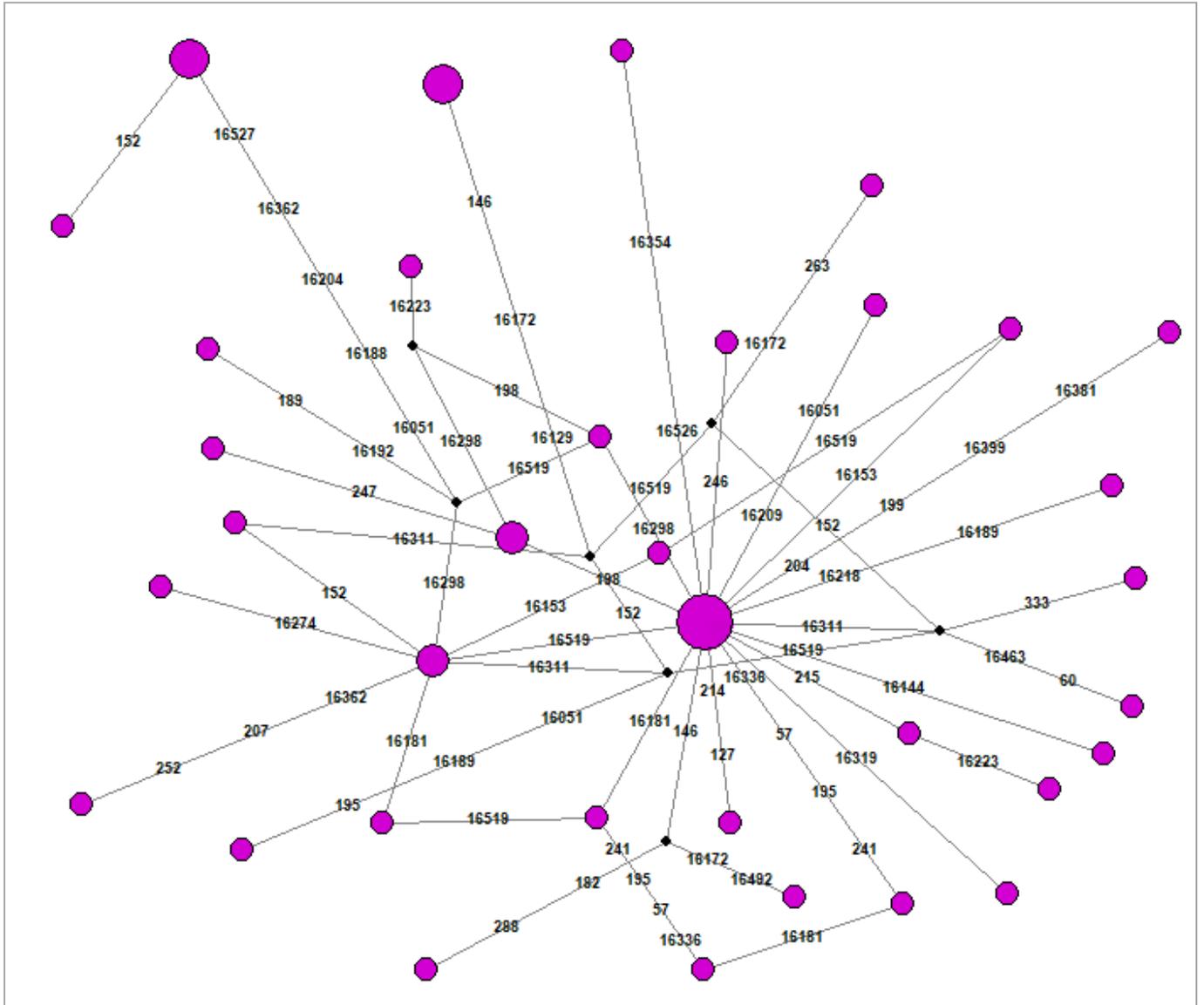


Fig. 18. Red media de haplotipos nativo americanos pertenecientes al haplogrupo C. Cada uno de 35 haplotipos pertenecientes al haplogrupo C, se representan con un nodo (círculo lila), el tamaño del mismo es proporcional a la frecuencia del haplotipo en la población; los números representan la posición de las mutaciones que presenta cada haplotipo y los vectores medios se encuentran representados por los puntos negros.

Haplotipos nativo americanos (Haplogrupo D)

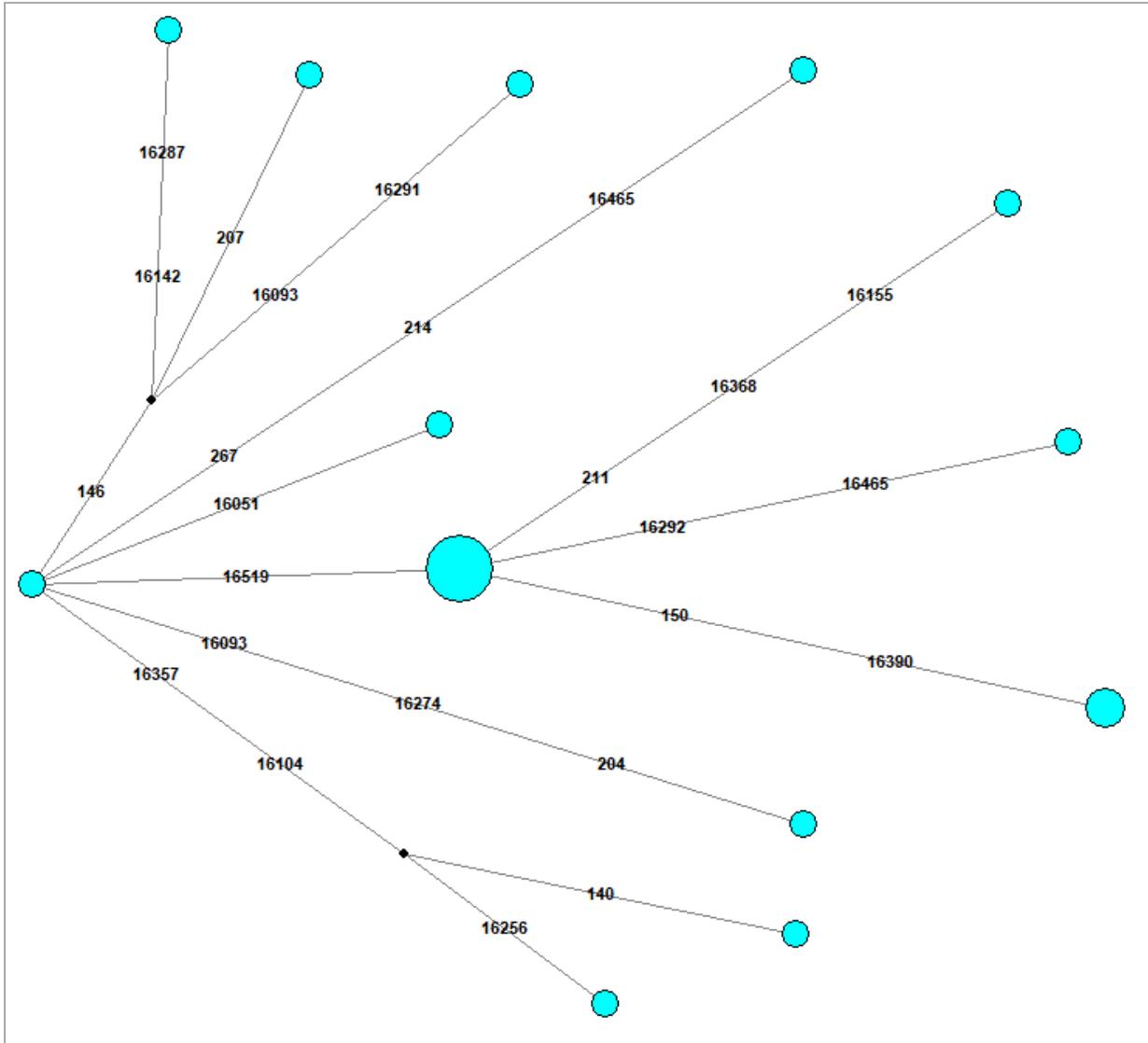


Fig. 19. Red media de haplotipos nativo americanos pertenecientes al haplogrupo D. Cada uno de 13 haplotipos pertenecientes al haplogrupo D, se representan con un nodo (círculo turquesa), el tamaño del mismo es proporcional a la frecuencia del haplotipo en la población; los números representan la posición de las mutaciones que presenta cada haplotipo y los vectores medios se encuentran representados por los puntos negros.

Haplotipos europeos y africanos

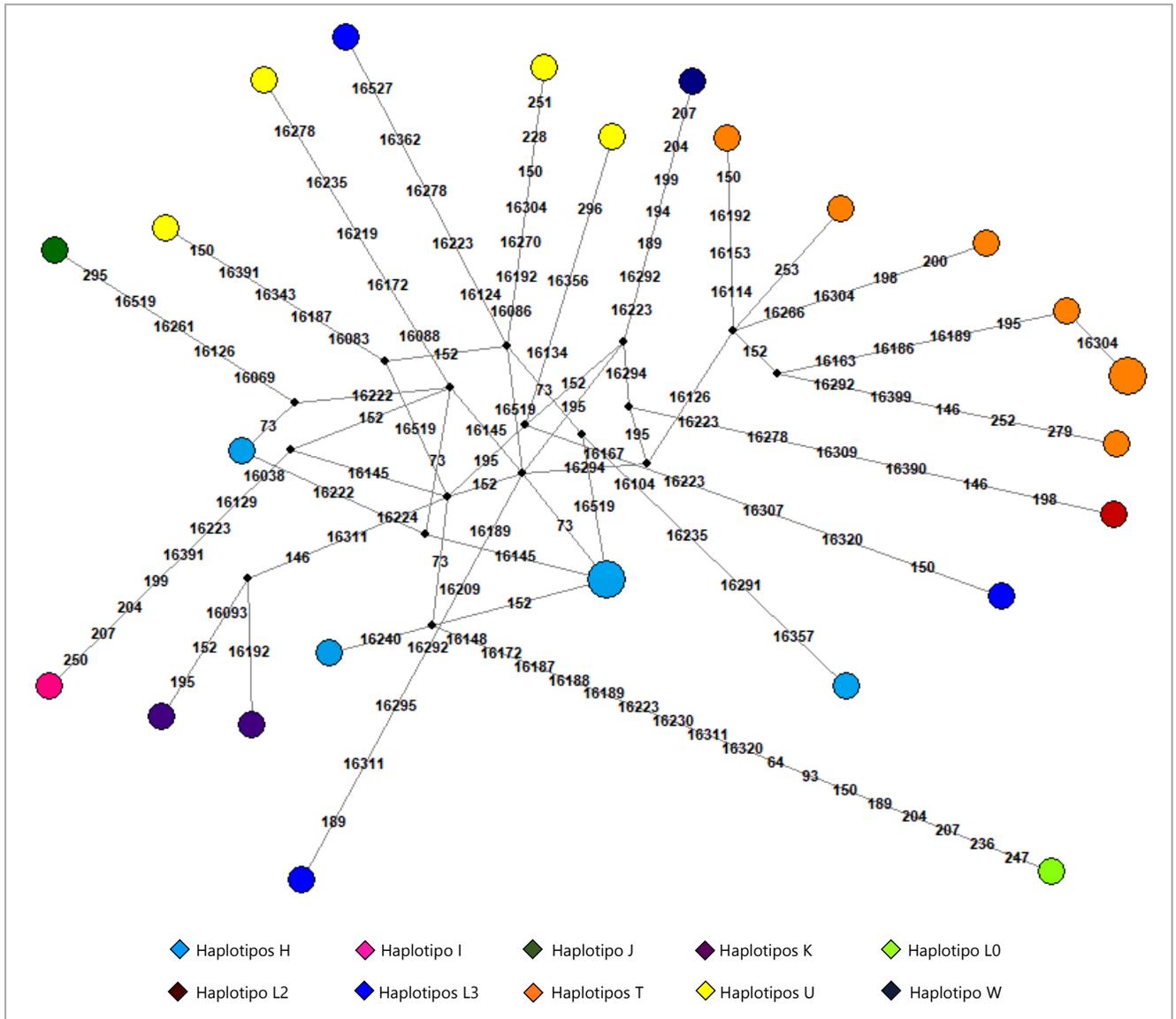


Fig. 20. Fig. V. Red media de haplotipos nativo americanos pertenecientes a los haplogrupos europeos y africanos. Cada uno de 26 haplotipos pertenecientes al haplogrupos europeos y africanos, se representan con un nodo (círculo de color de acuerdo al haplogrupos), el tamaño del mismo es proporcional a la frecuencia del haplotipo en la población; los números representan la posición de las mutaciones que presenta cada haplotipo y los vectores medios se encuentran representados por los puntos negros.