



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO

---

---

FACULTAD DE CIENCIAS

ESTRUCTURA GENÉTICA DE UNA POBLACIÓN  
MESTIZO MEXICANA, CON BASE EN CINCO  
MARCADORES NUCLEARES Y CINCO  
MITOCONDRIALES.

T E S I S

QUE PARA OBTENER EL TÍTULO DE

BIÓLOGA

P R E S E N T A :

ALICIA HUERTA CHAGOYA

TUTORA:

M en C. WENDY ARGELIA GARCÍA SUASTEGUI

COTUTOR:

DR. LUIS ENRIQUE EGUIARTE FRUNS

2007





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.

<p>1. Datos del alumno Huerta Chagoya Alicia 56918922 Universidad Nacional Autónoma de México Facultad de Ciencias Biología 30016365-8</p>
<p>2. Datos del tutor M en C. Wendy Argelia García Suastegui</p>
<p>3. Datos del cotutor Dr. Luis Enrique Eguiarte Fruns</p>
<p>4. Datos del sinodal 1 Dra. María Eugenia Gonsebatt Bonaparte</p>
<p>5. Datos del sinodal 2 Dr. Luis Medrano González</p>
<p>6. Datos del sinodal 3 Dra. Julieta Rubio Lighthbourn</p>
<p>7. Datos del trabajo escrito Estructura genética de una población mestizo mexicana, con base en cinco marcadores nucleares y cinco mitocondriales. 60p. 2007</p>

## DEDICATORIA

Pues bien, terminado el primer paso aunque aún falta mucho. Me enorgullece ser parte de la maravillosa Universidad Nacional Autónoma de México porque no sólo forma profesionistas, sino también seres humanos; espero algún día poder restituirle todo lo que he aprendido y seguiré aprendiendo en ella.

Jamás hubiera logrado esto sin el infinito amor, confianza, paciencia y apoyo de mis patos. Chinita, gracias por ser amigui, maestra, psicóloga, paño de lágrimas, enfermera, nutrióloga... pero sobre todo, la mejor mami del mundo. Gracias a ambos por creer en mí.

Va también para mis dos hermanos. Cali, nadie me pudo haber cuidado mejor que tú de pequeña, gracias por enseñarme a escribir... Betocha, tienes el corazón más grande del mundo.

A la tía Mari, que se sacó un 10 con mi hermanito. Julito, Egarito, Mauri y Adriana, gracias por haber donado su precioso líquido rojo sólo por ayudar a su prima y, por el contacto asiático. Tíos Medina, Karla y Valeria, gracias por ser más que familia. Sanpanchenses, son la opción!!

Gracias bicho por estar siempre presente, el camino no será más el mismo sin ti. Eres un gran regalo.

A Viri y Beto, aunque lejos siempre están conmigo. Arañita, tus carcajadas siempre fueron y serán de gran ayuda; Beto, eres maravilloso. Gracias a ambos por encender la luz.

Carmín, eres una gran lección de vida; este trabajo es compartido por ser tan valiente.

A todos los bioloquitos que hicieron mi vida mágica durante la carrera y la seguirán haciendo en adelante, especialmente a Loris, Mon y Fórido, la Biología no hubiera sido tan divertida sin ustedes. Drosi y Manis, siempre están en mi corazón. Es extensa la banda, pero gracias a todos por cada segundo que pasé con ustedes.

A la Dra. Gonsebatt, gracias por haberme permitido pertenecer a su grupo de trabajo, y por todo el apoyo y confianza que depositó en mí. Dra. Rubio, gracias por sus enseñanzas y por las palabras de aliento en todo momento.

A mi amiga y tutora de tesis, Wen, gracias por compartir conmigo más que técnicas de laboratorio.

A Rebe, que me introdujo en el mundo de los polimorfismos y, a Isa con la que compartí tantos tubos a rotular. A Marianis, Pavel, Manu, César, Miriam, Elsita, Valeria, Danny, Luisrael y Danna, que en algún momento me enseñaron algo; y a los más nuevitos, Karol, Camilo y Celeste, todos compañeros con los que trabajé en el laboratorio tanto tiempo.

Sara, Cu y Pel, gracias por prestar orejas.

Y por último y no menos importantes, puesto que sin ellos no se hubiera obtenido un solo dato, a todos los amables y voluntarios donadores. Sus muestras no han sido desperdiciadas.

#### AGRADECIMIENTOS

Este trabajo fue realizado en el Laboratorio de Medicina Genómica y Toxicología Ambiental, del Instituto de Investigaciones Biomédicas de la UNAM, a cargo de la Dra. Ma. Eugenia Gonsebatt Bonaparte; bajo la tutoría de la M. en C Wendy Argelia García Suastegui y cotutoría del Dr. Luis Enrique Eguiarte Fruns.

Se realizó bajo el financiamiento del proyecto número 46341-M del CONACYT, otorgado a la Dra. Gonsebatt.

Se agradece el apoyo del Dr. Luis Medrano González en el análisis de los datos y, de la M. en C García y del Dr. Pavel Petrosyan en la realización de los experimentos.

## ÍNDICE

### 1. INTRODUCCIÓN

1.1. Población	1
1.2. Procesos evolutivos que actúan sobre una población	2
1.3. Estratificación de la población	4
1.4. Marcadores genéticos	6
1.4.1. AIM	7
1.4.2. ADNmt	7
1.5. Divergencia de la población humana	8
1.6. Divergencia en el continente americano	10
1.7. Haplotipos mitocondriales	11
1.8. Población mexicana	14
1.9. Marcadores nucleares	14
1.9.1. Citocromos P450	15
1.9.2. GST	16
1.10. Marcadores mitocondriales	17
1.10.1. AIM de Europa	17
1.10.2. AIM de Asia	17
1.10.3. AIM de África	17
1.10.4. AIMs de América	18
1.11. Estudios de mezcla y variación étnica	18

JUSTIFICACIÓN 21

OBJETIVOS 22

HIPÓTESIS 22

### 2. MATERIAL Y MÉTODOS

2.1. Muestreo de individuos	23
2.2. Purificación de ADN	23
2.3. Genotipificación	23
2.3.1. Marcadores nucleares	24
2.3.1.1. CYP 1A1	24
2.3.1.2. GSTs	25
2.3.2. Marcadores genéticos mitocondriales	26
2.3.2.1. Marcador mitocondrial europeo	26
2.3.2.2. Marcadores mitocondriales amerindios	26
2.3.2.3. Marcadores mitocondriales asiáticos	27
2.3.2.4. Marcador mitocondrial africano	29

### 3. RESULTADOS

3.1. Distribución geográfica de la muestra	30
3.2. Marcador AFR2	30
3.3. Frecuencias genotípicas y alélicas de los marcadores	32
3.4. Índice de Fijación	34
3.5. Diferencias de las frecuencias encontradas respecto a otras reportadas previamente	35
3.6. Haplotipos mitocondriales y su interpretación de ancestría	36
3.7. Asociación de marcadores mitocondriales vs. nucleares	38
3.8. Análisis de distancia nucleotídica vs. distancia geográfica	39

### 4. DISCUSIÓN

4.1. Flujo génico a las poblaciones mestizas mexicanas	42
4.2. Marcadores nucleares y su variación interétnica	43
4.3. Marcadores mitocondriales y sus estimados de ancestría	45
4.4. Asociación GST M1*0 – haplotipos africanos	48
4.6. Correlación distancia geográfica – distancia nucleotídica	50

5. CONCLUSIONES 53

BIBLIOGRAFÍA            54  
APÉNDICE I

- I. Purificación de ADN (según Daly et al., 1996) 59
- II. Reacción en cadena de la polimerasa (Polymerase Chain Reaction). PCR y PCR múltiple 59
- III. Polimorfismos de longitud de fragmentos de restricción. RFLP (Restriction Fragment Length Polymorphisms) 60

## LISTA DE ABREVIATURAS

ADN Ácido Desoxirribonucleico

ADNmt Ácido Desoxirribonucleico mitocondrial

AFR Marcador mitocondrial 3592 HpaI

AFR2 Marcador mitocondrial 7055 AluI

AHH Aril Hidrocarburo Hidroxilasa

AIM Marcadores informativos de ancestría

AME Marcador mitocondrial 663 HaeIII

ASI Marcador mitocondrial 10394 DdeI

CYP P450 Citocromos P450

DEL Marcador mitocondrial delección 9 pares de bases

EUR Marcador mitocondrial 7025 AluI

GST Glutación S-Transferasa

AAH Aminas Aromáticas Heterocíclicas

Hap Haplotipo mitocondrial

Hap-A, B, C, D y X Haplotipos mitocondriales correspondientes al continente americano

n.mt. Ningún marcador mitocondrial

HAP Hidrocarburos Aromáticos Policíclicos

pb Pares de bases

*f* Frecuencia

PCR Reacción en cadena de la polimerasa

RFLP Polimorfismos de longitud de fragmentos de restricción

SNP Polimorfismos de un solo nucleótido

## RESUMEN

La población mexicana se formó con la mezcla de muchos grupos étnicos. Estudios hechos con marcadores genéticos indican que su estructura genética presenta diferencias en ancestría según entidad de origen. Conocer las contribuciones genéticas de nuestra población permite hacer inferencias de susceptibilidad si se comprueba asociación entre genotipos de riesgo y haplotipos de ancestría. Algunos polimorfismos frecuentes en ciertas regiones han sido relacionados con padecimientos; pero antes debe saberse si la población estudiada es homogénea para reducir errores en los cálculos de sus frecuencias alélicas y asociaciones –muy comunes cuando se usan categorías establecidas socialmente, como el índice de homogeneidad genética en la delimitación de poblaciones. Para evitar confusiones, conviene considerar los estimados de ancestría de un genotipo dado.

Este trabajo pretende estudiar la estructura genética de una población mestiza mexicana, así como la asociación entre marcadores de ancestría y riesgo a desarrollar padecimientos como cáncer. Se tomaron muestras sanguíneas de 148 habitantes del DF y área metropolitana, los cuales, indicaron su lugar de nacimiento, el de sus padres y abuelos. Se purificó el ADN y se utilizó para tipificar 5 polimorfismos nucleares de genes involucrados en el metabolismo de xenobióticos (CYP P4501A1 \*2A, \*2C y \*4; GST M1 y T1) y 5 mitocondriales, que reportan la ancestría del individuo (EUR, ASI, AFR, AFR2, AME y DEL), mediante PCR y RFLP.

Aunque la población estudiada radica en el centro del país, su historia genética –según datos de sus familiares– se amplía a casi todo el país. Las frecuencias de los polimorfismos nucleares en orden descendente fueron: GSTM1\*0, \*2C, \*2A, GSTT1\*0 y \*4; las de los mitocondriales fueron: AME, DEL, ASI, AFR, AFR2 y EUR. Con base en los 6 últimos, se encontraron 11 haplotipos con ancestría específica. En orden descendente, la población presentó contribución amerindia, asiática, no mezcla, africana y europea. Se halló asociación entre el polimorfismo GST M1\*0 y haplotipos africanos. Gracias a un análisis de correlación distancias geográfica/nucleotídica y con datos de las abuelas maternas, se identificaron 3 grupos de estados (Sureste, Centro y Bajío). El primero presentó marcadas diferencias en la frecuencia del polimorfismo GST M1 respecto a los otros.

Se concluyó que la población mestiza mexicana ha sufrido altos flujos de migración y está formada por una gran mezcla de genotipos de ancestría variada; aunque la amerindia es la más abundante. Hay asociación entre el polimorfismo GST M1\*0 y haplotipos africanos. Es posible identificar regiones de aislamiento con historias genéticas diferentes.

## ABSTRACT

Mexican population is compounded from the mixture of many ethnic groups. By using genetic markers it is possible to demonstrate that, according to birthplace, there are ancestry differences in its genetic structure. If the association between risk genotypes and ancestry haplotypes is proved, we can make susceptibility inferences. This can be done by studying the genetic contributions of our population. Some of the most frequently reported polymorphisms in certain regions have been related with morbidities. However, it must be known if the population is homogenous in order to reduce errors in the estimation of its allelic frequencies and associations –which is very common when socially established categories are used as a genetic homogeneity index during the delimitation of populations. It may be helpful to take into account a genotype's ancestry estimates to avoid confusions.

In this project we study the genetic structure of a Mexican mestizo population, as well as the association between ancestry markers and the risk of developing morbidities such as cancer. Blood samples were taken from 148 residents of Mexico City and the Metropolitan Area, whom were asked to indicate their birthplace and also their parent's and grandparents' one. DNA was purified and, by means of PCR and RFLP, it was used to typify 5 nuclear polymorphisms of genes involved in the metabolism of xenobiotics (CYP P4501A1 \*2A, \*2C and \*4; GST M1 and T1) and 5 mitochondrial polymorphisms –the last ones reported individual ancestry (EUR, ASI, AFR, AFR2, AME and DEL).

Even though the studied population concentrates in the center of the country, its genetic history – according to their genealogical data– can be tracked down in almost all the states of the country. In descending order, the resultant nuclear polymorphism frequencies are as follows: GSTM1\*0, \*2C, \*2A, GSTT1\*0 and \*4; the resultant mitochondrial polymorphism frequencies are as follows: AME, DEL, ASI, AFR, AFR2 and EUR. Based in the last six frequencies, eleven haplotypes with specific ancestry were found. In descending order, the population got Amerindian, Asian, no mixed, African and European genetic contributions. An association between polymorphism GST M1\*0 and African haplotypes was found. No correlation between geographic and nucleotidic distances was found. Considering this and the genetic data of the individual's grandmothers on the mother's side, three regions (groups of states) were identified: the Southeast Region, the Center and the Bajío. In the first region, significant differences between polymorphism GST M1 frequency and the others were detected.

We concluded that Mexican mestizo population have passed through many migratory flows and is compounded by a varied mixture of ancestry genotypes, though Amerindian genetic contribution is the most abundant. There is also an association between polymorphism GST M1\*0 and African haplotypes. Some isolated regions with different genetic histories may exist.

## 1. INTRODUCCIÓN

### 1.1. Población

La población es un grupo completo de organismos de una especie que comparten hábitat a un tiempo y que se pueden entrecruzar; consecuentemente, es un amplio grupo del que podría tomarse una muestra (Gardner *et al.*, 2003). Algunos autores delimitan a las poblaciones por áreas geográficas o por idioma; por eso, la definición mencionada no es aplicable estrictamente a las poblaciones humanas, dados los altos índices de migración y entrecruzamiento que se registran. Por esta razón, resulta difícil hacer diferencias poblacionales estrictas; es decir, las poblaciones humanas son entidades continuas más que discretas.

Se ha observado una gran variación genética entre los grupos humanos, sin contar con la existencia de variantes genéticas presentes en un solo grupo específico. En función de esto se han propuesto términos que diferencian cada población de acuerdo al grado de variación entre ellas. Una de estas clasificaciones se propone a continuación:

Una *etnia menor* se refiere a grupos localmente constituidos que se saben distintos a otros (otras etnias menores), aunque éstos tengan muchos factores genómicos, culturales y ambientales en común. El término "*etnia mayor*" describe el contraste entre grupos que comparten ancestría, pero que han divergido en términos de cultura. El concepto "*ancestría*" define comparaciones de grupos que son genómicamente divergentes, pero que comparten similitudes culturales y ambientales. Por último, el término "*raza*" es usado para caracterizar comparaciones de grupos que divergen en cualquier aspecto (Rebbek *et al.*, 2005). Según esta clasificación, una raza presenta el grado más alto de diferenciación genética, mientras que una etnia menor presenta el más bajo.

Hay mucha discusión en el uso de este tipo de términos para la identificación de poblaciones. Muchos argumentan que las "razas" en la especie humana no existen, debido a que no hay suficiente divergencia para ser reconocidas taxonómicamente según los estándares de sistemática molecular (Keita *et al.*, 2004). El término "*raza*" responde más a cuestiones sociales que a aquéllas que reflejan diferencias genéticas. Pese a ello, se reconoce la importancia de describir a las poblaciones en estudio; no sólo para identificar grupos humanos con mayor homogeneidad genética, sino para otros estudios como los de asociación, como se explica más adelante.

## **1.2. Procesos evolutivos que actúan sobre una población**

Una población natural es aquella que es finita en número y que está estructurada según factores geográficos, de manera que la mayor parte de los apareamientos suceden dentro de esa población. Está expuesta a distintos procesos evolutivos que la hacen diferenciarse de otras y, que determinan la dinámica de su variación genética. A continuación, se explican dichos procesos evolutivos:

### a) Mutaciones

Una mutación génica es una alteración en la secuencia del ADN (sin importar si habrán o no efectos fenotípicos como consecuencia de la misma). Dicha mutación tendrá efectos evolutivos sólo si es transmitida de generación en generación. Por ejemplo, si la tasa de mutación por locus fuera de  $10^{-5}$  (una en cada 100 mil gametos), la tasa de cambio en la frecuencia de un alelo debido a una mutación sería muy baja. Parecería entonces que las mutaciones no tienen importancia evolutiva; sin embargo, hay que considerar que el número de genes funcionales en el genoma humano es muy grande (30 mil genes). Esto implica que casi cada gameto tendría 0.3 mutaciones nuevas en su genoma ( $10^{-5}$  mutaciones por gen \* 30 mil genes = 0.3 mutaciones por genoma haploide en humanos). Además, si consideramos el número de individuos que conforman una población, entonces, es claro que se producen millones de nuevas mutaciones cada generación.

Se deriva que aunque las mutaciones juegan un papel muy importante en la población, aunque no tengan efectos fenotípicos, generan variación génica dentro de las poblaciones. La posibilidad de que haya un cambio evolutivo depende del número de mutaciones diferentes que generan un fenotipo particular; algunas veces es suficiente una única mutación. Los efectos de las mutaciones van desde generar una ventaja en la adecuación del organismo, no tener consecuencias, hasta generar desventajas; por lo que se les considera adaptativas, neutrales o deletéreas, respectivamente. Además, sus efectos en la adecuación dependen del ambiente de la población y también de la constitución genética existente. Así, en presencia de cierta condición ambiental, algunas mutaciones aumentan la adecuación del fenotipo; con otra condición pueden disminuir su adecuación e, incluso, disminuir la adecuación incondicionalmente.

Es importante aclarar que una adaptación es consecuencia de una mutación espontánea al azar seguida de selección natural y no de una mutación dirigida por el ambiente. Es decir, la probabilidad de que ocurra una mutación no está influenciada por el hecho de que el organismo esté en un ambiente en donde la mutación sería ventajosa (Futuyma, 1998).

## b) Selección natural

La selección natural es un proceso en una población que altera la frecuencia génica sin que haya un cambio en el material genético (mutación) o su introducción (inmigración). Implica una diferencia consistente en la adecuación entre las entidades biológicas (individuos, especies o poblaciones) fenotípicamente diferentes. Entendemos por “adecuación”, a la tasa de incremento *per capita*; es decir, el incremento en el número de descendientes debido a que el alelo seleccionado les confiere tasas altas de supervivencia y reproducción.

Que la selección opere sólo en fenotipos que difieren en su adecuación es importante, porque un alelo puede ser reemplazado en una población por deriva génica sin que la selección natural haya operado sobre él. En este caso, no se considera que los genotipos sean diferentes en su adecuación. La diferencia de adecuaciones entre fenotipos no ocurre al azar, sino que es producto de algunas diferencias características entre ellos. Por lo tanto, la selección natural es la diferencia en la tasa de incremento entre las entidades biológicas que no ocurre al azar.

Dentro de una población, la selección natural se puede entender como una competencia entre alelos, donde el ganador es aquél que confiere alguna característica benéfica en los organismos que portan uno de ellos. De lo anterior, cabe aclarar que los genes de un individuo interactúan de manera que los genotipos no son la suma neta de los efectos de sus genes. La adecuación de un heterocigoto, por ejemplo, no puede predecirse sólo del conocimiento de las actividades de los genes individuales. Las frecuencias alélicas cambian como resultado de los efectos fenotípicos de los genotipos (Futuyma, 1998).

## c) Deriva génica

Las frecuencias alélicas de una población pueden fluctuar al azar para tener diferentes constituciones genéticas. A este proceso se le llama “*deriva génica*”. Conforme pasan las generaciones, la población se conforma por descendientes con menos copias de los genes originales. Luego, y si no operan otros procesos evolutivos, la población será monomórfica para un alelo u otro; es decir, un alelo es fijado y el otro desaparece. Consecuentemente, la deriva génica disminuye la variación genética de la población porque elimina a los homocigotos para el alelo no fijado y después los heterocigotos.

La evolución por deriva génica ocurre más rápido en poblaciones pequeñas; en cada una puede fijarse un alelo diferente. La probabilidad de que un alelo se fije es igual a su frecuencia en un tiempo dado. Así, es más probable que un alelo raro (con frecuencia cercana a cero) se pierda, a que eso pase con un alelo común. La historia de cambios de frecuencia del alelo no afecta la

probabilidad de fijación a un tiempo dado. Por otro lado, la magnitud de la deriva génica depende del tamaño efectivo de la población; conforme el tamaño de la población es menor, mayor es la deriva génica. Esto sólo aplica cuando las poblaciones mantienen su tamaño constante, no se superponen y tienen reproducción al azar. De esta forma, un alelo mutante tendrá mayor probabilidad de ser fijado si el tamaño de la población es pequeña y sucederá más rápido que con una población grande.

#### d) Flujo génico

El movimiento de individuos entre grupos puede provocar intercambio genético; a esto se le denomina "*flujo génico*". Asimismo, puede haber un influjo de alelos que han alcanzado una frecuencia alta en otras poblaciones y son llevadas por migrantes que se unen a las poblaciones locales. Este evento suele generar mayor variación génica que la provista por nuevas mutaciones, pero cuando el flujo es alto y constante, las frecuencias entre los grupos en intercambio pueden igualarse homogenizando las poblaciones. El intercambio de individuos puede darse a diferentes tasas a lo largo de la población total, de acuerdo a la probabilidad de movimiento que tengan los individuos o a las superposiciones (*overlaps*) de las poblaciones, generando gradientes de frecuencias en la población total. Cabe señalar que los inmigrantes que no se reproducen con la nueva población, no contribuyen al flujo génico.

Todos estos procesos evolutivos pueden actuar al mismo tiempo en una población natural. Las frecuencias finales que se obtienen son resultado de la suma de todos ellos. Por ejemplo, una población que está evolucionando por deriva génica puede recibir, en algún momento, flujo génico por parte de algunos inmigrantes que introduzcan nuevamente el alelo que se está perdiendo. En una población evolucionando por deriva génica puede introducirse también un alelo nuevo que les confiera mayor adecuación y entonces la evolución será por selección natural. Todos los alelos están sujetos a deriva génica, pero sólo los que confieran mayor adecuación estarán sujetos a selección natural y tendrán mayor probabilidad de ser fijados si, además, son favorecidos por la selección natural.

### **1.3. Estratificación de la población**

Una población puede estar subdividida en pequeñas unidades (subpoblaciones) debido a factores geográficos, ecológicos o de comportamiento. Sobre cada una de dichas subpoblaciones, además de sobre la población total, actúan los procesos evolutivos ya explicados. Asimismo, las subpoblaciones pueden estar interconectadas dependiendo del grado de flujo génico entre ellas. Cuando el flujo génico es alto, la variación genética entre los grupos se homogeniza; cuando es

bajo, los grupos tienden a diferenciarse genéticamente por efecto de la deriva génica, la selección natural y las mutaciones.

Entonces, una subpoblación es una subestructura de una población que presenta diferencias en variación genética entre las otras subestructuras como resultado de procesos evolutivos. Las poblaciones naturales son entidades biológicas que están en continuo cambio y, si bien es posible identificar fácilmente poblaciones que están geográfica, ecológica o socialmente aisladas, la mayoría de las veces es difícil diferenciarlas debido a su cercanía o superposición. Lo es todavía más el identificar subpoblaciones, porque la variación genética es muy baja entre ellas –algunas veces de muy pocos alelos. No obstante, es posible diferenciarlas valiéndose de algunas técnicas moleculares cuyo propósito es identificar genotipos ancestrales, o sea, grupos genotípicamente divergentes. Estos estudios han sido ampliamente usados para fines evolutivos y antropológicos. Recientemente, se han utilizado también para fines epidemiológicos (Kaessmann *et al.*, 2002; Ioannidis *et al.*, 2004).

De una población total puede divergir una subpoblación que tendrá variaciones genotípicas con respecto a la población ancestral; en otras palabras, presentará diferencias en las frecuencias de algunos alelos, las cuales, pueden ser rastreadas. En caso de que los alelos tengan importancia biológica, se pueden identificar genotipos de interés presentes en ciertas subpoblaciones cuyas frecuencias varíen entre ellas.

Los estudios de asociación nos permiten identificar grupos de susceptibilidad o de alto riesgo; por ejemplo, con una incidencia alta de alguna enfermedad como el cáncer. Si los subgrupos no están representados igualmente en el estudio, es decir, si el genotipo de riesgo está mayoritariamente expresado debido a un muestreo sesgado de individuos, la existencia de subgrupos genéticos o subestructuras en una población puede causar asociaciones falsas. Si estos grupos son clasificados por su ancestría genética, el error causado por la estratificación de la población se corrige.

Esto es, una población estratificada está compuesta de varias subpoblaciones con frecuencias del alelo de interés diferentes (una con el alelo de riesgo altamente representado y otra con frecuencias bajas). En caso de que el investigador no tome en cuenta la estratificación mencionada, podría muestrear más individuos de la subpoblación con frecuencias altas, por ejemplo, y sobre-expresar el alelo de riesgo. Su conclusión sería que el padecimiento está asociado con ese alelo. No obstante, cuando se describe la estructura genética de la población, se pueden corregir estos falsos positivos en los estudios de asociación.

Estos errores son muy comunes porque se han usado categorías establecidas socialmente como un índice de homogeneidad genética en la delimitación de poblaciones. Para disminuir las confusiones y aumentar la eficiencia de un estudio, conviene considerar los estimados de etnicidad o ancestría específicos de un genotipo dado.

#### **1.4. Marcadores genéticos**

Un marcador genético es un factor hereditario que puede usarse para caracterizar a las poblaciones según la variación de su frecuencia en distintos grupos humanos (Lisker, 1981). El marcador genético puede ser cualquier gen o fragmento génico que es identificable y caracteriza a un individuo. Son usados para la detección temprana de patologías, mapeos génicos, estimaciones de diversidad genética, investigaciones médico-legales, etc. En el estudio de poblaciones, dado que no es posible caracterizarlas por los fenotipos de sus integrantes, es necesario analizar sus genotipos para distinguir unas de otras. Para esto, son útiles los marcadores genéticos. Entre ellos se encuentran los polimorfismos genéticos y son ampliamente utilizados.

Un polimorfismo genético es la presencia de dos o más variantes alélicas en una población, independientemente de sus efectos fenotípicos. Para ser considerados como polimorfismos, las variantes alélicas raras deben tener una frecuencia mayor a 1% en la población. La presencia de un polimorfismo involucra la existencia de un alelo silvestre o ancestral y alelos mutantes o recientes; cada uno posee mutaciones que alteran el producto génico de manera que la función proteica también podría estar alterada. De esta forma, las variantes pueden incluir: 1) cambios en la secuencia del ADN, pero no en la secuencia proteica, 2) cambios en la secuencia proteica sin cambios funcionales, 3) cambios que crean proteínas con diferentes actividades, 4) cambios que crean proteínas que no son funcionales, 5) deleciones que producen pérdida de la síntesis de la proteína.

Los polimorfismos pueden ser detectados sin importar si afectan o no el fenotipo con mapas de restricción. Una secuencia génica tiene sitios de restricción (sitios de corte para enzimas de restricción) que pueden ser modificados por la presencia de mutaciones. Por ejemplo, un alelo mutante puede tener fragmentos adicionales modificando así el patrón de restricción. Las diferencias en los mapas de restricción de diferentes genotipos son llamados *polimorfismos de longitud de fragmentos de restricción (RFLP, Restriction Fragment Length Polymorphisms)* (Apéndice I). Los polimorfismos causados por diferencias en una sola base nitrogenada son denominados *polimorfismos de un solo nucleótido (SNP, Single Nucleotide Polymorphisms)*, y son fácilmente identificados por RFLP.

La combinación particular de alelos o sitios de restricción (o de cualquier otro marcador genético) presentes en una región específica del genoma se llama *haplotipo*. Dejando de lado sus efectos fenotípicos, dan idea de qué tanto difiere genéticamente una población de otra. Se considera que una variante alélica es más antigua, si tiene el menor número de mutaciones con respecto a la variante silvestre o ancestral. Consecuentemente, a más variación acumulada, mayor tiempo de divergencia.

#### 1.4.1. Marcadores Informativos de Ancestría (AIM)

“Un marcador informativo de ancestría (*AIM, Ancestral Informative Marker*) es un marcador genético que presenta diferencias en sus frecuencias alélicas entre las poblaciones parentales de  $\geq 30\%$  y, por tanto, son más informativos en el estudio de poblaciones mezcladas que otros marcadores elegidos al azar” (Bonilla *et al.*, 2004). La genotipificación de estos marcadores puede dar un estimado de los antecedentes ancestrales de un individuo o un grupo y, por ende, permite la detección de diferencias genéticas entre poblaciones. Por tanto, se les utiliza con frecuencia para identificar y corregir la estratificación de las poblaciones. Existe ya una gran cantidad de AIM para diferentes grupos poblacionales; incluyen no sólo marcadores somáticos, sino también marcadores en cromosomas sexuales y en ADN mitocondrial (ADN mt).

#### 1.4.2. ADNmt

Las mitocondrias tienen su propia molécula de ADN que es de doble cadena, pero a diferencia del ADN nuclear, es circular y tiene 16.5 kilobases (kb); constituye menos del 1% del ADN total de la célula. Cada tipo celular contiene múltiples copias de ADN en sus mitocondrias; normalmente están distribuidos en grupos separados en el orgánulo y unidas a su membrana interna. Los óvulos contienen 100 000 mitocondrias aproximadamente con una molécula de ADN cada una, mientras que los espermatozoides tienen de 50 a 75 mitocondrias. Esto significa que aún cuando en la fertilización la aportación de mitocondrias fuera tanto paterna como materna, la paterna sería muy pequeña. Se sabe que las mitocondrias están contenidas en la parte media del espermatozoide y que ésta sí entra al óvulo; de tal modo, se ha sugerido que debe existir un sistema para eliminar las mitocondrias paternas del cigoto. Por lo tanto, la herencia mitocondrial es uniparental; es decir, las mitocondrias sólo son heredadas por vía materna.

La velocidad de sustitución de nucleótidos durante la evolución es 10 veces más rápida en los genomas mitocondriales que en los nucleares; algunas de las razones para esta tasa tan alta son: 1) alta concentración de radicales libres debido a su función de fosforilación oxidativa, 2) tasa de recambio más alta que lleva a más replicaciones por unidad de tiempo, 3) mayor permanencia en forma de cadena sencilla durante su replicación, que la hace más vulnerable, 4) ausencia de

empaquetamientos con histonas y, 5) menor efectividad de los sistemas de reparación en mitocondria. El número de productos génicos es bajo, por lo que la tasa de sustitución puede ser alta sin que se altere alguno de ellos. Gracias a que la velocidad de evolución del ADNmt es alta, las comparaciones entre las secuencias son muy útiles para estimar fechas de sucesos evolutivos relativamente recientes. Además, debido a que la herencia es uniparental, el tamaño efectivo de la población es bajo, lo cual, provoca un efecto de la deriva génica alto, así como la generación de estructuras genéticas.

### **1.5. Divergencia de la población humana**

El humano (*Homo sapiens*) forma parte de la familia *Hominidae*, al igual que los grandes simios (orangutanes, *Pongo*; gorilas, *Gorilla*; y chimpancés, *Pan*). Según los registros fósiles, el primer homínido de que se tiene conocimiento es el *Australopithecus*, que evolucionó en *Homo habilis* hace 2.4 a 1.6 millones de años (ma); éste, a su vez, evolucionó en *Homo erectus* aproximadamente de 1.6 ma a 300 mil años, hasta que finalmente se extinguió para dar paso al *Homo sapiens*. Los fósiles más antiguos reconocidos como *Homo sapiens* datan de 300 mil a 400 mil años; se les denomina *Homo sapiens neanderthalensis*. Los *Homo sapiens* modernos, indistinguibles de los humanos actuales, aparecieron en África aproximadamente hace 200 mil a 170 mil años. Su distribución ocurrió al mismo tiempo que la de los *H. neanderthalensis*, hasta que los segundos reemplazaron a los primeros.

Existen dos teorías para explicar el origen de las poblaciones humanas modernas (Fig. 1.1):

- 1) Hipótesis multirregional: propone que las poblaciones de *H. sapiens* antiguas en África, Europa y Asia evolucionaron en *H. sapiens* modernos gracias al intenso flujo génico entre ellas; es decir, la evolución a *H. sapiens* ocurrió en varios lugares del Viejo Mundo.
- 2) Hipótesis de reemplazo: propone que luego de que el *H. erectus* marchó de África a Asia y Europa (primera expansión) y que evolucionó al *H. sapiens* antiguo. Una población de *H. sapiens* antigua africana evolucionó a *H. sapiens* moderno y se esparció por todo el mundo reemplazando a las poblaciones de *H. sapiens* antiguos (segunda expansión).

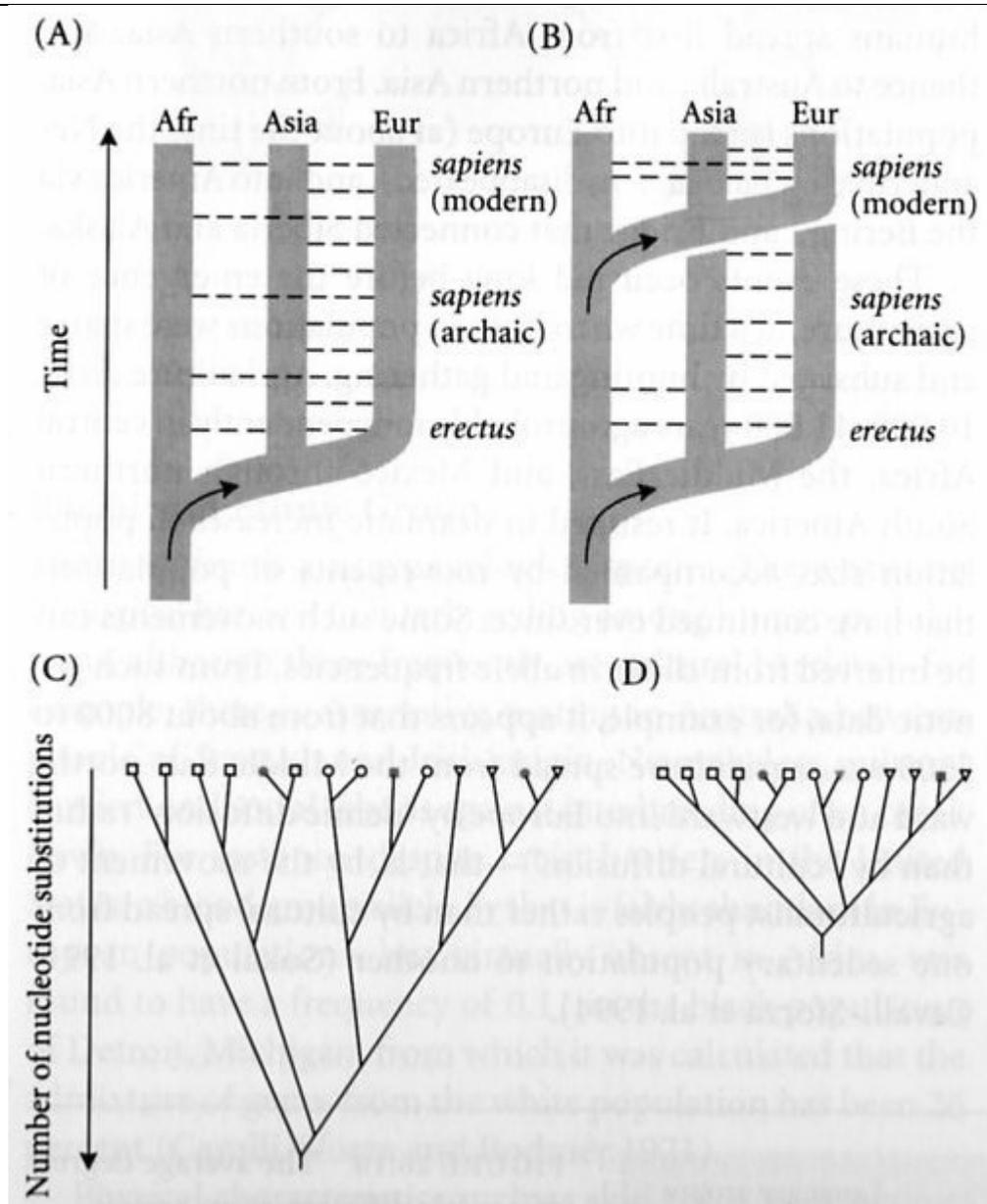


Figura 1.1. Hipótesis del origen de los humanos modernos. Fuente: Futuyma, 2001, fig. 26.8

Diversos trabajos apoyan estas hipótesis e incluso variantes de ellas; sin embargo, la mayoría de los estudios presentan filogenias muy similares con la utilización de diferentes marcadores genéticos como ADNmt, marcadores en el cromosoma Y, microsatélites y otros. Por lo que hay consenso en que la hipótesis de reemplazo explica la historia más eficientemente. A partir de ella, los estudios sugieren que los humanos modernos se extendieron de África al sur de Asia, de ahí a Australia y al norte de Asia, y luego a Europa (para entonces, el *Homo sapiens neanderthalensis* ya había desaparecido). Finalmente, los humanos modernos poblaron América a través del Estrecho de Bering hace 14 mil a 21 mil años (Fig. 1.2).

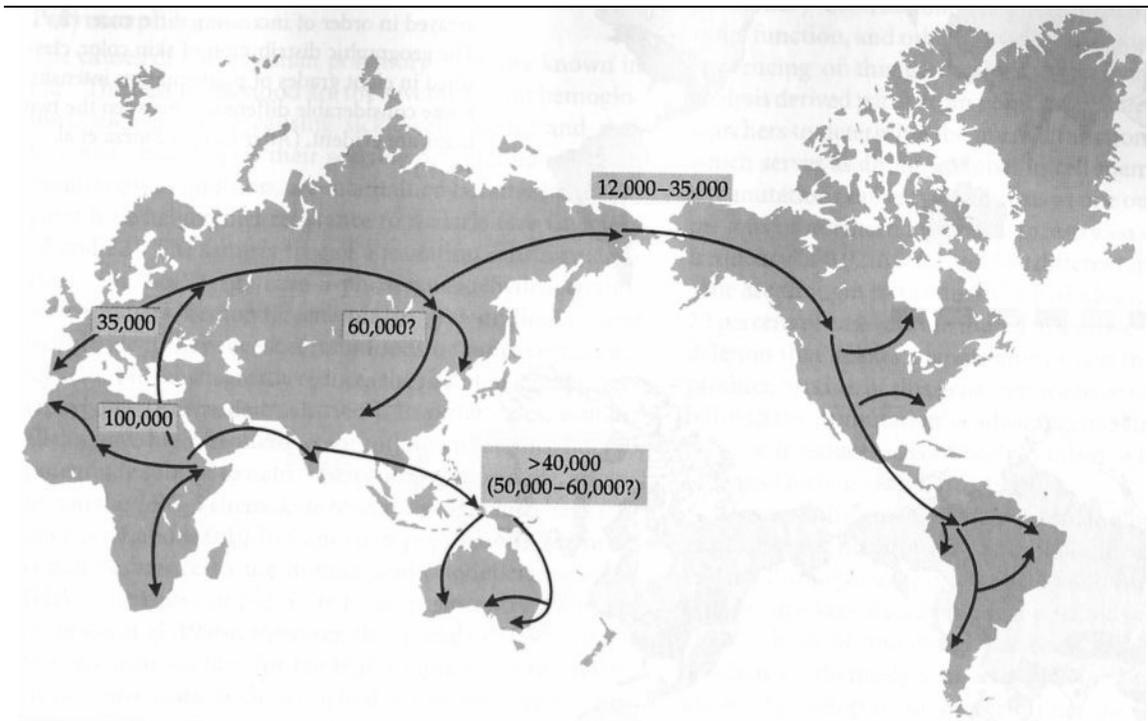


Figura 1.2. Posibles rutas de expansión de las poblaciones humanas modernas. Fuente: Futuyma, 2001, fig. 26.11.

### 1.6. Divergencia en el continente americano

Hay estudios que apoyan la idea de que las poblaciones fundadoras de nativos americanos migraron del norte de Asia a América. Aunque existen diferencias en los patrones de migraciones y las fechas, el consenso dice que sucedió a través del Estrecho de Bering hace 14 mil a 21 mil años aproximadamente. Al conjunto de fenómenos climáticos acaecidos en esos tiempos se le denomina "Era del Hielo"; incluyeron periodos de glaciación: el primero duró entre 65 mil y 35 mil años; el segundo, entre 30 mil y 13 mil años. Durante el segundo periodo, Norteamérica estaba unida a Asia por el Estrecho de Bering situado entre Siberia y Alaska, haciendo posible el paso de humanos de Asia a América. La ruta que siguieron las primeras poblaciones es aún discutida, pero en general, siguieron un curso continuo hasta Sudamérica (Fig. 1.3).

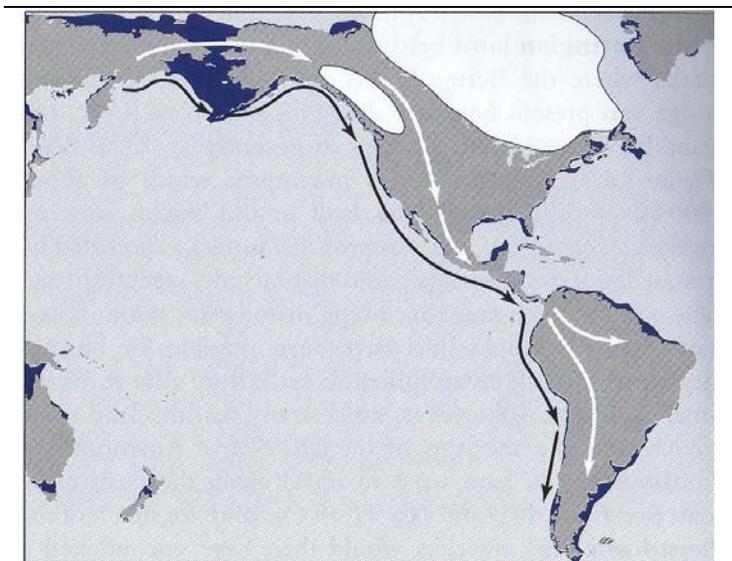
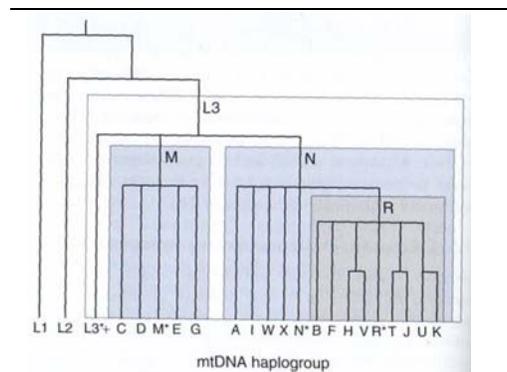
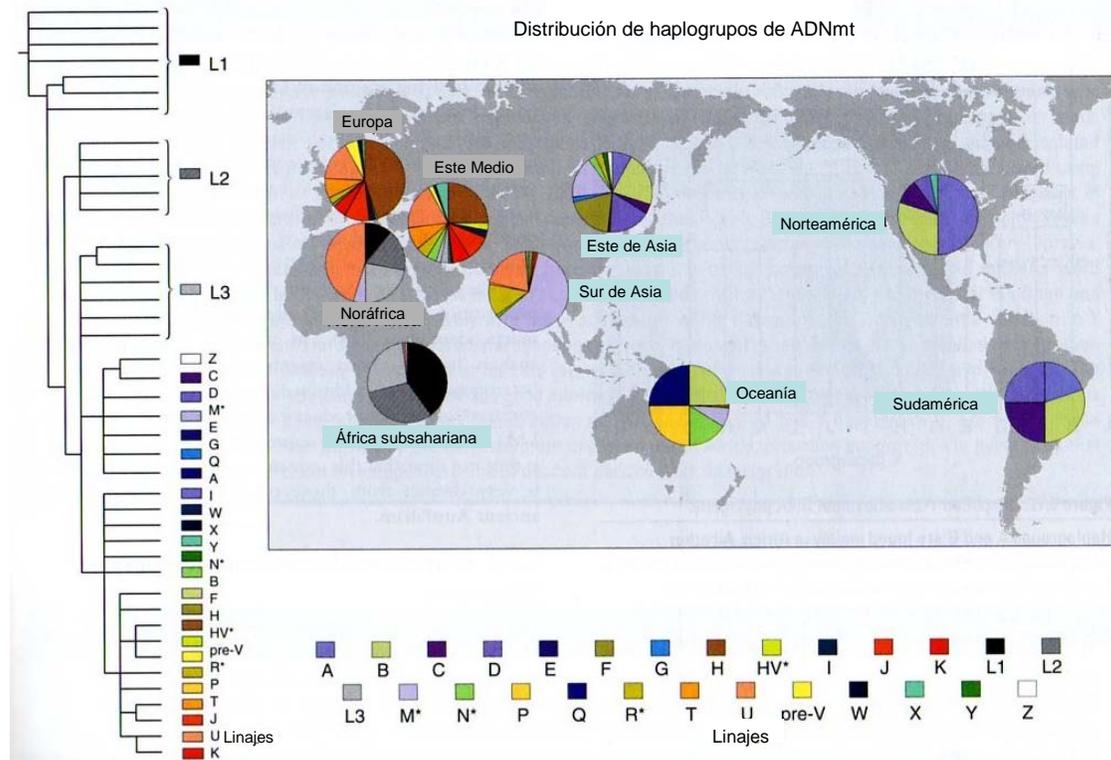


Figura 1.3. Posibles rutas de migración a América. Fuente: Jobling, 2004, fig.11.5

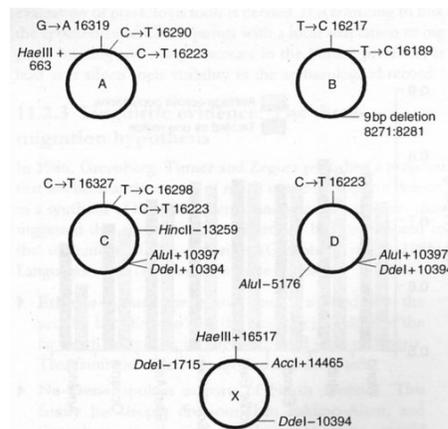
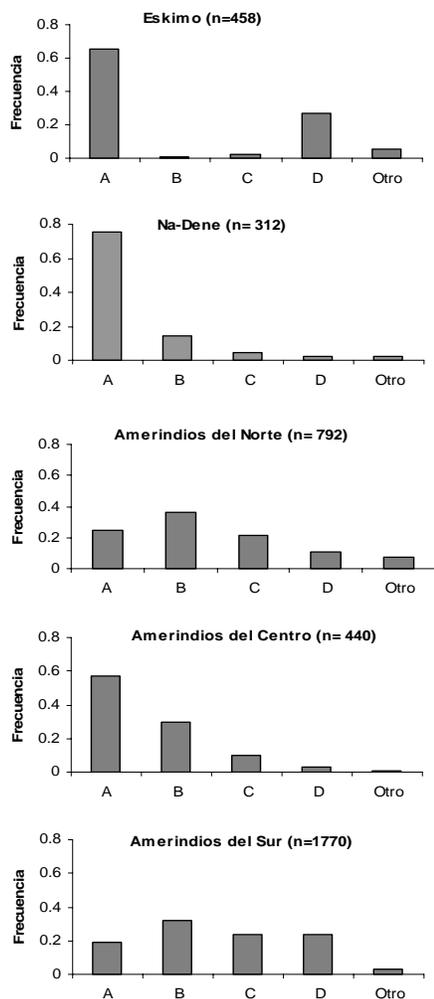
### **1.7. Haplotipos mitocondriales**

A partir del análisis de ADNmt de individuos de todo el mundo, se han encontrado haplogrupos con distribuciones geográficas bien definidas, a partir de los cuales es posible trazar filogenias que respaldan las hipótesis de migración de las poblaciones humanas antes mencionadas. La mayoría de los ADNmt de la población africana pertenecen a los haplogrupos L, mientras que los de la población no africana pertenecen a los haplogrupos M y N. El M y sus subdivisiones se encuentran principalmente en el sur y este de Asia, así como en América; el N se encuentra en Europa y el oeste de Asia. A partir de los nodos formados en las filogenias, es posible estimar fechas de divergencia (Figs. 1.4 y 1.5).



La población del continente americano se caracteriza por la existencia de cuatro haplogrupos: A, B, C y D; cada uno pertenece a diferentes poblaciones ancestrales que estuvieron aisladas por un tiempo relativamente largo antes de entremezclarse. Estas poblaciones migraron independientemente unas de otras y fueron divergiendo genéticamente en su paso. Ahora bien, aunque algunos investigadores proponen que América fue poblada a partir de varias migraciones correspondientes a cada haplogrupo, otros proponen un solo evento de migración. Sea cual sea el caso, el 95% de los ADNmt caen dentro de alguno de esos haplogrupos.

Poblaciones de ciertas regiones del continente asiático también presentan algunos de estos haplogrupos: los haplogrupos A, C y D también están presentes en Siberia y el este de Asia; el haplogrupo B se encuentra desde China central hasta el sureste de Indonesia, Polinesia y Micronesia. Dichos haplogrupos están representados diferencialmente en cada región del continente americano (Fig.1.6a). Además, se identificó un quinto haplogrupo que está presente en frecuencias muy bajas (3%) sólo en Norteamérica y además está presente en Europa, por lo que pudo haber sido introducido por otro linaje poblador. Se han identificado polimorfismos característicos de cada haplogrupo, de manera que es posible identificar a qué haplogrupo pertenece cada individuo (Fig.1.6b) (Jobling, 2004).



a)

b)

Figura 1.6. a) Frecuencias de los principales haplogrupos de ADNmt en poblaciones americanas agrupadas de acuerdo al lenguaje. b) Polimorfismos usados para identificar los 5 haplogrupos mitocondriales americanos. Fuente: Jobling, 2004, Fig. 11.11, y 11.12

### **1.8. Población mexicana**

La población mexicana se origina de una mezcla de genotipos. Hay dos eventos importantes que han influido en la estructura de la población mexicana: 1) Los nativos americanos se diversificaron a partir de los cuatro principales haplotipos fundadores analizados a partir de ADNmt y; 2) la mezcla entre nativos americanos y europeos, a partir de la Conquista en 1521. Sin embargo, a lo largo de la historia la población mexicana ha recibido flujo génico de individuos provenientes de otras regiones y, aunque no es tan alto, es posible encontrar diferencias en las frecuencias génicas a lo largo de las distintas poblaciones en el país.

Por esto, la diferenciación de las poblaciones humanas puede complicarse debido a que no tienen reproducción al azar, aún cuando tienen capacidad de reproducirse potencialmente. Los criterios de categorización de las poblaciones se organizan según los siguientes parámetros:

- a) Proximidad geográfica
- b) Lenguaje común
- c) Etnicidad compartida, cultura, religión

Ninguno de estos criterios clasifica a las poblaciones estableciendo límites estrictos entre ellas; sin embargo, para fines de algunos estudios pueden ser suficientes. No obstante, no es el caso de los estudios de estructuras poblacionales, ya que al evaluar una población podrían incluirse individuos que en realidad no forman parte de ella, ocasionando errores en el análisis.

Así pues, según el Instituto Nacional de Antropología, un mexicano mestizo es una persona que nació en el país, tiene un apellido de ascendencia española y tiene una familia de ancestros mexicanos de tres generaciones atrás. Sin embargo, varios estudios de marcadores genéticos han demostrado que la población mexicana tiene una estructura génica definida, pero que varía en cada región.

### **1.9. Marcadores nucleares**

Los marcadores genéticos nucleares que voy a utilizar en este estudio corresponden a enzimas de metabolismo de xenobióticos (cualquier compuesto químico que es ajeno a un organismo vivo). Cuando éstos son incorporados al organismo y se distribuyen en él, algunos serán almacenados y otros excretados. En cualquiera de ambos casos pueden ocurrir biotransformaciones, procesos por los cuales se incrementa la polaridad del xenobiótico para disminuir su difusibilidad y aumentar su solubilidad en agua, siendo así más fácilmente excretables.

Estas reacciones pueden desactivar o bioactivar al xenobiótico, lo que produce un metabolito más tóxico que el compuesto original y puede reaccionar con biomoléculas importantes como el ADN. Cabe mencionar que se usa la misma maquinaria bioquímica con la que se metabolizan los compuestos endógenos con estructura química similar. La mayoría de las bioactivaciones suceden durante la fase I del metabolismo de xenobióticos, aunque también puede suceder en la fase II.

Las enzimas de la fase I realizan reacciones de oxidación que introducen grupos funcionales en los xenobióticos de modo que los hagan reaccionar; en cualquier caso, estos metabolitos se convierten en sustratos de las enzimas de la fase II o son excretados. Las enzimas de fase II conjugan los metabolitos con glutatión, glucurónido o sulfato para facilitar su excreción.

#### 1.9.1. Citocromos P450

Los citocromos P450 (CYP P450) son una superfamilia de enzimas (más de 1000) presentes en los genomas de casi todos los organismos desde procariontes hasta mamíferos (Anzenbacher *et al.*, 2001). Participan en el metabolismo de fase I de xenobióticos y están involucrados en el metabolismo endógeno de esteroides. La familia CYP1A está íntimamente relacionada con la desintoxicación de xenobióticos carcinógenos. Dentro de ella, la subfamilia CYP1A1 es la principal responsable de la activación de la mayoría de los hidrocarburos aromáticos policíclicos (HAP) y las aminas aromáticas heterocíclicas (AAH) a especies más reactivas. Su proteína, la aril hidrocarburo hidroxilasa (AHH) cataliza la oxidación de los sustratos. Su gen se localiza en el cromosoma 15 y la proteína se compone de 512 aminoácidos. Se han descrito varios polimorfismos que pueden aumentar o disminuir la afinidad por el sustrato o la velocidad de reacción. Tres de ellos podrían tener relación con susceptibilidad a padecer cáncer (Fig. 1.7 y Tabla 1.1):

- 1) CYP 1A1 \*2A: es una transición T→C en posición 6235 en la región 3' no codificante. Es llamada también m1. Este cambio genera un sitio de restricción para la enzima MspI. El polimorfismo aumenta la inducibilidad del gen.
- 2) CYP 1A1 \*2C: consiste en la transición de una A→G en posición 4889 en el exón 7 cerca del grupo hemo, lo que genera un cambio de Ile→Val en el codón 462. También es llamada m2. La sustitución de esta base provoca la pérdida del sitio de restricción para la enzima BsrDI, e incrementa la actividad catalítica de la enzima.
- 3) CYP 1A1 \*4: es una transversión de una C→A en posición 4887 en el exón 7, lo que genera un cambio de Thr→Asn en el codón 461. También es llamada m4. Se utiliza la

enzima Bsal para su identificación, ya que al igual que el \*2C, el polimorfismo provoca la pérdida del sitio de restricción y también incrementa la actividad catalítica de la enzima.

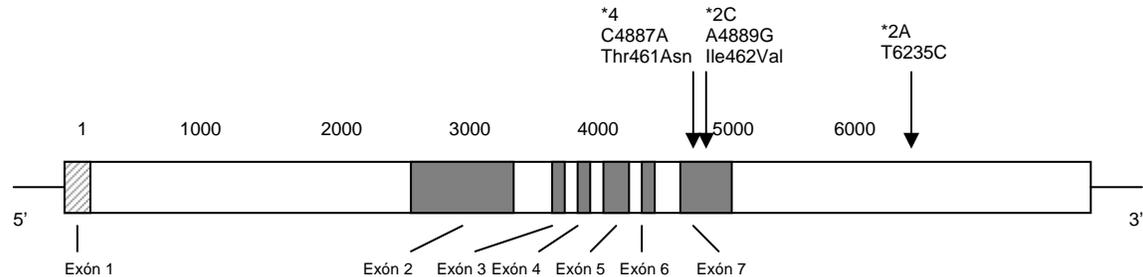


Figura 1.7. Mapa del gen CYP1A1 y localización de los 3 polimorfismos de interés. Fuente: traducida y modificada de Cascorbi, 1996, Fig. 1.

### 1.9.2. GST

Las glutatión S-transferasas (GST, *Glutathion S-Transferases*) son una superfamilia de genes de enzimas de metabolismo de fase II presentes en la mayoría de los organismos vivos. Para poder ser excretados conjugan los metabolitos producidos por las enzimas de fase I, como las CYP 1A1, con una molécula de glutatión. Se han identificado ocho familias principales de genes: alfa, kapa, mu, omega, teta, pi, sigma y zeta. Los siguientes son polimorfismos que son deleciones de los genes de las GST y, que se han asociado con el desarrollo de tumores (Tabla 1.1):

- 1) GST M: la familia *mu* consta de 5 genes situados *in tandem* en el cromosoma 1p13. Uno de los polimorfismos identificados en la isoforma GST M1 de este gen es una deleción de tamaño desconocido (GST M1\*0).
- 2) GST T: la familia *teta* consta de dos clases de GST (GSTT1 y GSTT2) y está en el cromosoma 22q11. Se ha identificado una deleción del gen GST T1 de tamaño desconocido (GST T1\*0).

Cualquiera de las dos deleciones impide la expresión de la proteína funcional y, por tanto, disminuye la capacidad de desintoxicación del organismo.

Tabla 1.1. Marcadores nucleares utilizados.

Marcador	Polimorfismo	Enzima de restricción	Respuesta molecular	Referencia
CYP1A1	*2C A4889G	BsrD I	Aumenta actividad catalítica	Hayashi <i>et al.</i> , 1991a; Cascorbi <i>et al.</i> , 1996
	*4 C4887A	Bsa I	Aumenta actividad catalítica	Cascorbi <i>et al.</i> , 1996
	*2A T6235C	Msp I	Aumenta su inducibilidad	Cascorbi <i>et al.</i> , 1996 ; Hayashi <i>et al.</i> , 1991b
GST	M1*0 Delección		Proteína no funcional	Abdel-Rahman <i>et al.</i> , 1996
	T1*0 Delección		Proteína no funcional	Pemble, <i>et al.</i> , 1994; Abdel-Rahman <i>et al.</i> , 1996

### 1.10. Marcadores mitocondriales

Cada uno de los polimorfismos que caracterizan cada haplogrupo mitocondrial es un AIM, por lo que su análisis nos da información confiable de la historia de mezcla de una población. Cada continente tiene polimorfismos comunes en su población; en otros estudios ya se han descrito sus frecuencias. Entre mayor sea la frecuencia, más confiable será la información que arrojen. A continuación, se describen los AIM utilizados en el presente trabajo (Tabla 1.2):

#### 1.10.1. AIM de Europa

El polimorfismo europeo-específico pertenece al haplotipo H; es la pérdida de un sitio de restricción para AluI en el nucleótido 7025 y tiene una frecuencia continental del 39%.

#### 1.10.2. AIM de Asia

El polimorfismo asiático-específico pertenece al haplotipo M; es la generación de un sitio de restricción para Ddel en el nucleótido 10394, con una frecuencia continental de 57%.

#### 1.10.3. AIM de África

El polimorfismo africano-específico pertenece al haplotipo L; es la generación de un sitio de restricción para HpaI en el nucleótido 3592 y tiene una frecuencia de 60 - 95% en poblaciones africanas subsaharianas y está presente en el 76% del total de los ADNmt africanos.

#### 1.10.4. AIMs de América

De los 4 haplogrupos principales descritos para los americanos nativos, los haplotipos A y B (Hap-A y Hap-B, respectivamente) son los más representados en la población americana.

- 1) Hap-A: para su identificación se busca la generación de un sitio de restricción para HaeIII en el nucleótido 663 y tiene una frecuencia promedio de 44% en poblaciones americanas. Más específicamente, tiene una frecuencia de 25%, 57% y 18% en Norteamérica, Centroamérica y Sudamérica, respectivamente (Salzano, 2002) (Fig.1.6).
- 2) Hap-B: para su identificación se busca una delección de 9 pb de los nucleótidos 8272 – 8281, con una frecuencia promedio de 22% en poblaciones americanas. Más específicamente, tiene una frecuencia de 37%, 29% y 31% en Norteamérica, Centroamérica y Sudamérica, respectivamente (Salzano, 2002) (Fig.1.6).

Tabla 1.2. Marcadores mitocondriales utilizados.

Grupo	Polimorfismo	Enzima de restricción	Frecuencia continental %	Referencia
Europa	C7025T	Alu I	39	Torrioni, <i>et al.</i> , 1995
Asia	A10394G	Dde I	57	Torrioni, <i>et al.</i> , 1995
África	C3592T	Hpa I	76	Torrioni, <i>et al.</i> , 1995
América	A663G	Hae III	44	Bonilla, <i>et al.</i> , 2006
	8271-8272	Delección 9 pb	22	Bonilla, <i>et al.</i> , 2006

#### 1.11. Estudios de mezcla y variación étnica

Para entender la estructura genética de la población mexicana se han hecho estudios utilizando diversos marcadores genéticos en pequeñas subpoblaciones y en algunos se han encontrado diferencias en ancestría al analizarse estados específicos del país (en Bonilla *et al.*, 2005, se hace una revisión de algunos trabajos). A pesar de que utilizando diferentes marcadores autosómicos, es posible encontrar diferencias de mezcla, la tendencia reportada siempre es la misma para las poblaciones mestizas: mayor ancestría nativa americana, mediana ancestría europea y muy baja ancestría africana. Sin embargo, en los estados de la costa este se incrementa la ancestría africana por haber recibido migración de individuos africanos, probablemente esclavos durante la Colonia (Tabla 1.3). Es claro que conforme los grupos indígenas están más aislados, presentan menor proporción de mezcla no indígena, de tal forma que ya se han caracterizado poblaciones que presentan niveles tan bajos de mezcla que son consideradas como poblaciones *parentales*. Tal es el caso de la población originaria de Tlapa, Guerrero, la cual, ha estado aislada

geográficamente por largo tiempo y, consecuentemente, se ha comportado endogámicamente y tenido poca mezcla (Bonilla *et al.*, 2005).

Es posible encontrar variaciones aún en el análisis de un mismo estado del país; esto podría depender de las características de muestreo. Se ha propuesto que factores como situación económica o regiones de vivienda, pueden ser determinantes en la generación de subestructuras; incluso, podrían muestrearse personas nacidas en otros estados. Los estudios suelen arrojar resultados contrastantes debido a que el grado de migración varía de región a región del país, según la clase social y ocupación de los habitantes en un determinado periodo de tiempo. Las diferencias de ancestría entre regiones pueden ser explicadas según contextos históricos que han moldeado a las poblaciones a través del tiempo.

Pese a esto, algunos autores argumentan que las poblaciones mestizas mexicanas son genéticamente homogéneas (Cerde-Flores *et al.*, 2002) e, incluso, que lo son también las poblaciones indígenas mexicanas (Buentello-Malo *et al.*, 2003). Es decir, que no presentan diferencias significativas en sus mezclas y que hay diferencias casi nulas en sus distancias genéticas.

Tabla 1.3. Proporciones de mezcla en poblaciones mestizo mexicanas determinadas con marcadores autosómicos, grupos sanguíneos y proteínas séricas.

Estado	Ciudad	Ancestría europea (%)	Ancestría nativo americana (%)	Ancestría africana del este (%)	Referencia
Yucatán	Mérida	43.1	51.0	5.9	Lisker <i>et al.</i> , 1996
Campeche	El Carmen	28.4	43.2	28.4	Lisker y Babinsky, 1986
Tabasco	Paraíso	30.9	47.4	21.7	Lisker y Babinsky, 1986
Oaxaca	Oaxaca	30.6	68.1	2.4	Lisker <i>et al.</i> , 1996
Veracruz	Veracruz	35.0	39.4	25.6	Lisker y Babinsky, 1986
Veracruz	Saladero	31.2	38.6	30.2	Lisker y Babinsky, 1986
Veracruz	Tamiahua	28.8	30.7	40.5	Lisker y Babinsky, 1986
Puebla	Puebla	33.0	56.3	10.7	Lisker <i>et al.</i> , 1988
Guerrero	Tlapa	4.2	94.5	1.3	Bonilla <i>et al.</i> , 2005
Tlaxcala	Tlaxcala	15.9	76.2	7.9	Crawford <i>et al.</i> , 1974
México	Cuanalán	20.8	61.1	18.1	Crawford <i>et al.</i> , 1976
DF	Cd. México	70.8	27.6	1.4	Tiburcio <i>et al.</i> , 1978
DF	Cd. México	40.9	56.2	2.9	Lisker <i>et al.</i> , 1986
DF	Cd. México	34.8	59.0	6.2	Lisker <i>et al.</i> , 1995
DF	Cd. México	50.1	49.0	0.9	Cerde-Flores <i>et al.</i> , 2002b
Guanajuato	León	40.4	51.1	8.5	Lisker <i>et al.</i> , 1996
Jalisco	Guadalajara	56.1	43.0	0.9	Cerde-Flores <i>et al.</i> , 2002b
Nuevo León	Monterrey	55.0	40.0	5.0	Cerde-Flores <i>et al.</i> , 2002
Coahuila	Saltillo	35.0	55.6	9.4	Lisker <i>et al.</i> , 1996

Las diferencias en los porcentajes de ancestría que se obtienen al utilizar marcadores mitocondriales y marcadores en cromosoma y se explican con base a su herencia uniparental; es decir, describen la historia femenina o masculina, respectivamente.

La distribución de los haplogrupos mitocondriales difiere tanto en el continente americano como en el país. En el caso de Centroamérica y México, la mayoría de los linajes maternos

pertenecen a los haplogrupos mitocondriales nativos americanos A y B, mientras que los haplogrupos C y D tienen muy bajas frecuencias (57%, 29%, 9.4%, 2.9%, respectivamente) (Salzano, 2002) (Tabla 1.4). Esto ocurre tanto para poblaciones indígenas como para mestizos. Sin embargo, hay variaciones en las frecuencias de los haplotipos comunes A y B en poblaciones de diferente etnicidad (Bonilla *et al.*, 2005). Por ejemplo, el Hap-A es más común en los nahuas (66.2%); por el contrario, en los mixtecos es más frecuente el Hap-B (61.1%) (Tabla. 1.4).

**Tabla 1.4. Diferencias en las frecuencias de los haplogrupos mitocondriales americanos en diferentes etnias de México.** Fuente: traducida de Bonilla, 2005, tabla 3

Distribución de haplogrupos de ADNmt en la población de Tlapa, global o subdividida por etnicidad <sup>1</sup> .									
Haplogrupo	Mixtecos	Nahuas	Tlapanecas	Mestizos DP	Mestizos DG	Etnicidad mezclada	Etnicidad desconocida	Población con MGD	Población sin MGD
A	4 (22.2)	45 (66.2)	5 (35.7)	8 (57.1)	5 (35.7)	12 (57.1)	1 (100.0)	80 (53.3)	75 (55.1)
B	11 (61.1)	14 (20.6)	5 (35.7)	7 (50.0)	7 (50.0)	8 (38.1)		50 (33.3)	43 (31.6)
C	2 (11.1)	3 (4.4)	3 (21.4)	1 (7.1)	1 (7.1)			10 (6.7)	9 (6.6)
D	1 (5.6)	4 (5.9)						5 (3.3)	5 (3.7)
AB		2 (2.9)						2 (1.3)	2 (1.5)
Otra			1 (7.1)		1 (7.1)	1 (4.8)		3 (2.0)	2 (1.5)
Desconocida <sup>2</sup>		2		3	1			6	5
Total	18	70	14	17	15	21	1	156	141

<sup>1</sup> En paréntesis, porcentaje basado en el número total de haplogrupos.

<sup>2</sup> Imposible de determinar el haplogrupo debido a problemas técnicos. No incluidos en el cálculo de los porcentajes. MGD, mestizos genéticamente definidos.

DP= definidos personalmente; DG= definidos genéticamente.

En cuanto a los marcadores nucleares, no sólo se han observado variaciones en las frecuencias registradas entre países, sino también en diferentes regiones de un mismo país (Kawajiri *et al.*, 1999). Para el caso de México, los datos son escasos, pero se ha reportado que variaciones en las frecuencias de diferentes estados no son significativas como para considerarlas subpoblaciones (Pérez, 2006) (Tabla 1.5).

**Tab3a 1.5. Variaciones en las frecuencias alélicas de los polimorfismos del CYP1A1 y GST en México.** Fuente: Pérez, 2006, tabla 3.2

Polimorfismo	M1*0	T1*0	Ile	Val	Thr	Asn	No. individuos
Chiapas	0.25	0.25	0.31	0.69	0.75	0.25	8
DF	0.33	0.15	0.39	0.61	0.94	0.06	317
Edo. Mex	0.46	0.12	0.42	0.58	0.95	0.05	32
Guanajuato	0.11	0	0.5	0.5	0.94	0.06	9
Guerrero	0.28	0.07	0.43	0.57	0.89	0.11	14
Hidalgo	0.53	0.13	0.57	0.43	0.97	0.03	15
Michoacán	0.25	0.37	0.25	0.75	1	0	8
Morelos	0.38	0.07	0.54	0.46	1	0	13
Oaxaca	0.43	0.06	0.38	0.62	1	0	16
Puebla	0.5	0.05	0.5	0.5	0.95	0.05	20
Tlaxcala	0.14	0.28	0.36	0.64	1	0	7
Veracruz	0.3	0.04	0.41	0.59	0.96	0.04	23
Suma	3.96	1.59	5.06	6.94	11.35	0.65	462
Promedio	0.33348	0.12739	0.42167	0.5783	0.94583	0.05417	40.1666667
Desv STD	0.1345	0.1345	0.09389	0.0939	0.07012	0.07012	87.4828122
Varianza	0.01809	0.01809	0.00882	0.0088	0.00492	0.00492	

\* En el análisis sólo se incluyeron los estados en los que teníamos al menos 7 individuos.

## JUSTIFICACIÓN

Se sabe que la población mexicana presenta una mezcla genética proveniente de muchos grupos étnicos; si bien los estudios son escasos, ya se ha comprobado variación interétnica entre habitantes de diferentes estados del país; y es claro que nuestra estructura genética es muy diferente a la de otros países. Por lo que analizar las proporciones de contribución de cada uno en poblaciones mestizas es importante para conocer su grado de heterogeneidad, además de que nos permitiría hacer inferencias de susceptibilidad si se comprobara asociación de genotipos de riesgo y haplotipos de ancestría específicos.

Algunos polimorfismos más representados en poblaciones de ciertas regiones geográficas han sido relacionados con ciertos padecimientos; no obstante, es necesario saber que la población estudiada es lo suficientemente homogénea como para reducir errores en los cálculos de sus frecuencias alélicas y sus asociaciones con los padecimientos. Por tanto, es necesario identificar y corregir la estratificación de las poblaciones en estudio.

El desarrollo de técnicas moleculares para la identificación de genotipos ancestrales ha sido crucial en la diferenciación de las poblaciones; la inclusión de los estimados de etnicidad o ancestría específicos de un genotipo dado que arrojan, ha permitido disminuir las confusiones y aumentar la eficiencia de los estudios de asociación. Existe ya una gran cantidad de marcadores genéticos indicadores de ancestría para diferentes grupos poblacionales; incluyen no sólo marcadores somáticos, sino también marcadores en cromosomas sexuales y en ADNmt (ADN mitocondrial).

Este trabajo pretende contribuir al conocimiento de la estructura genética de una población mestiza mexicana así como a la posible asociación entre marcadores de ancestría y marcadores de riesgo a desarrollar padecimientos como el cáncer.

## OBJETIVOS

- Calcular las frecuencias alélicas y genotípicas de cinco marcadores nucleares: CYP1A1 \*2C, \*4, \*2A (cromosoma 15) y GST M1\*0 (cromosoma 1p13) y T1\*0 (cromosoma 22q11), así como de cinco marcadores mitocondriales informativos de ancestría amerindia, asiática, europea y africana, en una población mestiza mexicana conformada por habitantes del Distrito Federal y Área Metropolitana. Al mismo tiempo, determinar si existe asociación entre alguno de dichos marcadores nucleares y los mitocondriales.
- Estimar las contribuciones genéticas provenientes de cada continente, con base en las frecuencias obtenidas de los marcadores mitocondriales en la población estudiada.
- Averiguar si las distancias geográficas que separan las diferentes entidades federativas del país influyen en la diferenciación genética de sus habitantes.

## HIPÓTESIS

- Dada la historia de la población americana, la población mestiza mexicana tendrá una alta contribución de ancestría amerindia y asiática, según la herencia materna del ADNmt.
- Dadas la variación interétnica tan marcada en las frecuencias alélicas de los polimorfismos nucleares en diferentes países, es muy probable que en nuestra población se encuentre asociación entre alguno de los locus nucleares y los mitocondriales.
- Dado que las contribuciones genéticas varían a lo largo de los estados del país, es probable que los más alejados entre sí, tengan habitantes con diferencias genéticas más acentuadas.

## **2. MATERIAL Y MÉTODOS**

### **2.1. Muestreo de individuos**

Se reunieron 148 individuos habitantes del Distrito Federal (DF) y su Área Metropolitana. A cada uno de ellos se les tomó una muestra de 5 ml de sangre y se les pidió indicaran su lugar de nacimiento, así como el de sus padres y abuelos.

### **2.2. Purificación de ADN**

El ADN se purificó de sangre entera mediante el método de sales y perclorato de sodio (Daly *et al.*, 1996), y se utilizó para la genotipificación de los 10 marcadores genéticos mencionados en la introducción a este documento (Apéndice I).

### **2.3. Genotipificación**

El conjunto de marcadores genéticos de ancestría, así como los marcadores nucleares elegidos para el análisis de esta población, fueron tomados de estudios anteriores (Tablas 1.1 y 1.2). Se tipificaron aquellos marcadores genéticos de ancestría que tuvieran grandes diferencias de frecuencia entre poblaciones.

Los polimorfismos del gen CYP1A1 que codifica para la enzima AHH y 4 marcadores mitocondriales se caracterizaron usando la técnica de RFLP; la delección de 9 pb (haplotipo mitocondrial B) por PCR; mientras que las de GST mediante un PCR múltiple (Apéndice I).

Para la amplificación de los templados, se utilizó un termociclador *Rapidcycler* de *Idaho Technology*; las condiciones de PCR se especifican después para cada marcador. Los análisis de restricción se realizaron siguiendo las indicaciones de los proveedores; información más detallada para cada marcador es descrita a continuación. Para la identificación de los polimorfismos, se utilizaron geles convencionales de agarosa o acrilamida, los cuales, fueron revelados por fluorescencia con bromuro de etidio y se utilizó el documentador de fotografías *Kodak Gel Logic 200 Imaging System* para conservar los registros.

### 2.3.1. Marcadores nucleares

#### 2.3.1.1. CYP 1A1

Para determinar los polimorfismos del CYP1A1 \*2C y CYP1A1 \*4, se realizó una amplificación con los oligonucleótidos M2F y M2R (Cascorbi *et al.*, 1996). Ya que el mismo amplicón de 204 pb obtenido contiene ambos polimorfismos, se le usó para realizar los RFLP en digestiones individuales. Para determinar el primer polimorfismo, se utilizó la enzima de restricción BsrDI y se obtuvieron los siguientes resultados: 1) para el homólogo silvestre, dos fragmentos de 149 pb y 55 pb; 2) para el heterólogo, tres fragmentos de 204 pb, 149 pb y 55 pb; 3) para el homólogo mutante, sólo un fragmento de 204 pb. Para determinar el segundo polimorfismo, se utilizó la enzima de restricción BsaI, obteniéndose: 1) para el homólogo silvestre, dos fragmentos de 139 pb y 65 pb; 2) para el heterólogo, tres fragmentos de 204 pb, 139 pb y 65 pb; 3) para el homólogo mutante, únicamente un fragmento de 204 pb (Tabla 2.1).

**TABLA 2.1. a) Condiciones de PCR y restricción para la tipificación de CYP1A1 \*2C y \*4. b) Patrón de bandeo obtenido en los RFLPs.** El mismo patrón de bandeo aplica para ambos; el polimorfismo provoca la pérdida del sitio de restricción. La primera cifra de peso de la banda corresponde al \*2C; la segunda cifra corresponde al \*4.

CYP1A1	*2C	*4
a) RIV1 5' CTG TCC CTC TGG TTA CAG G 3'	BsrDI	BsaI
RIV2 5' TTC CAC CCG TTG CAG CAG GAT 3'		
94°C, 2', 1 ciclo	Incubar a 65°C, 16 h;	Incubar a 50°C, 2 h;
94°C, 15"; 63°C, 15"; 72°C, 30", 40 ciclos	desactivar a 85°C, 20'	desactivar a 65°C, 20'
72°C, 2', 1 ciclo		
Pendiente 9.9		

b)

**CYP 1A1 \*4 y \*2C**

204 pb  
149, 139 pb  
55, 65 pb

100 pb

(-/-)      (+/-)      (+/+)

Para determinar el polimorfismo CYP1A1 \*2A, se utilizaron los oligonucleótidos P80 y C47 (Cascorbi *et al.*, 1996; Hayashi *et al.*, 1991b, respectivamente). Para realizar el análisis de restricción, se utilizó la enzima MspI. Aclarando que el amplicón es de 343pb, se obtuvieron: 1) Para el homólogo silvestre, únicamente un fragmento de 343pb; 2) para el heterólogo, tres fragmentos de 343pb, 204pb y 140pb; y 3) para el homólogo mutante, dos fragmentos de 204pb y 140pb (Tabla 2.2).

**TABLA 2.2. a) Condiciones de PCR y restricción para el CYP 1A1 \*2A. b) Patrón de bandeo obtenido en el RFLP. El polimorfismo genera el sitio de restricción.**

CYP1A1		*2A
a) P80 5' TAG GAG TCT TGT CTC ATG CCT 3'		MspI
C47 5' CAG TGA AGA GGT GTA GCC GCT 3'		
94°C, 2', 1 ciclo		Incubar a
94°C, 30"; 63°C, 30"; 72°C, 30", 35 ciclos		37°C, 17 h;
72°C, 5', 1 ciclo		desactivar a
Pendiente 9.9		65°C, 25'

**b)**

### 2.3.1.2. GSTs

Las isoformas GST M1 y GST T1 fueron genotipificadas en una misma reacción de amplificación (PCR multiplex) utilizando oligonucleótidos para la GST M1 y la GST T1, además de oligonucleótidos para el CYP1A1 (Abdel-Rahman *et al.*, 1996). Estos últimos sirvieron como control de amplificación. Se obtuvieron las siguientes bandas: 1) Para la GST T1, una banda de 480 pb; 2) para el CYP1A1, una banda de 312 pb; 3) para la GST M1, una banda de 215 pb. La ausencia de cualquiera de las bandas para las GST indica la presencia de una delección (Tabla 2.3).

**TABLA 2.3. a) Condiciones de PCR para GST. b) Patrón de bandeo obtenidas en el gel. La banda superior corresponde a GST M1; la banda inferior corresponde a GST T1.**

GST multiplex		b)
a) GST1 5' TTC CTT ACT GGT CCT CCA TCT 3'		
GST2 5' TCA CCG GAT CAT GGC CAG CA 3'		
GSM1 5' GAA CTC CCT GAA AAG CTA AAG C 3'		
GSM2 5' GTT GGG CTC AAA TAT ACG GTG G 3'		
CYP1 5' GAA CTG CCA CTT CAG CTG TCT 3'		
CYP2 5' CAG CTG CAT TTG GAA GTG CTC 3'		
94°C, 2', 1 ciclo		<p style="text-align: center;"><b>M+0      T+/M+      T+0</b></p>
94°C, 30"; 59°C, 30"; 72°C, 30", 40 ciclos		
72°C, 2', 1 ciclo		
Pendiente 9.9		

Los productos de PCR y digestiones para todos los marcadores nucleares fueron corridos en geles de agarosa al 3% y revelados por fluorescencia con bromuro de etidio.

## 2.3.2. Marcadores genéticos mitocondriales

### 2.3.2.1. Marcador mitocondrial europeo

Para el polimorfismo europeo, se amplificó un fragmento de 551 pb utilizando los oligonucleótidos EURfw y EURrv (Torrioni *et al.*, 1995); luego se le digirió con la enzima de restricción AluI. Debido a que el genoma mitocondrial es de cadena sencilla, sólo se obtienen dos patrones: 1) para los portadores del polimorfismo se observan dos bandas de 158 y 30 pb; 2) para los no portadores, una banda de 188 pb. Además, se observan 3 bandas comunes de 149, 139 y 75 pb (Tabla 2.4).

Las digestiones fueron corridas en geles de agarosa al 5% y reveladas por fluorescencia con bromuro de etidio.

**TABLA 2.4. a) Condiciones de PCR y restricción para el marcador europeo. b) Patrón de bandeo.** (-) Corresponde a la ausencia del polimorfismo y patrón europeo; (+) corresponde a la presencia del polimorfismo y patrón no europeo. El polimorfismo pierde el sitio de restricción.

<b>EUR 7025</b>	
<b>a)</b> Fw 5' AGA CCC CAT TCT ATA CCA AC 3' Rv 5' CGT AGG TTT GGT CTA GG 3'	AluI  Incubar a 37°C, 17 h; desactivar a 65°C, 10'
94°C, 2', 1 ciclo	
94°C, 15"; 52°C, 15"; 72°C, 30", 40 ciclos	
72°C, 2', 1 ciclo	
Pendiente 9.9	

**EUR 7025 AluI**

188 pb  
158 pb  
149 pb  
139 pb  
75 pb  
30 pb

100 pb

(-) (+)

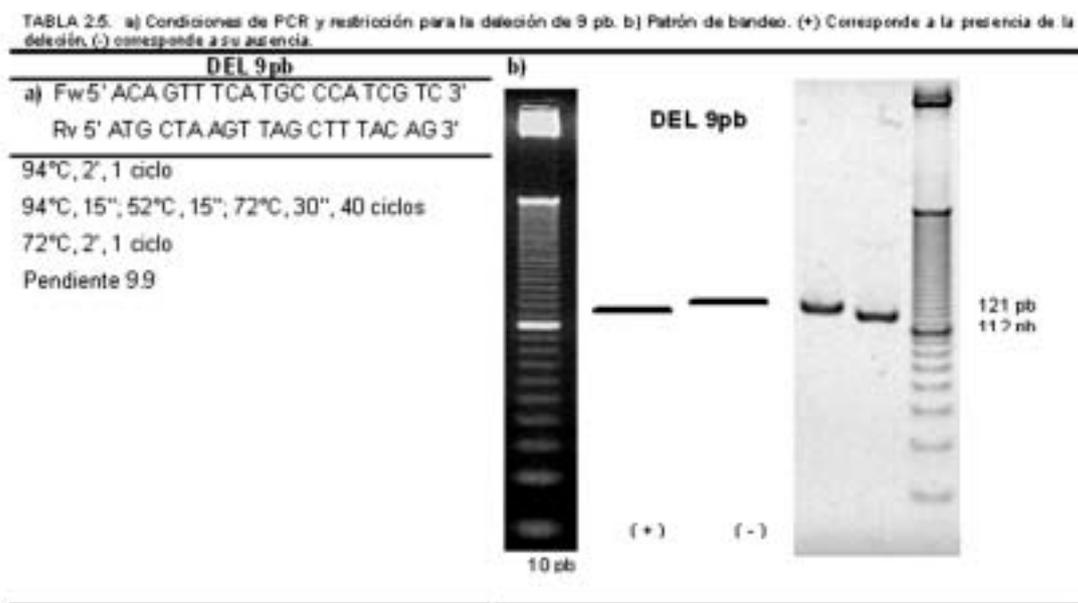
**b)**

### 2.3.2.2. Marcadores mitocondriales amerindios

Para evaluar la contribución genética amerindia, se utilizaron dos marcadores genéticos: 1) Una delección de 9 pb compartida por asiáticos y amerindios; 2) un polimorfismo específico del continente americano (AME). La presencia de la delección de 9 pb, fue determinada por medio de una amplificación utilizando los oligonucleótidos DELfw y DELrv (Bonilla *et al.*, 2005). Si no está

presente, se obtiene una banda de 121 pb; en caso contrario, se obtiene una banda de 112 pb (Tabla 2.5).

Los productos de PCR fueron analizados mediante geles de acrilamida al 10% y revelados por fluorescencia con bromuro de etidio.



Para el polimorfismo amerindio, se amplificó un fragmento de 80 pb con los oligonucleótidos AMEfw y AMErv (Bonilla *et al.*, 2005); se le digirió con la enzima HaeIII. Cuando el polimorfismo está presente, se obtienen dos bandas de 44 y 36 pb; en caso contrario, se obtiene una sola banda de 80 pb (Tabla 2.6).

Las digestiones fueron analizadas mediante electroforesis en geles de acrilamida al 10% y reveladas por fluorescencia con bromuro de etidio.

### 2.3.2.3. Marcadores mitocondriales asiáticos

Para el polimorfismo asiático se amplificó un fragmento de 658 pb con los oligonucleótidos ASIfw y ASIrV (Torrioni *et al.*, 1995); se le digirió con la enzima Ddel. Cuando el polimorfismo está presente se obtienen 4 bandas de 315, 175, 130 y 38 pb; en caso contrario, se obtienen bandas de 315, 213 y 130 pb (Tabla 2.7).

Las digestiones fueron analizadas mediante electroforesis en geles de agarosa al 4% seguido de fluorescencia con bromuro de etidio.

TABLA 2.6: a) Condiciones de PCR y restricción para el polimorfismo específico anerinda. b) Patrón de bandas. (+) Corresponde a la presencia del polimorfismo; (-) corresponde a su ausencia. El polimorfismo genera el sitio de restricción.

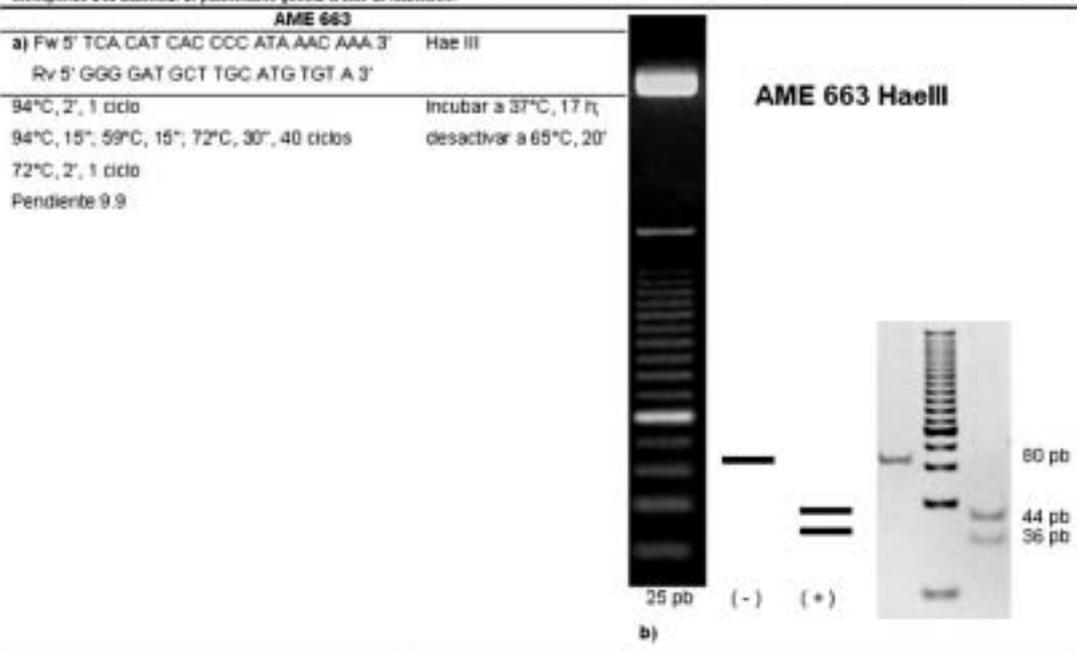
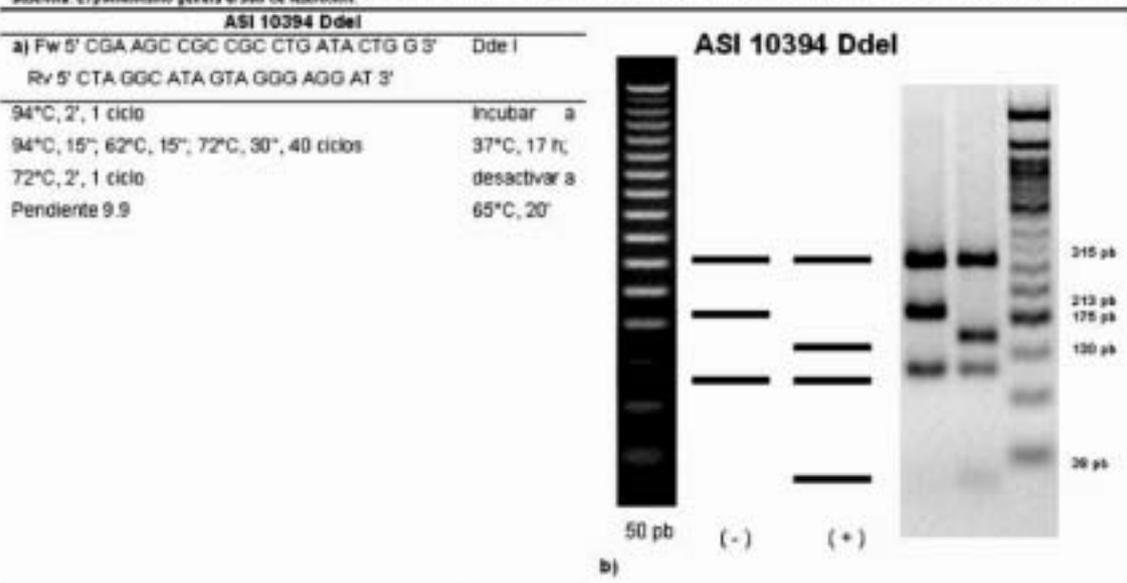


TABLA 2.7: a) Condiciones de PCR y restricción para el polimorfismo específico asálico. b) Patrón de bandas. (+) Corresponde a la presencia del polimorfismo; (-) corresponde a su ausencia. El polimorfismo genera el sitio de restricción.



### 2.3.2.4. Marcador mitocondrial africano

Para el polimorfismo africano, se amplificó un fragmento de 609 pb con los oligonucleótidos AFRfw y AFRrv (Torróni *et al.*, 1995); se le digirió con la enzima HpaI. Si el polimorfismo está presente se obtienen 2 bandas de 486 y 123 pb; por el contrario, si está ausente se obtiene una banda de 609 pb (Tabla 2.8).

Todas las digestiones se analizaron mediante electroforesis en geles de agarosa al 3%, seguido de fluorescencia con bromuro de etidio.

**TABLA 2.8. a) Condiciones de PCR y restricción para el polimorfismo africano. b) Patrón de bandeos.** (+) Corresponde a la presencia del polimorfismo; (-) corresponde a su ausencia. El polimorfismo genera el sitio de restricción.

<b>AFR 3592 HpaI</b>	
a) Fw 5' TTC AAA TTC CTC CCT GTA CG 3'	Hpa I
Rv 5' GGC TAC TGC TCG CAG TG 3'	
94°C, 2', 1 ciclo	Incubar a 37°C, 17 h;
94°C, 15"; 59°C, 15"; 72°C, 30", 40 ciclos	No se requiere
72°C, 2', 1 ciclo	desactivar
Pendiente 9.9	



100 pb  
b)

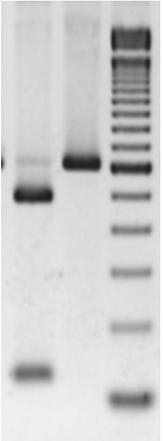
**AFR -3592 HpaI**

(-)



(+)





609 pb  
486 pb  
123 pb

### **3. RESULTADOS**

#### **3.1. Distribución geográfica de la muestra**

De los de 148 habitantes del DF y su Área Metropolitana, los cuales donaron una muestra sanguínea, 76 fueron mujeres y 72 hombres; sus edades fluctúan entre los 17 y 52 años de edad.

De los resultados arrojados por el análisis destaca el hecho de que, no obstante, la población estudiada radica en la zona centro del país, su historia genética no se limita a esta zona geográfica, sino que se amplía a muchos otros estados. Esto es observable si se rastrea el lugar de procedencia de las madres y las abuelas maternas de cada individuo. La ampliación aumenta conforme se retrocede en el estudio de generaciones anteriores hasta cubrir casi la totalidad del territorio (Fig. 3.1 y Tabla 3.1); aunque el DF y el Estado de México continúan siendo los más representados.

Algunos de los donadores no proveyeron datos completos sobre el lugar de nacimiento de sus abuelas; sin embargo, ésta no es una razón que debilite la afirmación hecha en el párrafo anterior.

#### **3.2. Marcador AFR2**

Durante la genotipificación del marcador mitocondrial europeo, se obtuvo otro patrón de bandeo no esperado en 3 de los individuos. Dicho patrón corresponde a un polimorfismo de ancestría africana; es la pérdida de un sitio de restricción para la enzima AluI en el nucleótido 7055 (Torrioni *et al.*, 1995) y tiene una frecuencia de 20.8% en ADNmt senegalense. Para distinguirlo del marcador africano AFR, se le llamó AFR2 (Tabla 3.2).

TABLA 3.1. Distribución geográfica de la muestra. Los estados están ordenados conforme aumenta la distancia con respecto al DF.

Donador			Madre			Padre			Abuela materna		
Estado	No.	%	Estado	No.	%	Estado	No.	%	Estado	No.	%
DF	124	83.78	DF	80	54.05	DF	76	51.35	DF	42	28.38
Edo Méx	14	9.46	Edo Méx	13	8.78	Edo Méx	12	8.11	Edo Méx	16	10.81
Puebla	4	2.70	Morelos	2	1.35	Tlaxcala	1	0.68	Morelos	1	0.68
Qro	1	0.68	Tlaxcala	1	0.68	Hidalgo	5	3.38	Tlaxcala	1	0.68
Jalisco	3	2.03	Hidalgo	6	4.05	Puebla	7	4.73	Hidalgo	7	4.73
Oaxaca	1	0.68	Puebla	8	5.41	Querétaro	2	1.35	Puebla	8	5.41
Guerrero	1	0.68	Jalisco	1	0.68	Sn L Potosí	1	0.68	Querétaro	1	0.68
			Guanajuato	8	5.41	Jalisco	3	2.03	Jalisco	2	1.35
			Michoacán	6	4.05	Guanajuato	9	6.08	Guanajuato	15	10.14
			Oaxaca	6	4.05	Michoacán	6	4.05	Michoacán	9	6.08
			Guerrero	2	1.35	Oaxaca	7	4.73	Oaxaca	7	4.73
			Veracruz	5	3.38	Guerrero	5	3.38	Guerrero	3	2.03
			Tabasco	2	1.35	Veracruz	5	3.38	Veracruz	6	4.05
			Chiapas	2	1.35	Chiapas	1	0.68	Tabasco	2	1.35
			Zacatecas	1	0.68	Coahuila	2	1.35	Chiapas	1	0.68
			Sinaloa	1	0.68	Sinaloa	1	0.68	Yucatán	1	0.68
			Durango	1	0.68	Durango	1	0.68	Zacatecas	1	0.68
			B.Cal. Sur	1	0.68	Chihuahua	1	0.68	Coahuila	1	0.68
						Guatemala	1	0.68	Sinaloa	1	0.68
									Durango	1	0.68
									Alemania	1	0.68
Sin dato				2	1.35		2	1.35		21	14.19
Total	148	100		148	100		148	100		148	100

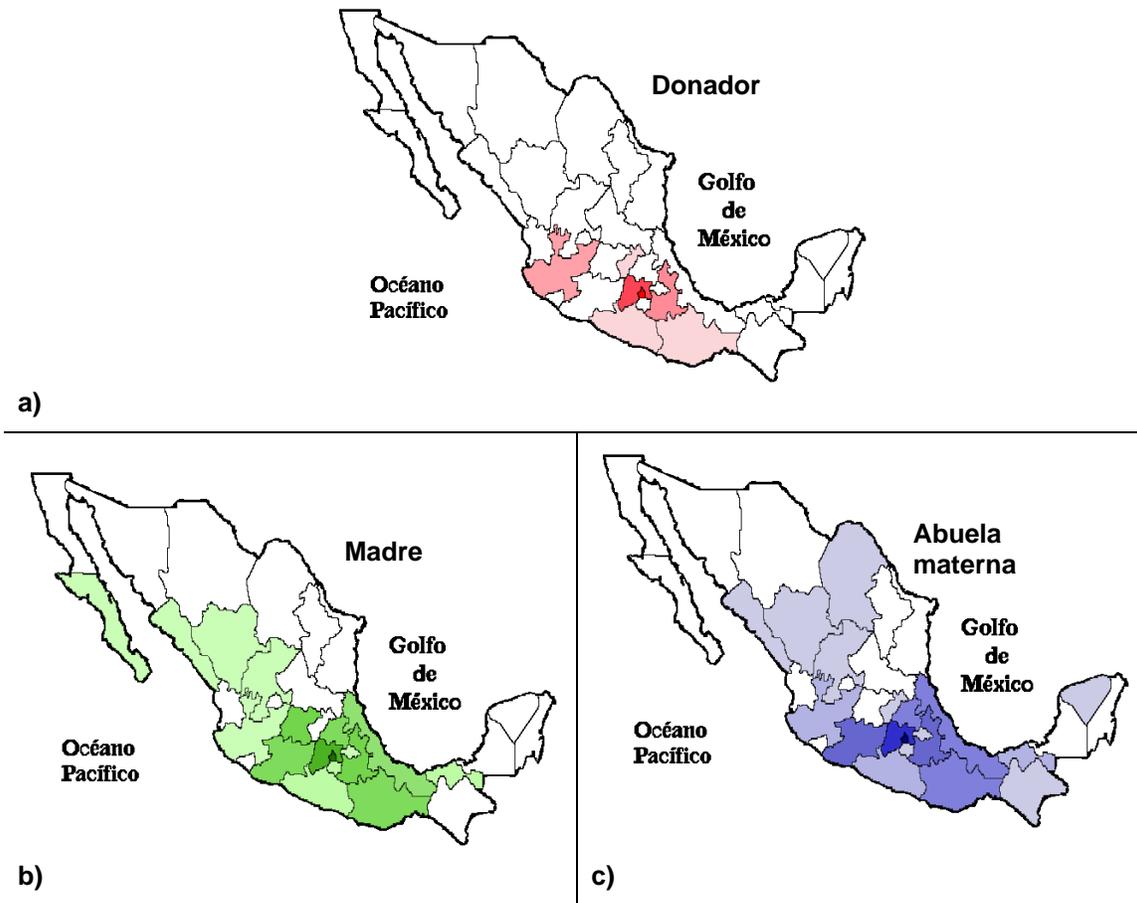


Figura 3.1. Mapas de distribución según la procedencia de: a) donadores, b) madres de los donadores, c) abuelas maternas de los donadores. La intensidad del color disminuye conforme se reduce el número de individuos nacidos en cada estado; es decir, el estado con tono más fuerte es el más representado.

**TABLA 3.2. a) Condiciones de PCR y restricción para el marcador africano AFR2. b) Patrón de bandeo.** Se muestra el patrón europeo (E), no europeo (NE), y el africano nuevo (AF). El polimorfismo pierde el sitio de restricción.

<b>AFR 7055</b>		<b>b)</b>
<p>a) Fw 5' AGA CCC CAT TCT ATA CCA AC 3'</p> <p>Rv 5' CGT AGG TTT GGT CTA GG 3'</p>	<p>AluI</p>	<p style="text-align: center;"><b>AFR 7055 AluI</b></p> <p style="text-align: center;">(E) (NE) (AF)</p>
94°C, 2', 1 ciclo	Incubar a 37°C, 17 h;	
94°C, 15"; 52°C, 15"; 72°C, 30", 40 ciclos	desactivar a 65°C, 10'	
72°C, 2', 1 ciclo		
Pendiente 9.9		

### 3.3. Frecuencias genotípicas y alélicas de los marcadores

El ADN se purificó de las 148 muestras y se tipificaron los 10 marcadores arriba mencionados. Fueron calculadas las frecuencias genotípicas y alélicas de los marcadores utilizando las fórmulas de la siguiente tabla:

**Tabla 3.3. Fórmulas para calcular las frecuencias genotípicas y alélicas observadas.**

Frecuencias genotípicas	Frecuencias alélicas
AA = AA/N	$p = (2AA + Aa)/2N$
Aa = Aa/N	$q = (Aa + 2aa)/2N$
aa = aa/N	
donde,	
N= no. total de la muestra	
AA= homócigo silvestre; Aa= heterócigo; aa= homócigo mutante	
p= alelo silvestre; q= alelo mutante	

Los polimorfismos nucleares con frecuencias más altas fueron el GST M1\*0, CYP 1A1\*2A y CYP 1A1\*2C; el polimorfismo GST T1\*0 presentó una frecuencia baja; mientras que el polimorfismo CYP 1A1\*4 mostró una casi nula. De los polimorfismos mitocondriales, el que presentó la frecuencia más alta fue el AME, correspondiente al haplotipo amerindio. Los polimorfismos DEL y ASI también presentaron frecuencias altas, el primero de ancestría asiática-amerindia y el segundo sólo asiática (Tabla 3.4 y Fig. 3.3).

Debido a que la técnica de detección de los polimorfismos de la GST no permite distinguir heterocigotos de homocigotos para cualquiera de las dos deleciones, sólo es posible calcular las frecuencias genotípicas de los homocigotos para cada deleción (M1\*0 ó T1\*0), así como la suma de las frecuencias de los individuos homocigotos silvestres y de los heterocigotos (M1/M1+M1/M1\*0 ó T1/T1+T1/T1\*0). Las frecuencias alélicas observadas no se pueden calcular porque solamente es posible distinguir entre individuos que portan o no la deleción; es decir, no sabemos cuántos de los que la portan son heterocigotos y cuántos son homocigotos mutantes. Consecuentemente, las frecuencias de las deleciones podrían subestimarse. Aún así, se calcularon las frecuencias alélicas esperadas para GST M1 y T1, con base en el número de homocigotos para la deleción encontrados para cada caso, asumiendo que la población se encuentra en equilibrio de Hardy-Weinberg. Por otra parte, dada la condición haploide del genoma mitocondrial, la frecuencia alélica de sus marcadores es la misma que la frecuencia genotípica.

**Tabla 3.4. Frecuencias de los marcadores genéticos.** a) Polimorfismos nucleares; b) Polimorfismos mitocondriales; +, presencia del polimorfismo, -, ausencia del polimorfismo. \* Frecuencia alélica esperada.

MARCADORES GENÉTICOS	GENOTIPO	No. INDIVIDUOS	f GENOTÍPICA OBSERVADA	f ALÉLICA	
<b>a) Polimorfismos nucleares</b>					
CYP 1A1	*2C A4889	Ile/Ile	36	0.243	Ile 0.497 Val 0.503
		Ile/Val	75	0.507	
		Val/Val	37	0.250	
	*4 C4887	Thr/Thr	135	0.912	Thr 0.956 Asn 0.044
		Thr/Asn	13	0.088	
		Asn/Asn	0	0	
	*2A T6235C	T/T	35	0.236	T 0.48 C 0.52
		T/C	72	0.486	
		C/C	41	0.277	
	GST	M1	M1 + M1/M1*0	94	0.635
M1*0			54	0.365	
T1		T1 + T1/T1*0	136	0.919	T1 0.72 * T1*0 0.28 *
		T1*0	12	0.081	
<b>b) Polimorfismos mitocondriales</b>					
EUR	+	3	0.02	0.97	
	-	145			
DEL	+	41	0.27	0.72	
	-	107			
AME	+	67	0.45	0.54	
	-	81			
ASI	+	36	0.24	0.75	
	-	112			
AFR	+	4	0.02	0.97	
	-	144			
AFR2	+	3	0.02	0.97	
	-	145			

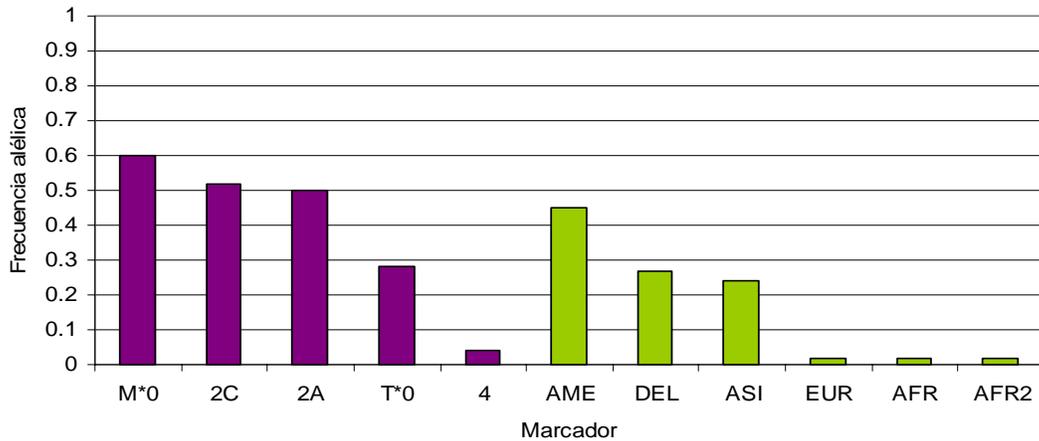


Figura 3.3. Frecuencias alélicas de los polimorfismos tipificados en la población analizada. Para el caso de las GST se grafican las frecuencias alélicas esperadas. En morado, los marcadores nucleares; en verde, los marcadores mitocondriales.

### 3.4. Índice de Fijación

Con el fin de comprobar que los marcadores nucleares están en equilibrio de Hardy-Weinberg, se calculó el índice de fijación (F) que mide la desviación de las frecuencias observadas con respecto a las frecuencias de heterocigotos esperadas según el equilibrio de Hardy-Weinberg. Cuando el valor de F resulta positivo, se podría sospechar de una deficiencia de heterocigotos; mientras que si es negativo, podría haber exceso de heterocigotos. Para verificar la significancia de los valores se realizó una prueba de *chi* cuadrada ( $\chi^2$ ), considerando un 95% de confianza y 1 grado de libertad (g.l.) (Tabla 3.5).

Tabla 3.5. Fórmulas para calcular F y  $\chi^2$ . Fuente: Hedrick, 2000.

$F = 1 - H/2pq$	$\chi^2 \text{ calculada} = F^2N$
H= no. heterocigotos observados	$\chi^2 \text{ establecido} = 3.84, 95\% \text{ confianza, } 1 \text{ g.l.}$

Debido a que ninguno de los valores de  $\chi^2$  calculados resultó mayor al  $\chi^2$  establecido, ninguno de los 3 polimorfismos nucleares del CYP1A1 tiene exceso o deficiencia de heterocigotos y, por lo tanto, están en equilibrio de Hardy-Weinberg (Tabla 3.6). El valor de F no se puede calcular para los marcadores de GST, porque se desconoce el número real de heterocigotos.

Tabla 3.6. Índices de fijación para los polimorfismos del CYP1A1.  $P < 0.05$

Marcador	F	Valor de $\chi^2$
CYP *2C	-0.014	0.029
CYP *2A	0.026	0.100
CYP *4	-0.046	0.313

### 3.5. Diferencias de las frecuencias encontradas respecto a otras reportadas previamente

Con el fin de evaluar las discrepancias encontradas entre las frecuencias estimadas y las ya reportadas, se realizó una prueba de *chi* cuadrada para bondad de ajuste de proporciones. Como poblaciones de referencia se utilizaron las ya reportadas y calculadas a partir de individuos del DF para los polimorfismos CYP1A1 \*2C, \*4 y GST M1 y T1 (Pérez, 2006), así como para el polimorfismo CYP1A1 \*2A (Fragoso *et al.*, 2005). Se realizó una prueba para cada uno de los locus nucleares.

En la tabla 3.7 se detallan las fórmulas y cálculos con una confianza de 95%, así como el valor de  $\chi^2$  establecido de 3.84 correspondiente a 1 g.l. Si la sumatoria de los valores de  $\chi^2$  de todas las categorías para cada locus es menor o igual que el valor de  $\chi^2$  establecido, será indicación de que hay poca discrepancia entre los valores observados en este estudio y los esperados de acuerdo a la proporción reportada y, por lo tanto, no hay diferencias significativas entre los estimados de la población muestreada y la población de referencia.

La prueba resultó significativa para los polimorfismos CYP1A1\*2C, \*2A y GST T1, donde el valor de  $\chi^2$  calculado es mayor al  $\chi^2$  establecido (7.07>3.84, 8.76>3.84 y, 5.51>3.84, respectivamente). Las frecuencias obtenidas de los otros dos locus, CYP1A1\*4 y GST M1, son estadísticamente similares para ambas poblaciones (Tabla 3.7).

Tabla 3.7. Valores de  $\chi^2$  calculados en la prueba de bondad. Las proporciones esperadas se tomaron de Pérez, 2006 y, Fragoso *et al.*, 2005. En negritas se resaltan los valores de  $\chi^2$  calculados superiores al valor  $\chi^2$  establecido. P<0.05

Categoría	Proporción reportada (P)	No. observado en este estudio (O)	No. esperado según referencia E= nP	$\chi^2$ calculado
Ile	0.39	73.5	57.72	4.31407484
Val	0.61	74.5	90.28	2.758179
				<b><math>\Sigma=7.07225384</math></b>
Thr	0.94	141.5	139.12	0.04071593
Asn	0.06	6.5	8.88	0.63788288
				$\Sigma=0.67859881$
T	0.599	71	88.652	3.514789334
C	0.401	77	59.348	5.250271349
				<b><math>\Sigma=8.765060682</math></b>
M1*0	0.33	54	48.84	0.545159705
M1	0.67	94	99.16	0.268511497
				$\Sigma=0.813671202$
T1*0	0.15	12	22.2	4.686486486
T1	0.85	136	125.8	0.827027027
				<b><math>\Sigma=5.513513514</math></b>

### 3.6. Haplotipos mitocondriales y su interpretación de ancestría

Con base en los seis marcadores mitocondriales tipificados, es posible clasificar a la población muestreada en 11 haplotipos mitocondriales. Se encontraron individuos que tienen solamente uno de los polimorfismos, otros que no tienen ninguno y otros más que tienen 2 ó 3. No se encontró ningún individuo que tuviera sólo el polimorfismo AFR, o bien, el AFR2. A cada una de las combinaciones le es atribuible una ancestría específica, de manera que su sumatoria evidencia las contribuciones genéticas de cada continente. Fueron calculadas las frecuencias de cada uno de los 11 haplotipos, siendo los más representados los que tienen ancestría amerindia y asiática (Tabla 3.8).

Tabla 3.8. a) Fórmula para calcular la frecuencia haplotípica. b) Haplotipos mitocondriales, sus frecuencias y su ancestría atribuible.

a) Frecuencia haplotípica = no. individuos por haplotipo/N				
b) Haplotipo	No. de individuos	f haplotípica	Ancestría	
Ame	58	0.392	Amerindia Hap-A	
Del	33	0.223	Amerindia Hap-B	
Asi	30	0.203	Asiática	
n.mt.	11	0.074	No mezcla	
Del/ame	7	0.0473	Amerindia Hap-A/B	
Asi/afr/afr2	3	0.0203	Africana	
Eur	2	0.0135	Europea	
Del/asi	1	0.0067	Asiática	
Eur/ame	1	0.0067	Probable amerindia	
Ame/asi	1	0.0067	Probable asiática	
Asi/afr1	1	0.0067	Africana	

Todos los marcadores mitocondriales estuvieron presentes en la población analizada. Las frecuencias de los polimorfismos en orden descendente son: AME>DEL>ASI>AFR>AFR2=EUR (Tabla 3.4). Sin embargo, para interpretarlos hay que considerar cuál es la ancestría para la que estos marcadores son informativos. A continuación, se detalla qué marcadores fueron utilizados para estimar la ancestría de la población analizada (Tabla 3.8). Con base en esta información, es posible hacer un estimado de ancestría para cada uno de los 11 haplotipos encontrados (Tabla 3.9):

1) *Ancestría amerindia*: como ya se mencionó, todos los nativos americanos caen en uno de los 5 haplogrupos mitocondriales: Hap-A, B, C, D ó X. Cada uno está caracterizado por la presencia de ciertos polimorfismos (Fig. 1.6.b), de tal forma que el polimorfismo AME es específico del Hap-A, mientras que el polimorfismo DEL caracteriza al Hap-B. Los otros tres haplogrupos (Hap-C, D y X) tienen también polimorfismos específicos que los discriminan del resto; sin embargo, dado que en la población mexicana esos tres últimos no son abundantes, los estimados de ancestría no se modifican si no se incluyen en el análisis. Por lo tanto, los polimorfismos AME y DEL reportan ancestría amerindia.

El haplotipo correspondiente al polimorfismo AME es el más abundante (39.2%), seguido del haplotipo correspondiente al polimorfismo DEL (22.3%); finalmente, los individuos con haplotipo DEL/AME pueden considerarse también de ancestría amerindia (4.7%). Tomando en cuenta lo explicado en el párrafo anterior, la ancestría amerindia en la población analizada podría expresarse como aproximada a la sumatoria de las frecuencias de los polimorfismos AME y DEL (específicos de los Hap-A y B, respectivamente) más la combinación AME/DEL (correspondientes a individuos con mezcla de Hap-A/B), arrojando un valor total de ~66.2% ( $\sum[P_{AME}+P_{DEL}+P_{AME/DEL}] \approx 66.2$ , donde P= proporción).

2) *Ancestría asiática*: el polimorfismo ASI está presente en proporciones muy altas en la población asiática, así como en los Hap-C, D y X de la población americana. No obstante, sólo se le utilizó como reportero de ancestría asiática debido a que, como ya se explicó, los haplogrupos amerindios C, D y X están casi ausentes en la población mexicana. Sin embargo, no se debe olvidar que los tres polimorfismos ya mencionados son compartidos tanto por la población asiática como por la americana, dada la divergencia tan cercana de sus poblaciones, de tal forma que ocurren combinaciones de éstos y se les puede incluir en cualquiera de las dos categorías. Para calcular la contribución asiática se consideró la proporción del haplotipo correspondiente al polimorfismo ASI (20.3%) y se le sumó la proporción con una combinación de polimorfismos DEL/ASI; dando un valor total para la ancestría asiática de ~20.97% ( $\sum[P_{ASI}+P_{DEL/ASI}] \approx 20.97$ , donde P= proporción).

3) *Ancestría europea*: el polimorfismo EUR se utilizó como reportero de ancestría europea. La proporción ancestral europea en la población analizada es de ~1.35% dada por el haplotipo EUR.

4) *Ancestría africana*: tanto el polimorfismo AFR como el AFR2 indican ancestría africana. El valor total de ancestría africana es de ~2.7% dada por los haplotipos ASI/AFR/AFR2 y ASI/AFR.

5) *No mezcla*: se encontraron además 11 individuos que no tenían ninguno de los 6 marcadores mitocondriales tipificados. Éstos no han adquirido algún haplotipo específico. Éste es el caso del ~7.4% de la población analizada.

6) El resto del porcentaje comprende a los individuos que tienen combinaciones de los marcadores cuya ancestría no es clara y, aunque podrían sumarse a los valores dados arriba, su contribución no puede asegurarse si no se tipifican más marcadores que los clasifiquen rigurosamente. Así, únicamente se tomaron en cuenta los haplotipos que dan información clara de la ancestría en las proporciones, aunque no varían los estimados dado que hay sólo dos de estos individuos. Por ejemplo, el individuo de haplotipo AME/ASI podría pertenecer a alguno de los HAP poco comunes en México; mientras que el individuo de haplotipo EUR/AME seguramente es también de ancestría amerindia.

**Tabla 3.9. Porcentajes de los estimados de contribución de ADNmt a la población analizada. N.mt=ningún marcador mitocondrial.**

Ancestría	Haplotipo	f haplotípica (%)	(%) de contribución a la población
Amerindia	AME	39.2	66.23
	DEL	22.3	
	AME/DEL	4.73	
Asiática	ASI	20.3	20.97
	DEL/ASI	0.67	
No mezcla	n.mt.	7.4	7.4
Africana	ASI/AFR/AFR2	2.03	2.7
	ASI/AFR	0.67	
Europea	EUR	1.35	1.35
Desconocida	AME/ASI	0.67	1.34
	EUR/AME	0.67	

### 3.7. Asociación de marcadores mitocondriales vs. nucleares

Para analizar si alguno de los polimorfismos nucleares está sobrerrepresentado en alguno de los haplotipos mitocondriales, se realizó una prueba de asociación de los 11 haplotipos mitocondriales con cada uno de los locus nucleares. Así, se calcularon las frecuencias genotípicas esperadas de cada locus para los 11 haplotipos mitocondriales encontrados en la población y, en consecuencia, conocer el número de individuos esperados para cada haplotipo. Estas cifras se compararon con el número de individuos observados mediante una prueba de *chi* cuadrada. Las fórmulas utilizadas para el cálculo de los valores de  $\chi^2$  se mencionan en la tabla 3.10. Si la diferencia entre el número de individuos observados y el número de individuos esperados para cada haplotipo, de acuerdo a sus valores de  $\chi^2$ , es estadísticamente significativa, se derivaría que dicho haplotipo está asociado con la presencia de un genotipo. Es decir, que los individuos de cierta ancestría con un haplotipo específico tienen probabilidades altas de ser portadores de un polimorfismo nuclear.

**Tabla 3.10. Fórmulas para calcular datos necesarios para la asociación.** Las frecuencias se calculan para cada haplotipo mitocondrial y para cada locus nuclear.

$$f \text{ genotípica esperada} = f \text{ haplotípica mitocondrial} * f \text{ genotípica nuclear}$$

$$\text{No. de individuos esperados} = f \text{ genotípica esperada} * N$$

$$\chi^2 = (\text{no. individuos observado} - \text{no. individuos esperados})^2 / \text{no. individuos esperados}$$

$$\chi^2 \text{ establecido} = 2.7, 90\% \text{ confianza, 1 g.l.}; \text{ o bien, } 3.84, 95\% \text{ confianza, 1 g.l.}$$

La prueba de asociación resultó positiva sólo para GST M1\*0 con el haplotipo ASI/AFR/AFR2, con un valor de  $\chi^2$  calculado de 3.32 y un valor de  $\chi^2$  establecido de 2.7 con una confianza del 90% (Tabla 3.11). Dada la diferencia significativa, se puede decir que el linaje materno ASI/AFR/AFR2 de ancestría africana se encuentra preferencialmente asociado al genotipo nuclear GST M1\*0. Los

demás polimorfismos no tuvieron diferencias significativas y, por lo tanto, ninguna asociación. Se realizaron nuevamente los cálculos de  $\chi^2$  uniendo los dos haplotipos africanos ASI/AFR/AFR2 y ASI/AFR; la asociación se conserva y se hace más significativa, arrojando un valor de  $\chi^2$  calculado de 4.42 y un valor de  $\chi^2$  establecido de 3.84 con una confianza del 95% (Tabla 3.11).

Tabla 3.11. Valores obtenidos en la prueba de asociación para la GST M1 y: a) haplotipo ASI/AFR/AFR2 ( $P < 0.10$ ) y; b) ambos haplotipos africanos ( $P < 0.05$ ).

a) Genotipo	f genotípica esperada	No. esperado	No. observado	Valor $\chi^2$
M1*0	0.007395909	1.094594595	3	3.316816817
M1	0.012874361	1.905405405	0	1.905405405
b) Genotipo	f genotípica esperada	No. esperado	No. observado	Valor $\chi^2$
M1*0	0.009861213	1.459459459	4	4.422422422
M1	0.017165814	2.540540541	0	2.540540541

Debido a que la prueba de  $\chi^2$  establece que el número de individuos esperados por clase debe ser mayor a 5, los resultados arrojados no pueden asegurar la asociación. Se recomienda la prueba de G para corregir tales errores; sin embargo, tampoco es aplicable en estos datos dado que para el caso de GST M1 hay clases con ningún individuo.

### 3.8. Análisis de distancia nucleotídica vs. distancia geográfica

Con el fin de averiguar si las distancias geográficas que separan las diferentes entidades federativas del país influyen en la diferenciación genética entre sus habitantes, se realizó un análisis de las distancias geográficas y nucleotídicas entre los estados. La información utilizada se basó en los lugares de procedencia de las abuelas maternas de los donadores. Para asegurar mayor representatividad de estados, sólo se incluyeron datos de los estados con registro de 2 o más abuelas maternas nacidas en ellos. Las distancias geográficas corresponden a la distancia en kilómetros entre la capital de un estado y otro; mientras que las distancias nucleotídicas se calcularon con base en las frecuencias haplotípicas mitocondriales por estado (Tabla 3.12). Entre más grande sea el valor de la distancia nucleotídica, más grande será la variación genética entre las entidades federativas comparados. El valor cero significa que son iguales genéticamente.

**Tabla 3.12. Matrices de datos de a) distancias geográficas en km y, b) nucleotídicas de los estados de nacimiento de las abuelas maternas de los donadores. Distancias geográficas tomadas de Guía Roji México, 2004.**

<b>a) Distancias geográficas (Km)</b>											
<b>Oax</b>	0										
<b>Gro</b>	711	0									
<b>Mich</b>	772	580	0								
<b>Jal</b>	1012	820	302	0							
<b>Tab</b>	609	1009	1070	1310	0						
<b>Ver</b>	395	643	704	944	492	0					
<b>Pue</b>	347	364	425	665	645	279	0				
<b>Hgo</b>	534	373	397	615	832	466	187	0			
<b>EMex</b>	536	344	236	476	834	468	189	161	0		
<b>DF</b>	470	278	302	542	768	402	123	95	66	0	
<b>Gto</b>	835	643	180	277	1133	767	488	392	349	365	0
	<b>Oax</b>	<b>Gro</b>	<b>Mich</b>	<b>Jal</b>	<b>Tab</b>	<b>Ver</b>	<b>Pue</b>	<b>Hgo</b>	<b>EMex</b>	<b>DF</b>	<b>Gto</b>
<b>b) Distancias nucleotídicas</b>											
<b>Oax</b>	0										
<b>Gro</b>	0.367	0									
<b>Mich</b>	0.436	0.086	0								
<b>Jal</b>	0.367	0	0	0							
<b>Tab</b>	0.367	0	0	0	0						
<b>Ver</b>	0.413	0.117	0	0.403	0.403	0					
<b>Pue</b>	0.137	0	0	0.145	0.145	0	0				
<b>Hgo</b>	0.232	0	0	0	0	0	0.146	0			
<b>EMex</b>	0.419	0.085	0	0.518	0.518	0.041	0.215	0.502	0		
<b>DF</b>	0.373	0.015	0	0.253	0.253	0	0.198	0.343	0	0	
<b>Gto</b>	0.234	0	0	0.025	0.025	0	0.115	0.128	0	0	0
	<b>Oax</b>	<b>Gro</b>	<b>Mich</b>	<b>Jal</b>	<b>Tab</b>	<b>Ver</b>	<b>Pue</b>	<b>Hgo</b>	<b>EMex</b>	<b>DF</b>	<b>Gto</b>

Con estos datos se realizó una prueba de Mantel para saber si existe correlación entre ambas matrices. El análisis demostró que no existe correlación entre las distancias geográficas y nucleotídicas ( $r = 0.0379$ ); sin embargo, es posible identificar al menos tres agrupaciones de los estados. Los tres grupos están conformados por los siguientes estados: 1) Sureste: Oaxaca, Veracruz, Tabasco y Guerrero; 2) Centro: Puebla e Hidalgo; 3) Bajío: Michoacán, Jalisco y Guanajuato (Fig. 3.6).

A partir de los tres grupos de estados identificados, fueron calculadas las frecuencias de los polimorfismos nucleares de los individuos cuyas abuelas maternas pertenecen a alguno de los grupos, con el fin de rastrear alguna posible tendencia de variación. Las frecuencias son muy similares para los tres grupos en todos los polimorfismos, a excepción de la GST M1 para el grupo del Sureste, que tiene los valores invertidos respecto a los otros dos grupos (Fig. 3.5). Por lo tanto, los grupos del Centro y Bajío son mucho más similares entre ellos.

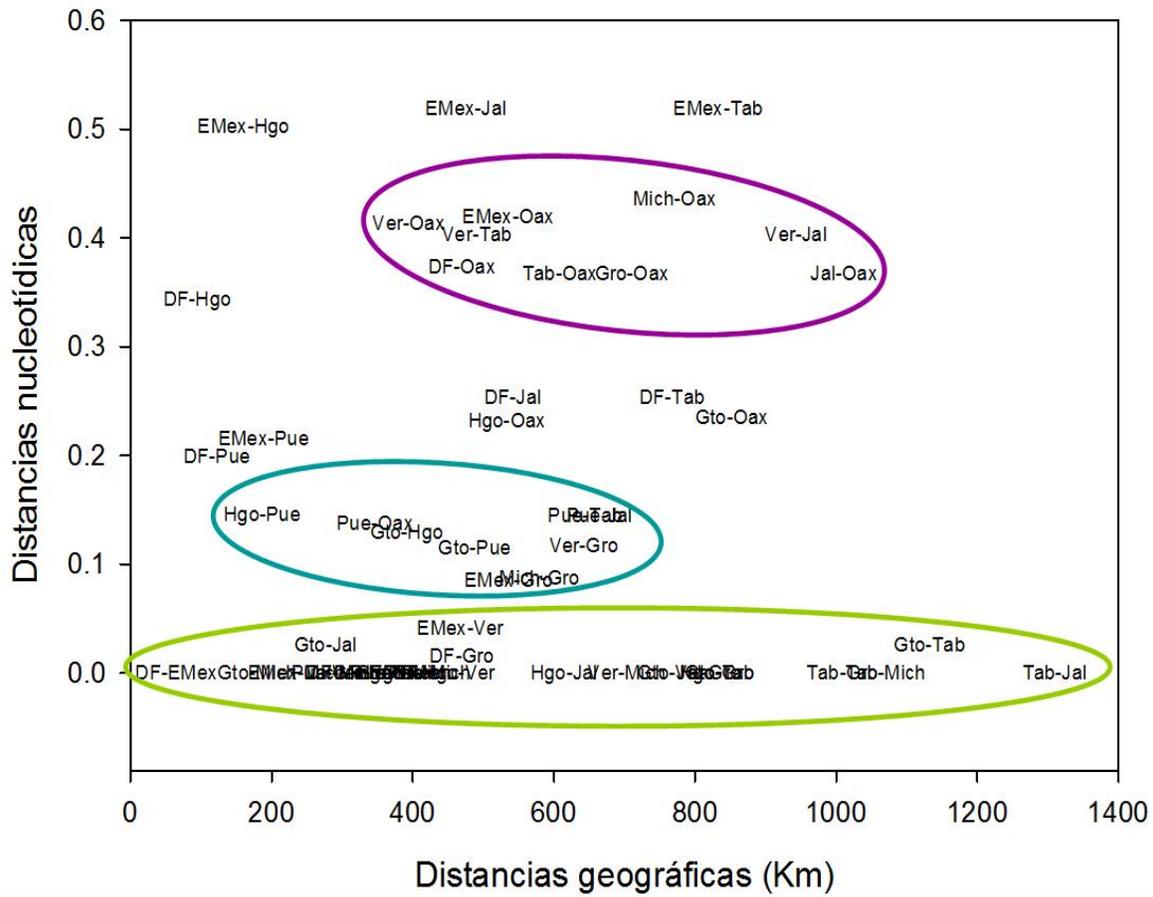


Figura 3.4. Gráfica de correlación entre entidades federativas.

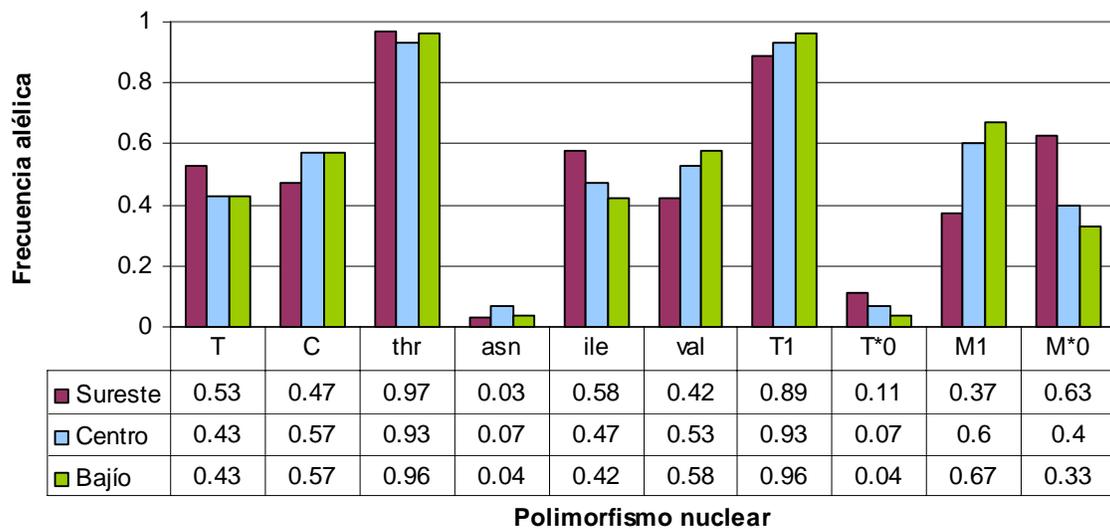


Figura 3.5. Gráfica de las frecuencias alélicas de los polimorfismos nucleares tipificados de los donadores de cada uno de los tres grupos de estados identificados según el lugar de procedencia de sus abuelas maternas.

## 4. DISCUSIÓN

### 4.1. Flujo génico a las poblaciones mestizas mexicanas

A pesar de que se intentó delimitar lo más posible la zona de muestreo para disminuir la variación, es claro que las fluctuaciones entre las frecuencias de la población mexicana son producto de una mezcla genética entre habitantes procedentes de muchos estados del país. Esto refleja que ha habido un gran flujo génico proveniente no sólo de los otros estados de la República, sino incluso de otros países. En este estudio, de los 148 individuos analizados, 4 dijeron tener abuelas maternas extranjeras (2.7%), sin contar a los que tenían familiares extranjeros de lado paterno. Sin duda, este fenómeno es mucho más marcado en el DF, puesto que por ser la capital del país, constituye una zona de convergencia para habitantes procedentes de todo el país y del mundo. Por lo tanto, recibe migrantes de origen étnico diverso. Por ejemplo, en una sola delegación del DF, es posible encontrar tanto mexicanos nacidos en el interior de la República (mestizos e indígenas) como extranjeros.

Lo anterior podría sugerir que la población mestiza mexicana analizada en este estudio es homogénea dada la contribución genética proveniente de diversas regiones del país, como se evidencia en los mapas de la Figura 3.1 y la Tabla 3.1. El registro de la historia genética de la población estudiada cubre casi la totalidad del territorio nacional rebasando los lugares de nacimiento, incluso cuando tan sólo se cuenta con información de dos generaciones anteriores. Con todo, debe considerarse que el influjo de individuos procedente de otros estados no es en todos los casos alto y constante; es decir, la tasa de intercambio de individuos es variable según la región y el tiempo. Este hecho rebate la afirmación de que la población puede ser homogénea. Además, es muy probable la existencia de variaciones interestatales, sin contar que la intensidad del flujo génico podría ser diferente para cada región del DF.

A pesar de que el flujo génico producido, sea alto o bajo, actúa como factor homogeneizante, el ocurrido en el DF, puede no ser suficiente para eliminar posibles subpoblaciones existentes. Más aún, el producido por individuos provenientes de otros países, que claramente tienen haplotipos diferentes a los nuestros, puede no sólo alterar las frecuencias alélicas, sino introducir nuevas variantes. Así, no únicamente la población mexicana en general es heterogénea, sino también lo podrían ser poblaciones mestizas más localizadas, como la población del DF.

Por población homogénea se entiende aquella cuyas frecuencias alélicas son similares a lo largo de ésta; consecuentemente, no presenta subpoblaciones. Para el caso mexicano, existen

estudios donde se demuestra que diferentes poblaciones indígenas son homogéneas genéticamente (Cerdeña-Flores *et al.*, 2002) y otros donde se prueba lo mismo para poblaciones mestizas de diferentes estados del país (Buentello-Malo *et al.*, 2003; Bonilla *et al.*, 2006). Esto significa que los individuos que las conforman, si bien tienen diferencias genéticas, éstas no son tan considerables como para distinguir subgrupos dentro de ellas. Ahora bien, que dos poblaciones sean homogéneas no significa que sean iguales; es decir, no presentan necesariamente las mismas frecuencias alélicas entre ellas.

Por ejemplo, dos poblaciones indígenas pueden ser homogéneas (sin subestructuras), pero no necesariamente iguales (pueden presentar variación interétnica en sus frecuencias alélicas). Si estas mismas no tuvieran diferencias significativas en sus frecuencias alélicas, podrían tomarse individuos de ambas poblaciones como muestra, sin que la representatividad de la población global se vea alterada. Si estudiáramos la población total del país como una sola, entonces, podríamos decir que es heterogénea dado que algunas regiones del país presentan variaciones interétnicas significativas y que es posible identificar subpoblaciones bien delimitadas, principalmente, aquéllas que están aisladas geográficamente y compuestas por individuos primordialmente indígenas.

Por otra parte, en los estudios de genética de poblaciones no es posible hacer generalizaciones, así que los resultados arrojados en un estudio sólo son válidos para la población analizada. No sólo no se puede afirmar que la población mestiza mexicana es homogénea en general; hasta no realizar muestreos representativos de los habitantes del estado completo, se podrá decir que la población mestiza de una sola entidad federativa lo es. Debido a que son pocos los estudios que se han hecho respecto a la población mexicana, se han utilizado los parámetros de otras poblaciones para hacer inferencias de la nuestra. Por esta razón, la caracterización genética de nuestra población es de gran importancia para conocer cuál ha sido su historia y filogenia, así como para analizar su estratificación, lo cual permitirá hacer comparaciones con otras poblaciones y disminuir errores en estudios de asociación.

#### **4.2. Marcadores nucleares y su variación interétnica**

Las frecuencias alélicas de los polimorfismos CYP1A1\*4 y GST M1\*0 analizados en esta población no presentan diferencias significativas a las ya reportadas en estudios anteriores hechos para el DF (Pérez, 2006). Sin embargo, la frecuencia de los polimorfismos CYP1A1\*2C y GST T1\*0 disminuye ligeramente en la población analizada, mientras que la del polimorfismo CYP1A1\*2A aumenta (Tabla 3.7). Dado que los estudios sobre la población mexicana son escasos, estas variaciones pueden deberse a que no se han realizado estudios cuyos muestreos limiten regiones

específicas de las poblaciones mestizas y que permitan descartar la existencia de subestructuras en el DF, las cuales, podrían ser la causa de las fluctuaciones en las frecuencias de los diferentes estudios. Así, es probable que la población mestiza del DF no sea totalmente homogénea.

Ahora bien, ya se han demostrado variaciones interétnicas en las frecuencias de algunos de estos polimorfismos entre mestizos y poblaciones indígenas mexicanas (Fragoso *et al.*, 2005). Éstas podrían ser suficientes como para proponerlas como subpoblaciones. Pérez propone que las diferencias interestatales de los polimorfismos GST M1, T1 y CYP1A1 \*2C y \*4 son muy bajas, por lo que podría considerarse a las poblaciones mestizas mexicanas similares genéticamente (2006). Sin embargo, para comprobar la hipótesis de este autor sería necesario realizar más estudios, ya que los tamaños de muestra utilizados para los estados (excepto para el DF) son estadísticamente insuficientes.

Además, hay reportes donde se encuentran diferencias entre las contribuciones genéticas de las poblaciones fundadoras ancestrales para los diferentes estados del país (Bonilla *et al.*, 2005). Según estos resultados, se esperarían variaciones en las frecuencias alélicas de algunos genes. Los estudios de varianza de estos polimorfismos nucleares en diferentes países son mucho más abundantes y quedan claras las diferencias entre sus frecuencias alélicas.<sup>1</sup> Así, se esperaría que regiones del país –e incluso de estados específicos– con alto flujo génico de una población A, tengan frecuencias similares a ésta, siempre que otros procesos evolutivos actúen en menor medida sobre ella.

Por otra parte, los índices de fijación calculados para los polimorfismos analizados del CYP1A1, no reflejaron exceso ni deficiencia de heterócigos, por lo que podemos concluir que los tres están en equilibrio de Hardy-Weinberg y, posiblemente, las frecuencias alélicas se han mantenido sin cambios. El equilibrio de Hardy-Weinberg establece que en una población, las frecuencias alélicas de un gen son iguales a las de la generación anterior debido a que no están sujetas a la evolución, lo cual, significa que las frecuencias de los marcadores utilizados pueden mantenerse constantes si ninguno de los procesos evolutivos antes mencionados comienzan a moldearlas. Sin embargo, es importante recalcar que para que sea observada alguna desviación pequeña (con un índice de fijación pequeño) en el equilibrio de Hardy-Weinberg, se necesitaría un tamaño de muestra muy alto. Además, algunos procesos como la deriva génica y las mutaciones, no provocan desviaciones significativas; y otras, como el flujo génico y el apareamiento al azar, aunque estén presentes pueden no cambiar las frecuencias genotípicas (Hedrick, 2000). Para el

---

<sup>1</sup> Garte *et al.*, (2001) hacen una revisión de varios estudios de diferentes países y combinan los cálculos de sus frecuencias dando por un lado, un valor para las poblaciones asiática, europea y africana y, por otro, un valor más específico a poblaciones de algunos países. En este metaanálisis queda clara la variación interétnica para algunos polimorfismos y sus combinaciones.

caso de los polimorfismos de la GST no se calcularon índices de fijación, ya que no es posible discriminar a los heterócigos por el método utilizado.

#### **4.3. Marcadores mitocondriales y sus estimados de ancestría**

En la tabla 3.9 se puede observar que los estimados de ancestría obtenidos a partir de las frecuencias alélicas de los marcadores mitocondriales tipificados, reportan que la población mestiza mexicana en estudio presenta en orden descendente contribución genética amerindia, asiática, no mezcla, africana y, europea.

Este estimado demuestra que la población mestiza mexicana analizada conserva mucha de la contribución genética de linaje materno proveniente de las poblaciones de los ancestros indígenas, aún cuando ha estado sujeta a altos y diversos flujos génicos procedentes de muchas regiones del planeta. Es muy posible que estos estimados se conserven en los estados cercanos al lugar de muestreo; sin embargo, podrían cambiar en estados del norte o costeros que han tenido – y aún están sujetos a– diferentes patrones de inmigración. Los reportes son contradictorios puesto que algunos autores reportan que las poblaciones mestizas de diferentes estados del país presentan los mismos porcentajes de ancestría (Cerdeña-Flores *et al.*, 2002), mientras que otros reportan lo contrario. Bonilla *et al.*, (2005) hacen una revisión de varios de esos trabajos (Tabla 1.3). Estas diferencias pueden estar basadas en el uso de diferentes marcadores genéticos para el análisis.

Por otra parte, los resultados indican que el Hap-A amerindio es el más frecuente en la población mestiza analizada. Esto concuerda con otros estudios, donde el Hap-A es el más abundante, seguido del Hap-B; mientras que las proporciones de los Hap-C y D son casi nulas. También se registran algunos individuos con la combinación Hap-A/B. En este trabajo, también se reporta que las comunidades indígenas nahuas se caracterizan por portar principalmente el Hap-A, mientras que las mixtecas portan principalmente el Hap-B (Bonilla *et al.*, 2005).

En orden de abundancia, sigue el estimado de ancestría asiática (como era de esperarse); dada la historia del continente americano donde las poblaciones fundadoras eran de origen asiático y pasaron a través del Estrecho de Bering hace unos 14 a 21 mil años, desplazándose hacia el sur del continente. Si bien la población mexicana adquirió una proporción considerable de contribución europea a partir de la conquista española, la mayoría de los inmigrantes eran hombres; de modo que los marcadores mitocondriales utilizados no la detectan porque sólo rastrean la rama materna dada su herencia uniparental. Otros estudios que utilizan marcadores autosómicos han reportado que los porcentajes de contribución española y amerindia son casi iguales en el DF y muy similares

en otras poblaciones mestizas de Nuevo León y Jalisco (Cerde-Flores *et al.*, 2002); resultados similares son arrojados por otros más (Bonilla *et al.*, 2005) (Tabla 1.3). El país sigue recibiendo migrantes no sólo de España sino de toda Europa, de manera que su contribución genética no es despreciable.

Aún cuando es baja, la proporción de ancestría africana refleja la inmigración de individuos africanos durante la época de la Colonia que, en el caso de México, ocurrió principalmente por zonas costeras del Golfo de México, así como por zonas del Pacífico cerca de Acapulco, además de aquellos que llegaron huyendo de plantaciones cercanas de los estados vecinos. Los principales países de exportación al continente americano fueron Marruecos, Costa de Oro y Senegal. En la actualidad el país recibe poca inmigración africana, a diferencia de Estados Unidos, países de Sudamérica y Antillas, cuyas poblaciones presentan mayor proporción de mezcla africana.

El polimorfismo AFR2 encontrado además de los 5 que de inicio se buscaban, refleja que la población mestiza mexicana es una gran mezcla de genotipos, y que algunos individuos tienen bien marcada la ancestría que aún conservan de otros continentes. Tal es el caso de los individuos con este polimorfismo que es muy específico del continente africano, y del cual no se reportan sus frecuencias, por no ser abundante (salvo en poblaciones como la cubana, con clara inmigración africana) (Torroni *et al.*, 1995). La existencia de 11 haplotipos en una población de 148 individuos apoya también la idea de que la población mestiza mexicana es muy variada genéticamente, tomando en cuenta que sólo se utilizaron 1 ó 2 marcadores genéticos por continente y, que si bien proporcionan bastante información, no todos los individuos del continente para el que son específicos lo portan. Por lo tanto, arrastran de por sí un error en la estimación de la ancestría, lo cual puede ser corregido si se incluyen más marcadores.

Además, la utilización de marcadores mitocondriales únicamente nos da información sobre la rama materna, que en el caso de la contribución genética europea a la población mexicana, podría subestimarse. En la actualidad, la magnitud del flujo génico en la población mexicana es similar tanto para mujeres como para hombres ya que: 1) los índices de migración interestatal son similares para mujeres y hombres, aunque podrían considerarse de tipo patrilocal puesto que el porcentaje de mujeres migrantes es ligeramente mayor (las mujeres son las que migran para establecerse en el lugar de origen del hombre al casarse); 2) los porcentajes de mujeres y hombres extranjeros inmigrantes en México también son muy similares. Sin embargo, el porcentaje de mujeres mexicanas que emigran a otros países es muy baja en comparación con el de hombres; por lo que, el ADNmt refleja con mucha fidelidad la historia de migración en las poblaciones mexicanas (*Estadísticas del INEGI sobre porcentajes de población inmigrante por sexo de 1950 a 2000: población nacida en otro país residente en México por entidad federativa según sexo; población emigrante a Estados Unidos de América por entidad federativa según sexo, 2000*).

Por ello, los marcadores utilizados fueron muy informativos sobre la contribución genética que los distintos continentes han tenido en la población mexicana, al menos para la población mestiza mexicana analizada; y si bien, podrían incluirse otros marcadores para aumentar la confianza en los valores estimados, los individuos de haplotipos de ancestría dudosa son sólo dos, de manera que los 6 marcadores utilizados resultaron muy informativos y suficientes para el análisis de esta población. Más aún, podrían ser utilizados en otras poblaciones mestizas mexicanas con la certeza de que se obtendrán datos confiables.

Lo anterior es apoyado por el hecho de que los estimados de ancestría arrojados por los seis marcadores mitocondriales usados en este estudio son similares a las ya reportadas utilizando marcadores autosómicos, grupos sanguíneos y marcadores proteícos (Bonilla *et al.*, 2005; Cerda-Flores *et al.*, 2002) (Tabla 1.3), considerando la variación de por sí producida por el patrón de herencia del tipo de marcador usado. La similitud de las proporciones de mezcla génica reportadas en estudios anteriores con las del presente, permite validar la fidelidad de los marcadores mitocondriales utilizados. Sin embargo, es importante recalcar que los patrones de mezcla génica pueden cambiar según el marcador usado de acuerdo a sus patrones de herencia; por ejemplo, si en la misma población en estudio, se utilizaran marcadores en cromosoma Y, entonces se esperaría un decremento en la proporción de ancestría asiática y amerindia, y un aumento en la europea. Los valores de las proporciones utilizando marcadores autosómicos caerían entre los estimados mitocondriales y los de cromosoma Y.

Pese a esto, es necesario subrayar que el ADNmt es muy útil para explorar en el pasado humano, puesto que aún en periodos cortos de tiempo, como el transcurrido de la conquista española a nuestros días, es posible estudiar la diversidad en la población y estimar fechas de sucesos evolutivos relativamente recientes debido a la tasa tan alta de mutación del ADNmt comparada con la del ADN nuclear ( $3.4 \times 10^{-7}$  y  $2.5 \times 10^{-8}$  por generación de 20 años, respectivamente). Consecuentemente, también se puede estimar su alta velocidad evolutiva (Jobling, 2004). Adicionalmente, el ADNmt tiene un menor número efectivo de la población (en un entrecruzamiento, así como se tienen 4 copias de cada autosoma, se tienen dos de ADNmt; sin embargo sólo se hereda uno, de modo que el tamaño efectivo del ADNmt es  $\frac{1}{4}$  del de los autosomas), así que es más sensible en la identificación de estructuras poblacionales y nos permite discriminar entre los eventos en los que estuvieron o no involucradas específicamente las mujeres. Su manipulación no representa ningún problema en la realización de los análisis; puesto que es igual de estable que el ADN genómico, no requiere técnicas especiales y dado que el número de copias de ADNmt por célula es alto, facilita el análisis de muestras escasas o antiguas.

#### **4.4. Asociación GST M1\*0 – haplotipos africanos**

Dadas las variaciones étnicas en las frecuencias alélicas de los marcadores nucleares analizados para algunas poblaciones, se pensó que las frecuencias altas reportadas de algunos de ellos en poblaciones específicas podrían estar relacionadas con su ancestría, lo cual es importante ya que significaría que individuos con cierta ancestría podrían ser portadores de uno o más de los polimorfismos nucleares. De encontrarse una asociación alta, el locus nuclear podría utilizarse como un marcador de ancestría. De mayor trascendencia es el hecho de que si el locus es de importancia médica, pueden identificarse zonas de riesgo. Tal es el caso de los polimorfismos nucleares utilizados en este estudio, que se han relacionado directamente con la susceptibilidad a desarrollar cáncer de pulmón por estar implicados en la desintoxicación del organismo de xenobióticos tales como los HAP, presentes en el humo del tabaco y en el aire urbano.

Los análisis de asociación se realizan con el fin de buscar alelos específicos en diferentes locus que aparecen juntos más de lo que podrían aparecer por azar. El presente análisis se realizó con base en lo anterior. Se trata de asociar un haplotipo mitocondrial con otro locus nuclear para probar que la alta probabilidad de aparecer juntos no se da al azar. Una diferencia significativa entre el número de individuos que se espera tenga un cierto haplotipo (según las probabilidades previamente calculadas) con los individuos observados representan una variación no al azar; por tanto, se sospecha de una asociación entre esos locus.

En el análisis sólo se encontró una posible asociación del haplotipo ASI/AFR/AFR2, de ancestría netamente africana, con la presencia del polimorfismo GST M1\*0 (Tabla 3.11). Como ya se mencionó, la prueba de *chi*-cuadrada utilizada plantea que el número de individuos esperados por clase debe ser mayor a 5; sin embargo, dado que la muestra analizada se reparte entre los 11 haplotipos mitocondriales encontrados, algunos tienen un número muy escaso de individuos, como sucede con el haplotipo en cuestión. Esto provoca que las frecuencias haplotípicas mitocondriales sean pequeñas y al multiplicarlas por las frecuencias genotípicas de cada locus arrojan números de individuos esperados menores a 5. Para corregirlo y asegurar la asociación, sería necesario coleccionar un número muy grande de individuos. En este caso, la probabilidad de obtener un individuo con haplotipo ASI/AFR/AFR2 y que sea GST M1\*0 es 0.0074; por lo que se necesitarían muestrear 676 individuos para encontrar un número de 5 individuos esperados.

Pero si se considera que el haplotipo mitocondrial ASI/AFR/AFR2 tiene una frecuencia de por sí baja en la población mestiza mexicana y, que los 3 individuos con ese haplotipo son además GST M1\*0, entonces es clara la relación que pueden tener estos dos locus. Más aún, el individuo con el otro haplotipo mitocondrial de ancestría africana, AFR/AFR2, también es M1\*0, lo cual, fortalece la asociación con un mayor nivel de confianza (Tabla 3.11). Dado que la frecuencia del

polimorfismo M1\*0 no es tan baja (0.365), era de esperarse que al menos 2 de los 4 individuos no portaran la delección. Por el contrario, se encontró que todos son GST M1\*0 (es decir, todos portan la delección). Esto indica que además los haplotipos africanos en poblaciones mexicanas podrían ser nulos para la GST M1 y, de comprobarse, serían más susceptibles a los efectos a los que estén relacionados, tales como susceptibilidad a desarrollar cáncer de pulmón.

La importancia de comprobar la existencia de asociaciones entre marcadores de ancestría y polimorfismos relacionados con padecimientos radica en la posibilidad de identificar grupos de riesgo. Para el caso en estudio, dichos grupos de riesgo están ubicados en los estados de la República Mexicana con mayor proporción de ancestría africana como Veracruz, Guerrero y Oaxaca. Las frecuencias del polimorfismo GST M1\*0 en poblaciones africanas no son superiores a las observadas en otras poblaciones como la caucásica; incluso son muy similares a las poblaciones mestizas mexicanas. Es importante aclarar que el hecho de que haya asociación GST M1\*0 y los haplotipos africanos, no implica que la frecuencia de este locus nuclear sea alta en poblaciones del continente africano debido a que la historia evolutiva del locus en la población mestiza a partir de su introducción es totalmente independiente de la ancestral. Más aún, si se toma en cuenta a la población africana como la más ancestral. Por esto, el hecho de que una población dada presente asociación entre dos locus, no significa que también la presenten las poblaciones anteriores.

Cuando un pequeño número de individuos migra y funda una nueva población, sus frecuencias alélicas serán modificadas principalmente por deriva génica hasta llegar al equilibrio de Hardy-Weinberg. Por probabilidad, los alelos raros son escasamente introducidos a la nueva población por los fundadores, de modo que desde el principio perderá variación genética. Luego, si la nueva población no aumenta en número rápidamente, la deriva génica actuará para eliminar la poca variación genética de los individuos fijando el alelo introducido más abundante ("Efecto Fundador"). Si la población crece, se conservará mayor variación genética y se generarán nuevas mutaciones, adoptando sus propias particularidades genéticas. Esto sucedió con las poblaciones americanas.

En poblaciones africanas la frecuencia del polimorfismo GST M1\*0 es de ~0.3; mientras que en poblaciones afroamericanas está presente en el 20 y 30% de sus individuos. Se han encontrado asociaciones entre la presencia del polimorfismo GST M1\*0 y la incidencia de cáncer de pulmón en poblaciones afroamericanas (Ford *et al.*, 2000). Esto es de gran relevancia, ya que las frecuencias de este polimorfismo son mucho más altas en otras poblaciones (0.5 en caucásicos [Ford *et al.*, 2000]) y no así su asociación con la incidencia de cáncer.

La ausencia de asociación para los demás locus nucleares es explicable debido a que la utilización de ADNmt no toma en cuenta la herencia paterna. El problema de los marcadores uniparentales es que solamente se obtienen las proporciones de uno de los progenitores; si bien esto es útil para ciertos fines como estimaciones de ancestría, en otros se pierde información como en el caso de análisis de asociaciones. Cabe la posibilidad de no encontrar asociación entre dos locus debido a que no es rastreada por el marcador uniparental utilizado, y no a que está realmente ausente. Pese a esto, la utilización de marcadores que rastreen información sobre ambos progenitores al mismo tiempo; es decir, en cromosoma *y*, *x* y/o mitocondriales, no asegura corregir el error porque la utilización de locus lejanos al gen de interés disminuye considerablemente la sensibilidad del análisis.

Esto es, la asociación de genes se da porque se heredan juntos y por lo tanto uno es marcador de la presencia de otro. Conforme más lejanos estén unos de otros, menos posibilidad de encontrar asociación entre ellos y, menos aún de encontrar asociación cuando se tratan de genes en genomas distintos (mitocondrial vs. nuclear). Una solución sería escoger aquellos AIM de ancestría esperada que se encuentren cercanos al gen de interés o, por lo menos, en el mismo cromosoma. El problema es que hay muy pocos marcadores de continente específicos, de modo que son necesarios muchos genes para poder clasificar a los individuos en una ancestría definida, lo que nuevamente hace menos sensible el análisis.

#### **4.6. Correlación distancia geográfica – distancia nucleotídica**

En la mayoría de las poblaciones, a mayor distancia geográfica, se genera un aislamiento por distancia, más aún si existe una barrera geográfica de por medio. De manera que sus integrantes tienen menor probabilidad de entrecruzamiento y, por tanto son más divergentes genéticamente (Futuyma, 1998). Actualmente, si bien en las poblaciones humanas también ocurre el aislamiento por distancia, las barreras geográficas impiden poco la migración y el intercambio genético; no obstante, las poblaciones siguen sin tener reproducción al azar debido a la existencia de aislamientos por lenguaje y cultura.

En la población analizada, la falta de una correlación entre las distancias geográficas y nucleotídicas demuestra que la lejanía entre los estados de la República no determina sus diferencias genéticas. Por consiguiente, estados alejados geográficamente pueden tener distancias nucleotídicas mínimas (Tabla 3.12). Esto sugiere que el flujo génico entre ellas ha sido de magnitud tal, que ha logrado homogeneizarlos más que otros y, por consiguiente, se encuentran menos diferencias genéticas entre ellos o, incluso, que hay rutas de migración preferidas a pesar de las distancias. No obstante, es posible identificar grupos de estados que tienen historias

genéticas más similares entre sí y pueden representar grupos de aislamiento geográfico (al menos, desde hace dos generaciones), puesto que sólo se utilizaron los datos de las abuelas maternas. Tal es el caso de los tres grupos de estados encontrados, que pudieron haberse hallado aislados geográficamente (Fig. 3.4).

Es probable que en la época que vivieron las abuelas maternas de los donadores, hubiera una barrera geográfica que aislara dichos grupos estatales, pero que pudo haberse diluido en el presente debido a los cambios sociales y económicos actuales. El análisis de variación de las frecuencias de los polimorfismos nucleares da un indicio de que tales aislamientos pudieron haber disminuido el flujo génico entre las regiones y, por consiguiente, otros procesos evolutivos actuaron diferencialmente entre ellos para moldear sus frecuencias (Fig. 3.5). De tal forma, la población de los estados del Centro parece ser genéticamente muy similares a la del Bajío; mientras que la población de estados del Sureste pudo haber enfrentado alguna barrera geográfica o cultural que los aisló más de los estados vecinos.

Lo anterior se puede observar en la figura 3.4, donde cualquier agrupamiento representa estados que comparten en mayor medida historial genético de modo que los grupos denominados como de Centro y Bajío se ubican muy cerca en la gráfica. No así el grupo del Sureste que claramente está bastante alejado del resto, lo cual, se explica dada la orografía del lugar y el aislamiento lingüístico y cultural que tienen respecto al resto del país. Es importante recalcar que tal aislamiento pudo haber sido eliminado o, al menos, disminuido; sin embargo, sus pobladores aún conservan características genéticas que los diferencian de los de otras regiones. El DF y el Estado de México son los estados que han tenido mayor intercambio genético con otros estados y sus distancias genéticas entre ellos son muy bajas o nulas, de manera que no presentaron una distribución específica en el análisis.

La historia de aislamiento de las regiones del país evidencia que los procesos evolutivos actuaron independientemente en cada región; por lo que las variaciones en las frecuencias polimórficas actuales de los grupos estatales aislados en el pasado dan cuenta de la presencia de subestructuras poblacionales. Las frecuencias obtenidas en este estudio no representan las correctas para cada una de las regiones, ya que el muestreo y el tamaño de la población estudiada no son los adecuados para este último análisis; sin embargo, sí indican que hay diferencias entre ellas, así como tendencias de variación.

Se observa una clara variación entre las frecuencias de la GST M1 en el grupo del Sureste respecto a los grupos del Centro y el Bajío. Esto significa que esa región ha tenido una historia genética diferente al resto del país, lo que apoya la asociación GSTM1 \*0-haplotipos africanos antes explicada. Sería importante ampliar los estudios de esta región para corroborar dicha

asociación. Respecto a los estados del norte, es importante mencionar que no se les incluyó en el análisis debido a que no se tenían datos de abuelas maternas originarias de esta zona del país.

El muestreo no fue diseñado para este tipo de análisis, de manera que no se tienen representantes de todos los estados de la República y no se pudieron incluir en la correlación. Esto no significa que no aporten diferencias; por el contrario, podrían presentar mayor variación respecto a otras zonas del país dada la distancia geográfica y las diferentes aportaciones genéticas –tal como ya ha sido reportado (Bonilla *et al.*, 2005). Con el fin de tener un panorama general de la población mexicana, sería necesario ampliar el análisis a la totalidad del territorio.

## 5. CONCLUSIONES

De la investigación expuesta en este trabajo, se derivan las siguientes conclusiones:

- La población mestiza mexicana analizada ha sufrido altos flujos de migración procedentes de individuos de la mayor parte del país y de otros países.
- Los marcadores nucleares CYP1A1\*2C, \*4 y \*2A están en equilibrio de Hardy-Weinberg. Sus frecuencias alélicas en la población fueron: 0.503, 0.04 y 0.52, respectivamente.
- Los marcadores nucleares GST M1\*0 y GST T1\*0 presentan una frecuencia en la población analizada de: 0.36 y 0.081, respectivamente.
- Los polimorfismos CYP1A1\*2C, \*2A y T1\*0 presentaron diferencias significativas con las reportadas por otros autores en estudios previos para poblaciones mexicanas mestizas.
- La población mestiza mexicana está conformada por una vasta mezcla de genotipos de ancestría variada, tal como lo evidencia la presencia de haplotipos de todos los continentes. Pese a esto se confirma la hipótesis propuesta, puesto que la población analizada presenta ancestría primordialmente amerindia (~66.23%) con base en marcadores de ancestría de ADNmt. Tampoco es despreciable la contribución asiática (~20.97%); el resto corresponde a contribuciones bajas de individuos con mezclas recientes, ancestría africana y ancestría europea (~7.4%, ~2.7% y ~1.35%, respectivamente).
- Podría haber asociación entre el polimorfismo GST M1\*0 y los haplotipos africanos ASI/AFR/AFR2 y ASI/AFR.
- En la población mestiza mexicana no existe correlación entre las distancias geográficas y nucleotídicas; no obstante es posible identificar regiones de aislamiento que pueden tener historias genéticas diferentes.

## BIBLIOGRAFÍA

- ABDEL-RAHMAN, S., El-Zein, R., Anwar, W., Au, W. 1996. "A multiplex PCR procedure for polymorphic analysis of GSTM1 and GSTT1 genes in population studies". *Cancer Letters*. 107 (2): 229-33.
- ALBERTS, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J. 2002. *Biología Molecular de la célula*. Ediciones Omega. 1385 pp.
- ANZENBACHER, P., Anzenbacherová, E. 2001. "Cytochromes P450 and metabolism of xenobiotics". *Cellular and molecular life sciences*. 58 (5-6): 737-47.
- BONILLA, C., Gutiérrez, G., Parra, E., Kline, C., Shriver, M. 2005. "Admixture analysis of a rural population of the state of Guerrero, Mexico". *American Journal of Physical Anthropology*. 128 (4): 861-9.
- BONILLA, C., Parra, E., Ptaff, C., Dios, S., Marshall, J., Hamman, R., Ferrell, R., Hoggart, C., McKeigue, P., Shriver, M. 2004. "Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping". *Annals of Human Genetics*. 68 (2): 139-53.
- BUENTELLO-MALO, L., Peñaloza-Espinosa, R., Loeza, F., Salamanca-Gómez, F., Cerda-Flores, R. 2003. "Genetic structure of seven Mexican indigenous populations based on five polymarker loci". *American Journal of Human Biology*. 15 (1): 23-8.
- CASCORBI, I., Brockmoller, J., Roots, I. 1996. "A C4887A polymorphism in exon 7 of human CYP1A1: population frequency, mutation linkages, and impact on lung cancer susceptibility". *Cancer Research*. 56 (21): 4965-9.
- CAVALLI-SFORZA, L., Menozzi, P., Piazza, A. 1994. *The history and geography of human genes*. Princeton University Press. 1088 pp.
- CERDA-FLORES, R., Villalobos-Torres, M., Barrera-Saldaña, H., Cortés-Prieto, L., Barajas, L., Rivas, F., Cariacedo, A., Zhong, Y., Barton, S., Chakraborty, R. 2002. "Genetic admixture in three Mexican mestizo populations based on D1S80 and HLA-DQA1 loci". *American Journal of Human Biology*. 14 (2): 257-63.

- CERDA-FLORES, R., Budowle, B., Jin, L., Barton, S., Deka, R., Chakraborty, R. 2002b. "Maximum likelihood estimates of admixture in northeastern Mexico using 13 short tandem repeat loci". *American Journal of Human Biology*. 14 (4): 429-39.
- CRAWFORD, M., Leyshon, W., Brown, K., Lees, F., Taylor, L. 1974. "Human biology in Mexico. II. A comparison of blood group, serum and red enzyme frequencies, and genetic distances of the Indian populations of Mexico". *American Journal of Physical Anthropology*. 41 (2): 251-68.
- CRAWFORD, M., Lisker, R., Pérez, R. 1976. "Genetic microdifferentiation of two transplanted Tlaxcaltecan populations". En: Crawford M., editor. *The Tlaxcaltecs: prehistory, demography, morphology and genetics*. Publications in anthropology 7. Lawrence, KS: University of Kansas. 169-175 pp.
- DALY, A., Steen, V., Fairbrother, K., Idle, J. 1996. "CYP2D6 Multiallelism". *Methods of enzymology*. 272:199-210.
- FORD, J., Li, Y., O'Sullivan, M., Demopoulos, R., Garte, S., Taioli, E., Brandt-Rauf, P. 2000. "Glutathione S-transferase M1 polymorphism and lung cancer risk in African-Americans". *Carcinogenesis*. 21 (11): 1971-75.
- FRAGOSO, J., Juárez-Cedillo, T., Hernández-Pacheco, G., Ramírez, E., Zuñiga, J., Izaguirre, R., De la Peña, A., Granados, J., Vargas-Alarcón, G. 2005. "Cytochrome P4501A1 polymorphisms in the amerindian and mestizo populations of Mexico". *Cell Biochemical Function*. 23 (3): 189-93.
- FUTUYMA, D. 1998. *Evolutionary Biology*. Editorial Sinauer. 763 pp.
- GARDNER, E., Simmons, M., Snustad, P. 2003. *Principios de genética*. Editorial Limusa Wiley. 648 pp.
- GARTE S., Gaspari, L., Alexandrie, A., Ambrosone, C., Autrup, H., Autrup, J., Baranova, H., Bathum, L., Benhamou, S., Boffetta, P., Bouchardy, C., Breskvar, K., Brockmoller, J., Cascorbi, I., Clapper, M., Coutelle, C., Daly, A., Dell'Omo, M., Dolzan, V., Dresler, C., Fryer, A., Haugen, A., Hein, D., Hildesheim, A., Hirvonen, A., Hsieh, L., Ingelman-Sundberg, M., Kalina, I., Kang, D., Kihara, M., Kiyohara, C., Kremers, P., Lazarus, P., Le Marchand, L., Lechner, M., van Lieshout, E., London, S., Manni, J., Maugard, C., Morita, S., Nazar-Stewart, V., Noda, K., Oda, Y., Parl, FF., Pastorelli, R., Persson, I., Peters, W., Rannug, A., Rebbeck, T., Risch, A., Roelandt, L., Romkes, M., Ryberg, D., Salagovic, J., Schoket, B., Seidegard, J.,

- Shields, P., Sim, E., Sinnet, D., Strange, R., Stücker, I., Sugimura, H., To-Figueras, J., Vineis, P., Yu, M., Taioli, E. 2001. "Metabolic gene polymorphisms frequencies in control population". *Cancer Epidemiology Biomarkers Preview*. 10 (12): 1239-48.
- HAYASHI, S., Watanabe, J., Nakachi, K., Kawajiri, K. 1991a. "PCR detection of an A/G polymorphism within exon 7 of the CYP1A1 gene". *Nucleic Acids Research*. 19 (17): 4797.
- HAYASHI, S., Watanabe, J., Nakachi, K., Kawajiri, K. 1991b. "Genetic Linkage of lung cancer-associated *MspI* polymorphism with aminoacid replacement in the haeme binding region of the cytochrome P450 IAI gene". *Journal of biochemistry*. 110 (3): 407-11.
- HEDRICK, P. 2000. *Genetics of populations*. Jones and Barlett Publishers. 553 pp.
- HORAI, S., Kondo, R., Nakagawa-Hattori, Y., Hayashi, S., Sonoda, S., Tarima, K. 1993. "Peopling of the Americas, founded by four major lineages of mitochondrial DNA". *Molecular Biologic Evolution*. 10 (1): 23-47.
- IOANNIDIS, J., Ntzani, E., Trikalinos, T. 2004. "'Racial' differences in genetic effects for complex diseases". *Nature genetics*. 36 (12): 1312-18.
- JOBLING, M., Hurles, M., Tyler-Smith, C. 2004. *Human Evolutionary Genetics. Origins, Peoples and disease*. Garland Science, India. 515 pp.
- KAESSMANN, H., Pääbo, S. 2002. "The genetical history of humans and the great apes". *Journal of Internal Medicine*. 251 (1): 1-18.
- KAWAJIRI, K. 1999. "CYP1A1 in Metabolic polymorphisms and susceptibility to cancer", en Ryder, W. (ed.). *IARC Scientific publications*. 148: 159-172.
- KEITA, S., Kittles, R., Royal, C., Bonney, G., Furbert-Harris, P., Dunston, G., Rotimi, C. 2004. "Conceptualizing human variation". *Nature Genetics*. 36 (11): 17-20.
- LEWIN, B. 1997. *Genes VI*. Editorial Oxford University Press. 1260 pp.
- LISKER, R. 1981. *Estructura genética de la población mexicana. Aspectos médicos y antropológicos*. Salvat. 158 pp.

- LISKER, R., Babinsky, V. 1986. "Admixture estimates in nine Mexican Indian group and five east coast localities". *Revista de Investigación Clínica*. 38 (2): 145-9.
- LISKER, R., Pérez, R., Granados, J., Babinsky, V., Rubens, J., Arrendares, S., Buentello, L., 1986. "Gene frequencies and admixture estimates in a Mexico city population". *American Journal of physical anthropology*. 71 (2): 203-7.
- LISKER, R., Pérez, R., Granados, J., Babinsky, V. 1988. "Gene frequencies and admixture estimates in the state of Puebla, Mexico". *American Journal of physical Anthropology* . 76 (3): 331-5.
- LISKER, R., Ramírez, E., González-Villalpando, C., Stern, M. 1995. "Racial admixture in a Mestizo population from Mexico city". *American Journal of Human Biology*. 7: 213-6.
- LISKER, R., Ramírez, E., Babinsky, V. 1996. "Genetic structure of autochthonous populations of Meso-America: Mexico". *Human Biology*. 68 (3): 395-404.
- PEMBLE, S., Schroeder, K., Spencer, S., Meyer, D., Hallier, E., Bolt, H., Ketterer, B., Taylor, J. 1994. "Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism". *Biochemical Journal*. 300 (1): 271-6.
- PÉREZ, R. 2006. *Frecuencia del polimorfismo CYP1A1\*2C y GSTM1\*0 en población mexicana y su posible asociación con la susceptibilidad a padecer cáncer pulmonar*. Instituto de Investigaciones Biomédicas, UNAM. [Tesis de Maestría en Ciencias Biológicas].
- PRITCHARD, J. Stephens, M., Donnelly, P. 2000. "Inference of population structure using multilocus genotype data". *Genetics*. 155 (2): 945-59.
- REBBECK, T., Sankar, P. 2005. "Ethnicity, ancestry, and race in molecular epidemiologic research". *Cancer Epidemiology Biomarkers Preview*. 14 (11): 2467-71.
- SALZANO, F. 2002. "Molecular variability in Amerindians: widespread but uneven information". *Anais da Academia Brasileira de Ciências*. 74 (2): 223-63.
- SILVA, W., Boniato, S., Holanda, A., Ribeiro-dos-Santos, A., Paixão, B., Goldman, G., Abe-Sandes, K., Rodríguez-Delfín, L., Barbosa, M., Pacó-Larson, M., Petzl-Erler, M., Valente, V. Santos, S., Zago, M. 2002. "Mitochondrial genome diversity of native americans supports a single

early entry of founder populations into America". *American Journal Human Genetics*. 71 (1): 187-92.

STRANGE, R., Fryer, A. 1999. "The glutathione S-transferases: influence of polymorphism on cancer susceptibility in Metabolic polymorphisms and susceptibility to cancer", en Ryder, W. (ed.). *IARC Scientific publications*. 148: 231-243.

TIBURCIO, V., Romero, A., De Garay, A. 1978. "Gene frequencies and racial admixture in a Mestizo population from Mexico city". *Annals of Human Biology*. 5 (2): 131-6.

TORRONI, A., Brown, M., Lott, M., Newman, N., Wallace, D., Cuba Neuropathy Field Investigation Team. 1995. "African, native American, and European mitochondrial DNAs in Cubans from Pinar del Rio province and implications for the recent epidemic neuropathy in Cuba". *Human mutation*. 5 (4): 310-7.

## APÉNDICE I

### I. Purificación de ADN (según Daly *et al.*, 1996)

El ADN fue purificado a partir de 5 ml de sangre entera. A cada muestra se le agregó 45 ml de buffer de lisis (320 mM sacarosa, 5 mM MgCl<sub>2</sub>, 1% Tritón X-100, 10 mM Tris-HCl, pH 7.4). Después de agitar bien para romper las células, el lisado se centrifugó a 2000 g por 5 minutos (min) a 4 °C. La pastilla se resuspendió en 2 ml de buffer de suspensión (150 mM NaCl, 60 mM EDTA, 1% SDS, 400 mM Tris-HCl, pH 7.4) y 0.5 ml de perclorato de sodio 5 M.

Se mantuvo la suspensión en rotación a temperatura ambiente por 15 min y luego se incubó a 65 °C por 30 min. Al terminar, se agregaron 2 ml de cloroformo a -20 °C y se mezclaron por rotación a temperatura ambiente por 10 min. Luego, se volvió a centrifugar a 1400 g por 10 min. La fase acuosa superior que contiene al ADN fue transferida a un tubo nuevo y se le agregaron 2 volúmenes de etanol absoluto a 4 °C. El ADN fue precipitado por inversión y la hebra de ADN fue recuperada en un tubo *ependorf*.

El ADN recuperado se lavó con etanol al 70% y se secó en un *DNA 110 Speed Vac* a una velocidad media. El ADN fue resuspendido en 200 µl de TE (10 mM Tris-HCl, 1 mM EDTA, pH 7.4), incubando a 60 °C toda la noche. El ADN fue almacenado a 4 °C hasta su utilización.

### II. Reacción en cadena de la polimerasa (*Polymerase Chain Reaction*). PCR y PCR múltiple

La técnica de PCR permite amplificar muchas copias de genes o de fragmentos de genes individuales. Para flanquear la región de interés, se utilizan oligonucleótidos (secuencias cortas de nucleótidos) sintetizados. La reacción se lleva a cabo en un termociclador, aparato que ejecuta ciclos de diferentes temperaturas y duración, durante los cuales, el templado de ADN se desnaturaliza a alta temperatura, se alinean los oligonucleótidos en su secuencia específica a temperaturas específicas para cada par de oligonucleótidos, y se sintetizan las nuevas copias (amplicones) de los genes a temperatura alta, gracias a una polimerasa termoestable (Taq polimerasa). Estos ciclos se repiten cuantas veces sea necesario para obtener la cantidad de copias deseadas. Es posible comprobar que se obtuvo el amplicón deseado por medio de una electroforesis.

Como se mencionó, cada par de oligonucleótidos (uno para cada hebra de ADN) debe tener una temperatura de alineamiento específica, pero si los oligonucleótidos tienen las mismas

temperaturas de alineamiento, se pueden incluir en la misma reacción y así amplificar varios genes o fragmentos génicos al mismo tiempo. A esta variante del PCR se le llama “PCR múltiple”.

Para todos los experimentos se usó la siguiente mezcla de reacción para preparar un volumen total de 10 µl: 1 µl de ADN (~100 ng), 1 µl de dNTP's (2mM), 1 µl de oligonucleótidos (5mM), 0.1 µl de Taq polimerasa, 1 µl de buffer 20 mM MgCl<sub>2</sub>, y 5.9 µl H<sub>2</sub>O. Las condiciones de amplificación varían para cada polimorfismo y se detallan en la metodología anexa a este trabajo.

### **III. Polimorfismos de longitud de fragmentos de restricción. RFLP (*Restriction Fragment Length Polymorphisms*)**

Los RFLP son variaciones de las longitudes de los fragmentos de ADN producidos por el corte del mismo por medio de enzimas de restricción. Las enzimas de restricción se unen a secuencias específicas del ADN y cortan en esos sitios, pero cuando existe un cambio en la secuencia, las enzimas no reconocen el sitio por lo que no se unen y no cortan. Tal es el caso de los polimorfismos, en los que hay sustituciones, deleciones o adiciones de pares de bases, de modo que es posible identificarlos usando esta técnica. Algunas veces, el corte se produce cuando está presente el polimorfismo; otras, cuando está ausente, dependiendo de la enzima que se utilice. Estos RFLP pueden observarse cuando los fragmentos productos de la digestión se separan por electroforesis.

Es posible identificar heterocigos y homocigos según el patrón de bandeo. Un individuo heterocigo tendrá un alelo con el polimorfismo (genotipo mutante), y otro sin el polimorfismo (genotipo silvestre o ancestral); por lo tanto, un alelo es cortado y otro no. En un individuo homocigo, ya sea para el genotipo mutante o para el silvestre, ambos alelos son reconocidos de la misma forma por la enzima de restricción; es decir, o bien ambos se cortan, o ninguno de los dos lo hace. Saber la longitud del fragmento total de ADN y las longitudes de los fragmentos esperados, permitirá identificar heterocigos y diferenciar entre homocigos.

Para las restricciones se usó la siguiente mezcla de reacción: 2 µl de producto de PCR, 1 unidad de enzima, completando con buffer adecuado a un volumen final de 8µl. Las condiciones de restricción varían para cada enzima y se detallan en la metodología anexa a este trabajo.