

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS
Departamento de Zoología y Antropología Física



**ATLAS TOPOGENÉTICO DE GRUPOS
INDÍGENAS MESOAMERICANOS: UNA
APROXIMACIÓN MOLECULAR.**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

Amaya Gorostiza Langa

Bajo la dirección de los doctores

Antonio González-Martín
Sonsoles Rueda González

Madrid, 2011

ISBN: 978-84-694-4161-9

© Amaya Gorostiza Langa, 2011

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE BIOLOGÍA
DEPARTAMENTO DE ZOOLOGÍA Y ANTROPOLOGÍA FÍSICA



**ATLAS TOPOGENÉTICO DE GRUPOS INDÍGENAS
MESOAMERICANOS:
UNA APROXIMACIÓN MOLECULAR**

AMAYA GOROSTIZA LANGA

Otoño 2010

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE BIOLOGÍA
DEPARTAMENTO DE ZOOLOGÍA Y ANTROPOLOGÍA FÍSICA

**ATLAS TOPOGENÉTICO DE GRUPOS INDÍGENAS
MESOAMERICANOS:
UNA APROXIMACIÓN MOLECULAR**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

Presentada por

AMAYA GOROSTIZA LANGA

Bajo la dirección de los doctores

Antonio González-Martín
Dpto. Zoología y Antropología Física
Facultad de Biología
Universidad Complutense de Madrid

Sonsoles Rueda González
Departamento Técnico
Genómica S.A.U.
Grupo Zeltia

Otoño 2010

SONSOLES RUEDA GONZÁLEZ, DIRECTORA TÉCNICA DE GENOMICA S.A.U., EMPRESA DEL GRUPO ZELTIA, Y **ANTONIO GONZÁLEZ-MARTÍN**, PROFESOR CONTRATADO DOCTOR DEL DEPARTAMENTO DE ZOOLOGÍA Y ANTROPOLOGÍA FÍSICA DE LA FACULTAD DE BIOLOGÍA DE LA UNIVERSIDAD COMPLUTENSE DE MADRID,

CERTIFICAN QUE:

La Memoria titulada **“Atlas topogenético de grupos indígenas mesoamericanos: Una aproximación molecular”** que presenta **Amaya Gorostiza Langa** para optar al Título de Doctora en Biología, ha sido realizada en Genomica S.A.U. empresa del grupo Zeltia, y en el Departamento de Zoología y Antropología Física de la Facultad de Biología de la Universidad Complutense de Madrid bajo nuestra dirección. Autorizamos su presentación a la Junta de Facultad al considerar que constituye un trabajo de tesis.

Para que así conste, firmamos este certificado en Madrid a 30 de septiembre de 2010.

Fdo. Antonio González-Martín

Fdo. Sonsoles Rueda González

A toda mi familia, aquellos con los que comparto genes y aquellos con los que no.

A los que están y sobre todo a los que se fueron.

A Antonio.

Una tarde que estaba la virgen
cosiendo y cantando sentaíta al sol,
llamó al niño y el niño no estaba
y atribulaíta la virgen queó.
Pero apareció que en el templo
a doctores y obispos
les enseñaba el dogma de la religión.

Hasta Herodes, Campanilleras

Pepe Menese

Estaba próxima la Pascua de los judíos, y subió Jesús a Jerusalén. Encontró en el templo a los vendedores de bueyes, de ovejas y de palomas, y a los cambistas sentados; y haciendo de cuerdas un azote, los arrojó a todos del templo, con las ovejas y los bueyes, derramó el dinero de los cambistas y derribó las mesas; y a los que vendían palomas les dijo: Quitad de aquí todo eso y no hagáis de la casa de mi Padre casa de contratación.

Juan 2,13-22

A los directores de esta tesis. A la Dra. Sonsoles Rueda por su paciencia y su apoyo. Al Dr. Antonio González-Martín, por sus comentarios y observaciones, por su apoyo incondicional y su confianza. A ambos por tener fe en mí, muchas veces más que yo misma.

A todo el equipo directivo de GENOMICA SAU, en especial a Rosario Cospedal, por su apoyo y su confianza.

A todos mis compañeros de GENOMICA SAU, en especial a Chon, Nuria, Laura e Irene, que me han animado, apoyado y consolado siempre que ha sido necesario.

A todos los compañeros del Departamento de Zoología y de Antropología Física de la Facultad de Biología de la Universidad Complutense de Madrid, que me han apoyado y que siempre me han recibido con una sonrisa. En especial a Jorge, que compartió conmigo las múltiples dimensiones de la estadística.

Al Departamento de Genética de la Facultad de Biología de la Universidad Complutense de Madrid que me apoyó cuando fue necesario.

A la Universidad Autónoma del Estado de Hidalgo (UAEH), a la Escuela Nacional de Antropología e Historia (ENAH), y al Instituto Mexicano del Seguro Social (IMSS) de Hidalgo, que apoyaron los distintos proyectos de los que ha surgido esta tesis.

A algunas personas, su falta de respeto y las trabas que han puesto al desarrollo de esta tesis. Su mezquindad ha alimentado mi fuerza y mi pasión hasta el final de este trabajo.

A todos y cada uno de los miembros de mi familia mexicana: los Acuña-Alonzo incluyendo los no bípedos, el Pollo, Mirsha, Paola, Celta, Ana, Juan Carlos, Julio Granados, Héctor Rangel, David Sámano y el Padre David, Cherif, Flo e Inés. Con todos ellos comparto un camino más allá de los cánones occidentales que no ha hecho más que empezar.

A Víctor Ballesteros que me animó a hacer esta tesis aunque nunca sabrá de ella.

A las comunidades indígenas mexicanas, por su generosidad, curiosidad y por haberme permitido conocer otras realidades. Lo que se refleja en esta tesis es una ínfima parte de lo que compartí y aprendí de ellas.

A todo el equipo de Ángel Carracedo de la Universidad de Santiago de Compostela, en especial a él mismo, a Paula y a María.

A la familia de *Falco tinnunculus* que habita la fachada norte de la Facultad de Biología, por su compañía.

A Carlos Martín y Silvia Gutiérrez, por su amistad y su apoyo con el instrumental.

A Madrid, mi casa, por haberme acogido. A Barcelona, mi hogar, por haberme dejado marchar.

Al Bar Los Toreros de Barcelona, donde nació esta tesis. Y al Bar Los Morales de Madrid, Los Caracoles, donde maduré.

...a Antonio, mi luz del alma. Mía divina. Que a mí me alumbró mi corazón.

ÍNDICE

Objetivos.....	1
I. Introducción.....	5
I.1. Antecedentes.....	7
I.2. La información molecular como herramienta.....	8
I.2.I. Características generales de las mitocondrias	
I.2.II. Características del ADNmt humano	
I.2.III. Genes mitocondriales codificantes.	
I.2.IV. Región no codificante o <i>D-Loop</i>	
I.2.V. Tasa de mutación	
I.2.VI. Herencia materna	
I.2.VII. Ausencia de recombinación	
I.2.VIII. Polimorfismos y heteroplasmia en el ADNmt	
I.3. Aplicaciones del ADNmt.....	20
I.3.I. Primeros estudios	
I.3.II. Relaciones genealógicas entre primates y <i>Homo sapiens</i>	
I.3.III. Utilidad forense: la familia Romanov	
I.3.IV. Dispersión humana y origen geográfico	
I.3. V. Uso potencial del ADNmt en biomedicina	
I.4. Expansión de <i>Homo sapiens</i>	25
I.4.I. África, el origen	
I.4.II. El éxodo africano	
I.4.III. El sudeste asiático y Australia	
I.4.IV. El centro del continente asiático	
I.4.V. Europa occidental	
I.4.VI. La convivencia con Neandertales	
I.4.VII. Consideraciones climatológicas	
I.4.VIII. La velocidad de expansión	
I.5. El poblamiento de América.....	39
I.5.I. La cultura Clovis: presencia humana en el Nuevo Mundo	
I.5.II. Información osteológica	
I.5.III. Familias lingüísticas y poblamiento	
I.5.IV. Estudios y evidencias moleculares del poblamiento de América	

I.6. Regiones culturales: Oasisamérica, Aridoamérica y Mesoamérica.....	55
I.6.I. Occidente de México	
I.6.II. Norte	
I.6.III. Centro de México	
I.6.IV. Golfo de México	
I.6.V. Guerrero	
I.6.VI. Oaxaca	
I.6.VII. Costa Sur	
I.6.VIII. Centroamérica	
I.6.IX. Maya	
I.7. Cronología del México antiguo.....	63
I.7.I. Período Preclásico temprano	
I.7.II. Período Preclásico medio.	
I.7.III. Período Preclásico tardío	
I.7.IV. Período Clásico temprano.	
I.7.V. Período Clásico tardío.	
I.7.VI. Período Posclásico temprano.	
I.7.VII. Período Posclásico tardío.	
I.7.VIII. Período colonial	
I.8. Clasificaciones de las lenguas indígenas.....	74
I.9. Estudios moleculares en la región de estudio.....	82
I.10. Zonas y grupos indígenas estudiados.....	87
I.10.I. Pimas	
I.10.II. Mayos	
I.10.III. Huicholes	
I.10.IV. Nahuas	
I.10.V. Otomíes	
I.10.VI. Tepehuas	
I.10.VII. Mayas	
II. Material y métodos.....	99
II.1. Muestreo de las poblaciones.....	101
II.2. Planificación del muestreo.....	101
II.3. Estrategias de muestreo.....	102
II.4. Toma de muestras en las comunidades indígenas.....	103
II.4.I. Pimas	
II.4.II. Mayos	

II.4.III. Huicholes	
II.4.IV. Nahuas	
II.4.V. Otomíes	
II.4.VI. Tepehuas	
II.4.VII. Mayas	
II.5. Extracción y cuantificación de ADN.....	108
II.5.I. Sangre	
II.5.II. Pelo con raíz mediante Chelex®	
II.5.III. Pelo con raíz mediante ChargeSwitch Kit (Invitrogen)	
II.5.IV. Epitelio bucal	
II.6. PCR: Amplificación de la región control o D-Loop.....	114
II.7. Electroforesis en geles de agarosa.....	125
II.8. Purificación de amplificadores y reacción de secuenciación.....	127
II.9. Purificación de secuencias y análisis.....	130
II.10. Alineamiento de secuencias y determinación del haplogrupo.....	131
II.11. <i>Software</i> y determinaciones estadísticas.....	134
III. Resultados.....	139
III.1. Consideraciones metodológicas.....	141
III.2. Resultados D-Loop para población total.....	144
III.3. Resultados HVI y HVII para población total.....	152
III.4. Estudios poblacionales mediante D-Loop.....	153
III.5. Estudios poblacionales mediante HVI y HVII.....	160
III.6. Análisis intrapoblacionales: subestructura interna.....	164
III.6.I. Otomíes del Valle vs Otomíes de la Sierra	
III.6.II. Nahuas de la Huasteca	
III.6.III. Estado de Hidalgo	
III.7. Análisis interpoblacionales.....	194
III.8. <i>Networks</i>	204
III.9. Barreras geográficas vs barreras genéticas.....	205
IV. Discusión.....	219
IV.1. Consideraciones sobre el muestreo.....	221
IV.2. Consideraciones sobre las muestras.....	223
IV.3. Haplogrupos y haplotipos detectados.....	225
IV.4. Análisis por grupos indígenas.....	234
IV.5. Otomíes del Valle vs Otomíes de la Sierra.....	241

IV.6. Relaciones internas entre comunidades Nahuas.....	245
IV.7. Relaciones entre los grupos de Hidalgo.....	250
IV.8. Análisis por grupos indígenas en el contexto americano.....	254
V. Conclusiones.....	259
VI. Bibliografía.....	265
VII. Anexos.....	301

OBJETIVOS

Los objetivos que se pretenden conseguir con el desarrollo de la presente tesis doctoral son los siguientes:

- Diseñar una estrategia óptima de muestreo que permita la obtención de muestras en las comunidades indígenas.
- Poner a punto los protocolos de laboratorio para proceder al análisis optimizando cada una de sus fases (extracción, amplificación, secuenciación).
- Secuenciar la región control o D-Loop de todas las muestras.
- Caracterizar genéticamente grupos indígenas inéditos desde el punto de vista molecular.
- Caracterizar genéticamente grupos indígenas en proceso de extinción.
- Estudiar pormenorizadamente las comunidades indígenas de una región geográfica concreta.
- Hacer una aproximación a la historia de las poblaciones de Mesoamérica.
- Hacer una valoración del efecto que geografía, lengua e historia tienen en la subestructuración genética de las poblaciones humanas.
- Determinar la relación entre Mesoamérica y las regiones geográficas cercanas.
- Detectar y cuantificar las posibles barreras en el ámbito estudiado.

I. INTRODUCCIÓN

I.1. Antecedentes

El *Homo sapiens* es una especie relativamente reciente con baja diversidad genética (Chakravarti 1999; Jorde et al., 2001; Przeworski et al., 2000) y estructurada geográficamente; el 85% de la diversidad humana se encuentra dentro de las poblaciones, un 10% se detecta entre los grandes grupos continentales y el 5% restante procede de las poblaciones del mismo continente (Lewontin 1972; Livshits and Nei, 1990; Relethford and Harpending, 1994; Batzer et al., 1994; Jorde et al., 1997; Barbujani 1997; Stoneking 1997). Estas características implican que las diferencias genéticas entre las poblaciones humanas sobre todo si las comparamos con otras especies, son relativamente pequeñas. De cualquier manera son lo suficientemente grandes como para ser detectadas y estudiadas, sobre todo si se conocen los mecanismos de herencia y el efecto que las fuerzas evolutivas han tenido sobre ellas. Aplicando las herramientas moleculares adecuadas y los *software* estadísticos específicos, a partir de la diversidad genética de grupos humanos actuales pueden reconstruirse aspectos de su pasado como su origen, los procesos a los que han estado sometidos, la relación biológica con otros grupos, etc. En definitiva, a partir de la información genética y aplicando los métodos adecuados, se puede inferir en el pasado de nuestra especie y en la historia de grupos humanos concretos.

La presente tesis doctoral se centra en la utilización de la información genética, concretamente del ADN mitocondrial (ADNmt), procedente de grupos indígenas de Mesoamérica para inferir en su pasado. Para ello y al margen de la información molecular, también se ha utilizado la información procedente de otras disciplinas muy distantes en su metodología para crear un marco cultural e histórico en el que interpretar la información molecular. En ocasiones será preciso profundizar en aspectos históricos muy concretos, como la clasificación cronológica de Mesoamérica, o diferenciar las implicaciones demográficas que tiene para un grupo humano la utilización de una estrategia de subsistencia determinada. Toda esta información permitirá interpretar de una forma plausible la información molecular, permitiendo aproximarnos al pasado de los grupos humanos que habitan una de las regiones más interesantes del continente americano.

Los grupos humanos estudiados son excepcionales por muchos motivos, desde un punto de vista antropológico. Por una parte, algunos de ellos son los herederos de las culturas más relevantes del continente americano; por otra, muchos están en proceso de desaparición, tanto cultural como biológicamente. Estos aspectos convierten a este material en único y de un valor insustituible en el campo de la antropología.

I.2. La información molecular como herramienta

El ADN humano admite múltiples clasificaciones, una de ellas basada en su tamaño y estructura, permite diferenciar un genoma complejo o nuclear (ADNn), y un genoma más sencillo o mitocondrial. El primero de ellos se encuentra en el núcleo celular organizado en 23 pares de cromosomas, y representa más del 99% del ADN contenido en una célula, con aproximadamente 3.200.000 pares de bases (pb).

El ADNn se encuentra en el núcleo celular y está organizado en estructuras independientes llamadas cromosomas, 23 pares en nuestra especie. Los cromosomas están constituidos por una doble hélice de ADN estructurada en torno a unas proteínas llamadas histonas.

El ADN celular es idéntico en todas las células excepto en las germinales. La diferencia más importante entre estos tipos de células es que las somáticas contienen 23 pares de cromosomas, mientras que las germinales contienen únicamente 23 cromosomas, la mitad.

La composición cromosómica de las células somáticas es pareada, es decir, existen dos dotaciones para cada cromosoma; cada uno de estos pares se conoce como cromosomas homólogos y proceden de nuestros progenitores, uno del masculino y otro del femenino. Células con esta doble dotación cromosómica y mediante el fenómeno de la meiosis, reducen su composición cromosómica a la mitad, produciendo células germinales constituidas por 23 cromosomas. La distribución de los cromosomas durante el proceso de reducción de 46 a 23 se realiza al azar mediante un proceso conocido como segregación cromosómica. La información contenida en una célula germinal según proceda de uno u otro progenitor, será un proceso probabilístico, el número de gametos que se pueden generar considerando exclusivamente el proceso de segregación cromosómica es de 2^{23} , es decir, más de 8.000.000 de gametos distintos.

La unión de dos células germinales formará, mediante la adición de sus dotaciones cromosómicas, una nueva célula compuesta por 23 pares de cromosomas, un cigoto que se duplicará, conservando su dotación cromosómica, mediante el proceso de mitosis. Durante el desarrollo embrionario las nuevas células, con idéntica dotación genética, se diferenciarán para formar cada uno de los órganos y estructuras.

Durante este proceso de reducción y duplicación cromosómica hay un fenómeno que debe tenerse en cuenta para aplicar la información molecular en la reconstrucción de la historia de nuestra especie. Se trata de la *recombinación genética*, basada en el

intercambio de material genético entre cromosomas homólogos. En realidad, cualquier cromosoma está constituido por un mosaico de información procedente de los cromosomas homólogos de sus progenitores, este mecanismo evolutivo permite mantener una alta variabilidad genética entre las generaciones progenitoras y sus descendientes.

El ADNn admite otra clasificación, los autosomas o *cromosomas autosómicos*, son 22 pares y no están involucrados en la determinación del sexo, mientras que los *cromosomas sexuales* son los que definirán el sexo del individuo: cromosoma X y cromosoma Y. El primero lo presentan ambos sexos en diferentes proporciones, mientras que el segundo es exclusivo de los varones: un individuo masculino se describe como 44 + XY, es decir, 44 cromosomas autosómicos más dos cromosomas sexuales, X e Y, un individuo femenino será 44 + XX.

La segregación cromosómica y la recombinación genética son dos fenómenos especialmente destacables que han de tenerse en cuenta al aplicar la genética molecular en la reconstrucción de nuestro pasado. Una de las fases más importantes en este proceso de inferencia es asignar variantes genéticas a una región geográfica concreta, lo que permite reconstruir las migraciones que se han producido en el proceso de expansión de *Homo sapiens*. La presencia de una de estas variantes en un individuo permitirá hacer una inferencia en el origen de sus antecesores. Tanto la segregación cromosómica como la recombinación genética pueden enmascarar este proceso, ya que es imposible determinar de qué predecesor procede un marcador concreto determinado, representando un problema complejo cuando los dos progenitores proceden de diferentes regiones geográficas. De todas formas, en la actualidad se están desarrollando nuevas técnicas para determinar a partir del estudio de marcadores del ADNn el origen ancestral de los individuos, y por extrapolación de las poblaciones (Tian et al., 2008; Novembre et al., 2008; He et al., 2009).

Pero al margen de la complejidad y las posibilidades del ADNn existe otro genoma en el interior de unos orgánulos citoplasmáticos, el ADNmt. Este material por sus características excepcionales, que se comentarán más adelante, ha sido utilizado como herramienta para investigar el pasado, estudiar la evolución humana y reconstruir los patrones de poblamiento y migraciones. Es uno de los sistemas genéticos más explorados en Antropología Molecular.

I.2.I. Características generales de las mitocondrias

El ADNmt se localiza en los orgánulos citoplasmáticos llamados mitocondrias, que tienen la función de suministrar energía a la célula. Su carácter extranuclear así como otras de sus características pueden explicarse por su origen. Existen varias hipótesis sobre el origen de la mitocondria y del genoma que contiene, pero en la actualidad se acepta la Teoría de la Endosimbiosis postulada a finales del siglo pasado por Lynn Margulis (Sagan, 1967; Margulis 1981) sobre el origen de mitocondrias y cloroplastos.

Esta teoría propone el origen bacteriano de la mitocondria, aceptando que aproximadamente 1.500 millones de años antes del presente (AAP) una protobacteria se incluyó en una célula proto-eucariota, beneficiando a ambos organismos. Este proceso culminó con la inclusión definitiva de la bacteria como orgánulo citoplasmático en las células eucariotas. La aceptación de esta teoría explica la singular estructura del ADNmt, así como sus características y sistema de transmisión (Margulis and Sagan, 1987). Por ejemplo, los ribosomas mitocondriales tienen características muy similares a los bacterianos, tanto en su estructura como en su sensibilidad a antibióticos, la cicloheximida inhibe la síntesis proteica en todos los ribosomas del citoplasma en eucariontes, pero no afecta a los ribosomas mitocondriales.

En general, se considera que la mayoría de las proteínas sintetizadas a partir de ribosomas mitocondriales son producto del ADNmt. Sin embargo, la condición simbiótica a partir de la cual la mitocondria se originó como orgánulo celular, implica una estrecha relación con el núcleo. La subunidad 9 de la F1 ATPasa es codificada por el ADNmt en organismos como la levadura, mientras que en otras especies de hongo como *Neurospora crassa* o los mamíferos, es codificada por el ADNn poniendo de manifiesto una compleja interacción entre ambos genomas.

Las mitocondrias son orgánulos de forma ovalada de 0,5 μm de diámetro por 2-5 μm de longitud. Presentan una membrana externa y otra interna extraordinariamente plegada, que contiene una matriz fluida. En el interior del orgánulo se desarrolla la fosforilación oxidativa, fuente primaria de la energía celular, así como una serie de ciclos necesarios para sintetizar el ATP (trifosfato de adenosina), molécula de intercambio energético celular. La producción de energía está basada en una red muy compleja de proteínas que actúan de forma secuencial hasta la síntesis del ATP. De hecho, durante la primera mitad del siglo XIX, los trabajos realizados sobre la mitocondria se centraban en su bioquímica y en la función de este orgánulo celular como estructura generadora de energía. .

I.2.II. Características del ADNmt humano

El ADNmt es una molécula de doble cadena circular compuesta por 16.569 pb cuya estructura le confiere mayor estabilidad frente a procesos degradativos. El ADNmt ha sido secuenciado en toda su longitud y la numeración de la molécula es la establecida por Anderson en 1981 (Anderson et al., 1981; Andrews et al., 1999). Su contenido en Guanina + Citosina (G+C) es del 41%. La composición de bases de las dos cadenas del ADNmt es significativamente distinta: hay una cadena pesada (H, del inglés *heavy*), rica en guaninas y una cadena ligera (L, del inglés *light*) que es rica en citosinas. Las purinas (adenina, A, y guanina, G), tienen un peso molecular mayor que las pirimidinas (citosina, C, y timina, T) y por esa razón una cadena es más pesada que la otra.

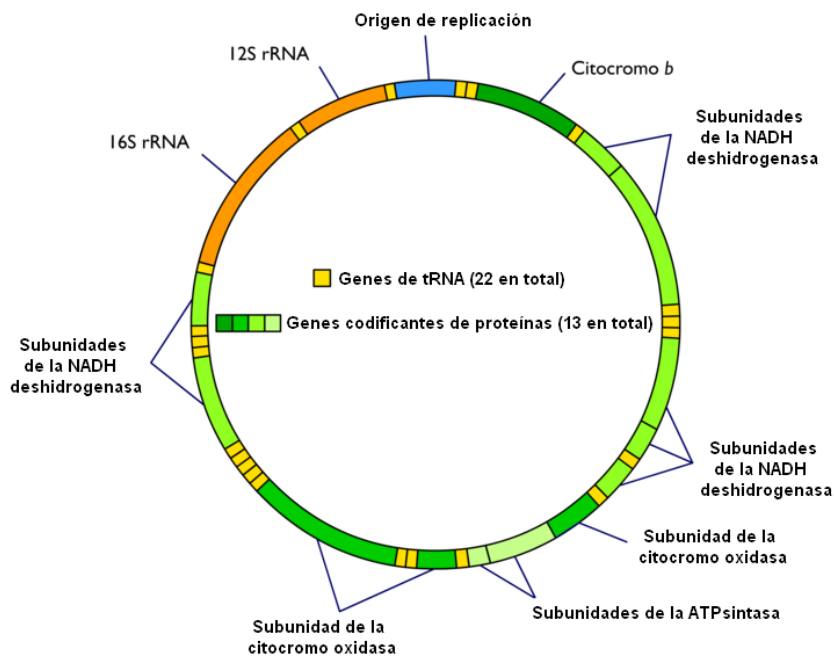


Figura I.1. Esquema del ADNmt humano. La zona señalada como origen de replicación corresponde a la región D-Loop o región no codificante que, a su vez, contiene las regiones hipervariable I (HVI) e hipervariable II (HVII). El resto del ADNmt representa la región codificante con diferentes funciones relacionadas con la producción celular de energía.

Cada célula contiene un número variable de mitocondrias, en función de la demanda energética, que oscila entre 10 y 10.000 copias. En una célula típica es de 2.000 lo que representa hasta el 0,5% del contenido de ADN total. Aunque la información contenida en el genoma mitocondrial es muy pequeña en comparación con el genoma nuclear, la cantidad de ADNmt es relativamente grande (Chen et al., 1995). Esta

característica es una ventaja para su manipulación en el laboratorio frente a otros genomas.

El ADNmt se puede dividir según su funcionalidad en dos regiones; una codificante, que representa el 93% del total, y una no codificante, conocida también como D-Loop (*displacement loop*) o Región Control (Figura I.1).

I.2.III. Genes mitocondriales codificantes

La región codificante está formada por 15.448 pb y contiene 37 genes, 28 de ellos codificados por la cadena H y 9 por la cadena L. La tecnología actual permite el estudio de los puntos polimórficos de esta región para discriminar el origen geográfico de la molécula. Casi una tercera parte de los genes (13) codifican para polipéptidos involucrados en el proceso de fosforilación oxidativa, el resto codifican para dos tipos de moléculas de ARN, 22 de transferencia (RNAt) y 2 ribosómicos (RNAr), 12S y 16S, formando parte de la síntesis proteica de la mitocondria.

Presenta marcadas diferencias respecto al ADNn ya que es extremadamente compacto y sus genes no presentan intrones, encontrándose un gen cada 0,45 Kb. Las secuencias de algunos genes presentan solapamientos y en la mayoría de los casos, las regiones codificantes de genes adyacentes o bien son contiguas o están separadas por una o dos bases no codificantes.

Los análisis de las secuencias ponen de manifiesto diferencias en el código genético mitocondrial respecto al nuclear, por ejemplo el codón UGA que significa terminación en el núcleo es leído como triptófano en la mitocondria humana. Lo mismo sucede con ATA, que codifica metionina en la mitocondria mientras que en el núcleo codifica por isoleucina.

La mitocondria contiene ribosomas que son responsables de toda la actividad de síntesis proteica que tiene lugar en el orgánulo, cuyas subunidades están codificadas en el mismo ADNmt. Sin embargo, las proteínas ribosómicas (entre 70 y 100) que forman parte de los ribosomas mitocondriales están codificadas en el genoma nuclear.

Otra característica específica del genoma mitocondrial es que el codón AUG de iniciación del RNA mensajero (RNAm) está en el extremo 5' de la molécula, y no tras una secuencia *leader* como sucede con los RNAm citoplasmáticos; además falta la estructura *cap 5'*. Este contraste nos indica que la síntesis proteica mitocondrial es bastante distinta de la nuclear, ya que el ribosoma y el mecanismo de iniciación

reconocen el extremo del RNAm sin ningún requerimiento para la unión a un punto del ribosoma o a una estructura 5'.

El hecho de no tener espaciadores entre transcritos de RNAm y RNAt, hace que tras el proceso de maduración se generen RNAt y RNAm monogénicos. Estos largos transcritos son sometidos a *splicing* y son reconocidos por su estructura característica.

Otro aspecto a destacar es que la secuencia de ADN para la mayoría de RNAm no codifica una cadena con codón de terminación. En lugar de eso, la replicación termina con U o UA en la misma pauta de lectura que el codón de inicio, y la poliadenilación en 3' característica de los RNAm parece suplir la parte perdida del codón de stop UAA.

I.2.IV. Región no codificante o *D-Loop*

La región control o *D-Loop* representa aproximadamente el 10% del ADNmt (Figura I.1). Contiene elementos reguladores de la replicación y transcripción del genoma mitocondrial, así como puntos de unión de varios factores de transcripción y secuencias relacionadas con el inicio y finalización de la replicación. El mayor grado de variación en el ADNmt en los individuos se encuentra en esta región, que está compuesta por 1.021 pb y está dividida a su vez en dos sub-regiones no consecutivas: HVI comprendida entre las posiciones 16.024 y 16.365, y HVII limitada por los nucleótidos 73 y 340 (Vigilant et al., 1989). Por ejemplo, las regiones HVI y HVII entre europeos se diferencian en una media de 8 posiciones (Budowle et al., 1999). Estas dos sub-regiones son de gran interés por dos de sus particularidades: no están sometidas a presión selectiva y tienen una alta tasa de mutación (μ). Ambas características hacen que presente una elevada diversidad nucleotídica. En general el ADNmt presenta valores de μ superiores al ADNn, ya que al ser un genoma expuesto a agentes oxidativos y careciendo de histonas, presenta una baja eficacia en los sistemas de replicación y reparación celulares.

El ADNmt es de gran utilidad para estudios antropológicos sobre evolución humana, debido a la alta variabilidad de la región control. Las diferencias interpoblacionales a nivel mitocondrial son un reflejo molecular de sucesos históricos acontecidos durante los últimos milenios.

Esta característica hace al ADNmt ideal para apreciar relaciones evolutivas entre grupos de organismos muy cercanos, y le confiere gran importancia en los estudios moleculares y en sus aplicaciones filogenéticas. Por ejemplo en los estudios sobre las relaciones entre humanos, chimpancés, gorilas, orangutanes y gibones, un grupo cuyas relaciones evolutivas precisas han sido causa de grandes controversias

(Andrews et al., 1999; Arnason et al., 1996a, 1996b, 1998, 2000; Gokey et al., 2004; Green et al., 2008; Horai et al., 1995; Ingman et al., 2000; Xu and Arnason, 1996).

I.2.V. Tasa de mutación

La tasa de mutación (μ) es la frecuencia con la que ocurren cambios en la secuencia nucleotídica del ADN, en el genoma mitocondrial este parámetro es mayor que el nuclear (Parsons et al., 1997) y en algunas regiones puede ser de 6 a 17 veces la tasa de los genes nucleares de copia única (Brown et al., 1979, 1982; Neckelmann et al., 1987; Wallace et al., 1987, 1991, 1997).

Para comprender esta elevada tasa de mutación en el ADNmt han de considerarse varios factores.

- El consumo en este orgánulo de más del 90% del oxígeno celular en el metabolismo oxidativo, produce un constante flujo de radicales libres en la matriz mitocondrial.
- La ausencia de histonas en el genoma mitocondrial frente a los daños oxidativos.
- La mayoría de las mutaciones aparecen durante la replicación del material genético y el ADNmt tiene muchos más ciclos replicativos que el ADNn. Deben tenerse en cuenta los ciclos adicionales de la oogénesis que supone además, un cuello de botella.
- A diferencia del ADNn, la replicación del ADNmt tiene lugar durante todo el ciclo celular. Los orígenes de replicación para las cadenas L y H están localizados en regiones diferentes. Durante la replicación el ADN permanece como cadena sencilla durante un período de tiempo considerable, y éste es particularmente proclive a las mutaciones espontáneas. Por ejemplo la tasa de desaminación espontánea de C-U es unas 200 veces mayor que para el ADN de cadena doble. Esto podría explicar la pronunciada asimetría en la mutación del ADNmt: las transiciones G-A en la cadena L son aproximadamente 9 veces más frecuentes que en la cadena H (Tanaka and Ozawa, 1994).

Los segmentos hipervariables de la región control presentan una tasa de mutación aún más elevada que la de la región codificante (Howell et al., 1996). Estos elevados valores de μ en la región control nos permiten utilizarla como cronómetro: si se conoce el tiempo medio en el que se produce una variación en una región genética determinada (μ), basta con contar el número de diferencias medias que tienen dos

individuos o especies para estimar el tiempo que hace que estas tuvieron un antecesor común. Este aspecto convierte a la D-Loop en una región ideal para apreciar relaciones evolutivas entre grupos de organismos cercanos y en estudios intraespecíficos. Este principio es el fundamento en el que se basa el concepto de *reloj molecular*.

Este concepto fue definido en 1962 por Emile Zuckerkandl y Linus Pauling, y trabajado por Allan Wilson, Vincent Sarich y Motoo Kimura en 1968 (Kimura 1968; Sarich and Wilson, 1967; Zuckerkandl and Pauling, 1962). Actualmente es una herramienta muy utilizada en sistemática molecular ya que asume que la tasa de cambio de cualquier secuencia es aproximadamente constante a lo largo del tiempo y de diferentes linajes, por lo tanto la acumulación de diferencias entre dos secuencias puede utilizarse para medir el tiempo.

En estudios basados en la diversidad interna de una especie, sobre todo si es relativamente joven como *Homo sapiens*, es muy importante que μ sea elevada puesto que nos permite hacer una estima temporal muy detallada.

El punto más débil del cálculo temporal de la evolución de una especie a través del ADNmt es el valor de μ , cuya tasa media se estima, sin considerar la región control, en 0.017×10^{-6} sustituciones por año (Ingman and Gyllensten, 2003). De la misma manera unos valores más o menos aceptados de μ para las regiones HVI y HVII, oscilan entre $0,075$ a $0,165 \times 10^{-6}$ sustituciones por año (Tamura and Kei, 1993) con un valor medio de 0.47×10^{-6} sustituciones por año (Howell et al., 2003).

Aunque se han hecho gran cantidad de estudios de las regiones HVI y HVII en poblaciones de todo el mundo, recientes trabajos ponen en duda que estas sean las más adecuadas para hacer inferencias en el pasado de nuestra especie (Ingman et al., 2000). A pesar de todo los últimos trabajos utilizan un valor de $1,675 \times 10^{-7}$ (Santos et al., 2005).

I.2.VI. Herencia materna

Otra de las características del ADNmt que lo hace interesante por su aplicación en antropología, es su exclusivo sistema de herencia por vía materna. Los óvulos humanos tienen aproximadamente unas 100.000 mitocondrias mientras que en los espermatozoides el número oscila entre 50 y 75 distribuidas en el axonema, en la región más gruesa de la cola. Esta porción de la célula germinal masculina no interviene directamente en la fecundación con lo que las mitocondrias heredadas pertenecerán al óvulo, es decir, a la progenitora femenina (Manfredi et al., 1997). De

esta manera, el linaje mitocondrial de las mujeres con descendientes varones únicamente o que no tengan hijos, se perderá a no ser que tengan hermanas y estas a su vez, sobrinas. Por lo tanto, analizar el ADNmt de una persona es estudiar el ADNmt de sus ancestros por vía femenina. En la figura I.2 se ilustra este sistema de herencia en comparación con el ADNn.

El sistema de herencia materna del ADNmt no es común para todos los seres vivos (Zouros et al., 1992; Ankel-Simons and Cummins, 1996) y en los últimos años se han detectado excepciones (Schwartz and Vissing, 2002) que han abierto un debate sobre la conveniencia de asumir este sistema de herencia como estrictamente cierto (Bromham et al., 2003). A pesar de estas voces críticas la mayor parte del colectivo científico sigue apoyando la exclusiva transmisión femenina del ADNmt y su uso para la reconstrucción de la historia evolutiva de *Homo sapiens* (Jobling et al., 2004).

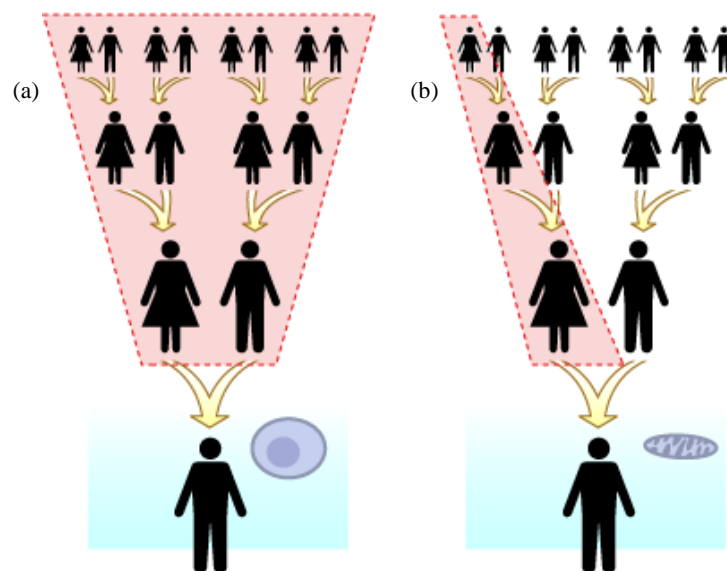


Figura I.2. Comparación entre el sistema de transmisión del ADNn (a) y el ADNmt (b).

I.2.VII. Ausencia de recombinación

La recombinación es el proceso por el cual dos cromosomas homólogos pueden intercambiar material genético. Este proceso es de suma importancia en la generación de variabilidad y de nuevas combinaciones genéticas en la descendencia, pero puede ser problemático a la hora de rastrear el origen de una porción de material genético. Si

ha habido recombinación, intercambio de información entre los cromosomas procedentes del progenitor masculino y femenino, no se podrá determinar de qué progenitor se ha heredado un fragmento de ADN concreto. Si por el contrario no existe recombinación y la herencia es materna, es obvio que todo el material genético procederá de la madre, y se podrá reconstruir fidedignamente la genealogía femenina de un individuo y de una especie. Es así como el ADNmt se transmite intacto, por vía femenina, y de generación en generación. El único sistema posible por el que se pueden introducir variaciones serán por tanto, las mutaciones.

Aunque generalmente no se produce recombinación (Hutchison et al., 1974; Giles et al., 1980), esta característica del ADNmt ha sido cuestionada en los últimos años. Gran parte de estos trabajos críticos se basaban en estudios de desequilibrio de ligamiento e interpretaciones de los árboles filogenéticos (Awadalla et al., 1999; Eyre-Walker et al., 1999; Eyre-Walker and Awadalla, 2001) aunque algunos autores incluso llegaron a afirmar que habían detectado recombinación genética en el ADNmt en poblaciones humanas de la Melanesia (Hagelberg et al., 1999). Estudios posteriores demostraron que todos estos resultados eran erróneos y se basaban en la utilización de datos defectuosos, en algoritmos matemáticos inadecuados o en errores de alineamiento de las secuencias (Arctander 1999; Jorde and Bamshad, 2000). Otros estudios llevados a cabo recientemente sobre la secuenciación completa de ADNmt han confirmado nuevamente esta ausencia de recombinación en el ADNmt (Ingman et al., 2000).

La ausencia de recombinación favorece el rastreo y reconstrucción de los marcadores genéticos, ya que inequívocamente estos se heredarán a través de líneas femeninas, eliminando cualquier otra posible vía de herencia y permitiéndonos una reconstrucción fidedigna de la genealogía femenina.

Esta misma reflexión es válida para realizar comparaciones entre poblaciones, ya que el trazado de los linajes maternos proporciona un recurso para examinar la estructura familiar en poblaciones.

I.2.VIII. Polimorfismos y heteroplasmia en el ADNmt

En 1966 Ford define el término polimorfismo como un carácter mendeliano monogénico, que se presenta al menos bajo dos formas alternativas en una población, y aunque su origen pueda ser la mutación su mantenimiento no puede depender de la recurrencia de la misma (Ford 1966). Cuando existen modificaciones en un gen a nivel de un locus específico en una población, se dice que este locus es polimórfico. Bajo

esta definición, para que un locus sea polimórfico el alelo más común debe tener una frecuencia menor del 99%.

Existen numerosos ejemplos de polimorfismos en el genoma humano. Los polimorfismos de ADN pueden ser de diversos tipos, desde la mutación de una sola base hasta el cambio en el número de unidades repetidas en tándem de ciertas regiones del ADN. Este trabajo se centra en los polimorfismos de secuencia, producidos por el cambio de uno (mutación puntual) o más nucleótidos en una secuencia de ADN. Los polimorfismos de sustitución simple son los que caracterizan fundamentalmente al ADNmt.

Muchos genes, tanto del ADN nuclear como del ADNmt, pueden expresarse de diferentes formas, y bajo estas circunstancias se dice que un gen es polimórfico, es decir, que puede tener varias formas de expresión. Los genes polimórficos son excelentes herramientas para estudiar y caracterizar las poblaciones humanas.

Los polimorfismos en el ADNmt pueden ser de dos tipos:

- 1) Polimorfismos de secuencia. El tipo predominante de sustitución son las transiciones, las purinas tienden a estar sustituidas con más frecuencia que las pirimidinas; aunque también podemos hallar transversiones, sustitución de una base por otra de distinta clase, es decir, pirimidina por purina o a la inversa. La mayoría de la variación en secuencia se encuentra dentro de la denominada región control (Greenberg et al., 1983).
- 2) Polimorfismos de longitud. Es frecuente encontrar en el ADNmt inserciones o deleciones de una o más bases nucleotídicas. Estas inserciones y deleciones están localizadas entre las posiciones 16.184 hasta 16.188, y entre la 16.190 hasta 16.193 de la región HVI; y entre las posiciones 303 hasta la 309, y entre la 311 hasta la 315 en la región HVII.

Las moléculas de ADNmt que contiene un individuo habitualmente son iguales, cuando existen diferentes tipos estamos frente al fenómeno de la *heteroplasmia*. Con bastante frecuencia coexisten moléculas con diferente número de inserciones dentro de una misma mitocondria o en distintas mitocondrias dentro de una misma célula. En la actualidad se considera que el 14% de la población humana tiene un segundo tipo de ADNmt con una frecuencia menor del 1% (Tully et al., 2000).

Debe tenerse en cuenta que el mantenimiento de la heteroplasmia para mutaciones asociadas a enfermedades puede ser seleccionada positivamente. De este modo el

cuello de botella que se produce en la oogénesis tendría el efecto de mecanismo amplificador, pudiendo quedar reflejado en la población (Behar et al., 2004). Simultáneamente se produce la amplificación en las moléculas de ADNmt (por 100) y la del número de mitocondrias (por 1000), lo que en términos de tipos mitocondriales podría conllevar una segregación si partimos de dos tipos iniciales. En un único ciclo total podemos encontrar gran número de copias de ADNmt de dos tipos distintos.

Es relativamente frecuente detectar en el ADN repeticiones de fragmentos cortos que se conocen con el nombre de *minisatélites* y *microsatélites*, en función del número de nucleótidos que se repitan. De manera que un individuo puede ser caracterizado por el número de veces que tiene repetida una secuencia corta de ADN. Los microsatélites son una excelente herramienta utilizada en genética forense (Balding 1999), aunque también se utilizan con muy buenos resultados para reconstruir la historia biológica de las poblaciones (Ruíz-Linares et al., 1999).

En el ADNmt únicamente se ha descrito un microsatélite (Bodenteich et al., 1992) en la posición 514 de la región D-loop. Se trata de un microsatélite dinucleotídico (AC)_n con un bajo grado de heterocigosidad.

Un método muy utilizado en el laboratorio para la determinación de polimorfismos de secuencia es el RFLPs (Restriction Fragment Length Polymorphism). Los RFLPs se basan en la elección de fragmentos específicos del ADN en los que existen cambios o mutaciones mediante un enzima de restricción. Estos enzimas son endonucleasas producidas por bacterias como mecanismo de defensa contra ADN extraño. Cada enzima corta la doble cadena de ADN en una secuencia que reconoce de forma específica (diana de restricción). Si el fragmento a analizar presenta la diana la endonucleasa la reconoce y corta el ADN en esa posición, si el ADN no presenta la diana, la cadena no se corta. En el primer caso se obtendrían dos fragmentos de la cadena original de tamaños conocidos, en el segundo caso un único fragmento. La presencia o ausencia de un fragmento puede utilizarse cuando la divergencia entre los grupos comparados es menor al 15%. Por encima de este valor el tamaño de los fragmentos puede deberse a una convergencia, es decir, la longitud del fragmento es la misma pero las dianas de restricción que lo han producido son distintas. Para evitar este problema se construyen mapas de restricción de la molécula en estudio.

I.3. Aplicaciones del ADNmt

Los análisis de ADNmt representan una visión amplificada de la variabilidad humana gracias a la elevada tasa de mutación que posee y junto con el resto de sus características, lo convierten en una de las herramientas más útiles en estudios antropológico-moleculares. Este marcador ha hecho grandes aportaciones en estudios de evolución humana durante los últimos años. Pero es en la actualidad cuando su aplicación permite realizar un análisis detallado de nuestra diversidad y, por extrapolación, de nuestra historia. Para llegar a este punto ha sido necesaria la convergencia de varias circunstancias: caracterización de gran cantidad de poblaciones, la secuenciación completa de algunos ADNmt, los avances en los protocolos de análisis y el desarrollo tecnológico de *software* específicos.

I.3.1. Primeros estudios

Uno de los primeros métodos empleados en el análisis del ADNmt fue el análisis de fragmentos con enzimas de restricción o RFLP's. Wesley Brown y Douglas Wallace fueron pioneros en esta clase de estudios a partir de la década de los 70. Estas primeras investigaciones se iniciaron mediante digestión con una única enzima y con un amplio número de muestras (Denaro et al., 1981); en otros casos, se empleaban varias enzimas con pocas muestras (Brown 1980).

La primera investigación en ADNmt se llevó a cabo digiriendo la molécula entera con algunas enzimas de restricción. En 1980 el estudio pionero de Brown con 21 humanos de diversas etnias y orígenes geográficos indicaban que los estudios en ADNmt con los patrones de RFLP, podían ser utilizados para trazar la historia genética humana (Brown 1980). En base a la diversidad observada en una muestra mundial, Brown obtuvo el sorprendente tiempo estimado de coalescencia de 180.000 años para la variación de ADNmt global.

Estas primeras experiencias dieron la pauta para comprender que fuera cual fuera el número de muestras, se requería siempre un análisis enzimático lo más completo posible. Posteriormente aumentó el número de enzimas empleadas para el análisis de mayor número de muestras (Scozzari et al., 1988), y estos resultados se representaban como una filogenia de la variación global del ADNmt. Estas filogenias fueron representadas mediante árboles en los que a partir de datos menos variables se generaban agrupaciones, que a su vez iban originando la aparición de nuevas agrupaciones, afines por presentar menor variación que el grupo que las originó.

Esta forma de construcción de árboles filogenéticos fue llamada *starlike* o topología estrella, en la que un conjunto de elementos que comparten ciertos caracteres se estructuran y dan lugar a otros conjuntos con características propias. Es decir, un haplotipo central (universal) debió ser asumido para representar al ancestro común más reciente (MRCA) de todos los ADNmt en el mundo.

I.3.II. Relaciones genealógicas entre primates y *Homo sapiens*

En el año 1987 y después de otros trabajos previos, Rebecca Cann publicó uno de los primeros estudios en que se empleó el ADNmt como herramienta en la investigación de la relación genealógica entre humanos y primates, así como en investigaciones acerca de las relaciones entre humanos (Cann et al., 1982, 1984, 1987).

A partir de ese momento han sido innumerables los intentos de crear una filogenia molecular entre los humanos, los chimpancés y otros parientes cercanos. Morris y Goodman propusieron en 1994 una filogenia basada en marcadores autosómicos (Goodman et al., 1994), y Kim y Takenaka una alternativa fundamentada en marcadores del cromosoma Y (Kim and Takenaka, 1996). Aunque en la actualidad el árbol filogenético molecular de los grandes simios se basa en la comparación de ADNmt completos (Figura I.3). Es esta figura se detalla el árbol de máxima parsimonia creado a partir de los ADNmt completos de diferentes especies, que interrelaciona a los grandes simios con neandertales y con dos secuencias humanas, la de referencia (Cambridge) y la de un individuo sub-sahariano (Yoruba). Los códigos en los extremos de las ramas son la referencia de los ADN mitocondriales en Genbank.

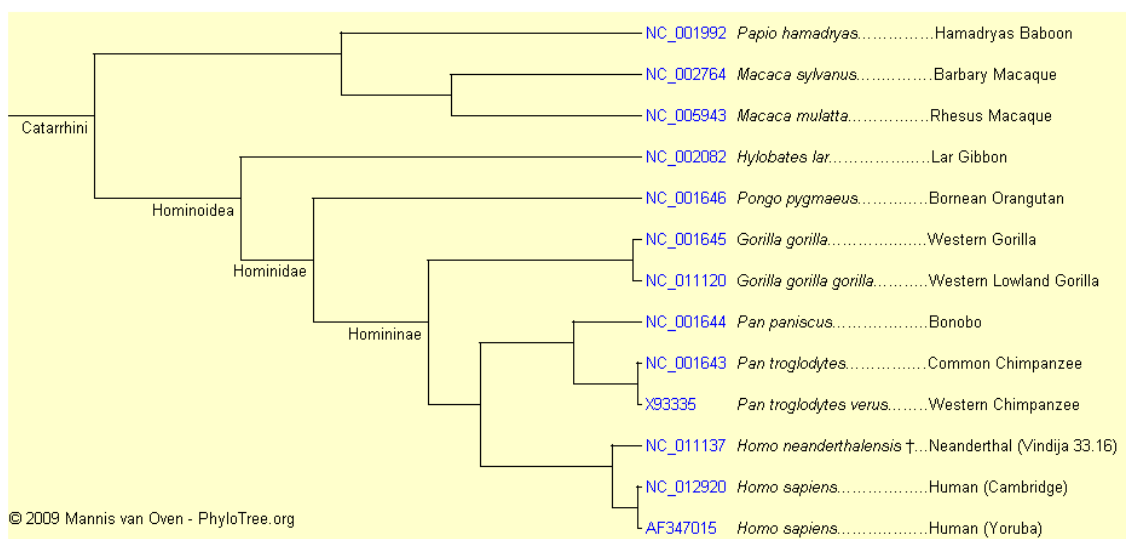


Figura I.3. Árbol de máxima parsimonia basado en alineamiento del ADNmt completo de diferentes especies (<http://www.phylotree.org>)

Los resultados son obvios y uno de los aspectos más relevante es la cercanía entre humano y neandertal y las diferentes especies de chimpancé (*Pan sp*).

I.3.III. Utilidad forense: la familia Romanov

Una de los campos en que más relevancia ha tenido el conocimiento de la estructura y dinámica del ADNmt ha sido en el campo de la genética forense. El elevado número de copias de esta molécula en las células la hace especialmente adecuada tanto en casos forenses como en cualquier caso en el que la muestra sea de análisis complejo, bien sea por el estado en el que se encuentre como en la cantidad disponible de vestigio biológico. El análisis del ADNmt es una herramienta muy eficaz y utilizada de rutina hoy en día en este campo (Holland and Parsons, 1999).

Un caso muy representativo y con gran divulgación fue el de la familia Romanov (Gill et al., 1994; Coble et al., 2009; Rogaev et al., 2009). Restos humanos que podrían pertenecer a los miembros de la última familia real rusa (Zar Nicolás II de Rusia, Zarina Alexandra y sus hijos) ajusticiados por los bolcheviques en 1918, fueron encontrados cerca de la ciudad de Ekaterimburgo en los Urales, Rusia central.

Los análisis del ADN cromosómico confirmaban que se trataba de una familia, pero fue necesario comparar las secuencias del ADNmt con descendientes de la familia imperial rusa para probar que realmente se trataba de los Romanov. Comparaciones entre las muestras de los cuerpos indicaban que cuatro muestras eran idénticas (la Zarina Alexandra y tres de sus hijos) y las cinco restantes diferían en cuatro nucleótidos (el Zar Nicolás II, tres sirvientes y el médico de la familia). En el ADNmt obtenido del supuesto Zar Nicolás II se detectó heteroplasmia por una única base de la región control. Esto obligó a una comparación con dos descendientes de la rama materna para confirmar que se trataba del último zar ruso.

I.3.IV. Dispersión humana y origen geográfico

Debido a que el proceso de diferenciación molecular del ADNmt es relativamente rápido y ocurrió mayoritariamente durante y después del proceso de dispersión de *Homo sapiens* desde África hacia distintas regiones del mundo, los tipos de ADNmt tienden a estar habitualmente restringidos a una determinada área geográfica y población, por lo que la variación genética entre poblaciones humanas muestra que los grupos que viven en el mismo continente son relativamente homogéneos (Bamshad et al., 2004; Underhill and Kivisild, 2007).

Recientes análisis de esta variación al máximo nivel posible de la resolución molecular nos están permitiendo determinar el origen y la diversidad en distintas regiones del mundo, con el objetivo de reconstruir la historia de la dispersión humana. En la actualidad hay múltiples trabajos centrados en la secuenciación completa de ADNmt (Saillard et al., 2000; Lewis et al., 2004; Jackson et al., 2005; Zlojutro et al., 2006).

Al principio de los 90 en paralelo a los estudios del origen de *Homo sapiens*, los análisis de ADNmt empezaron a ser aplicados a los distintos continentes de manera individual y con gran número de muestras, con el objetivo de determinar el origen de los humanos en las grandes áreas geográficas. Estos estudios se llevaron a cabo mediante una combinación de RFLPs de alta resolución y la secuenciación de HVSI (Macaulay et al., 1999), y permitió una reconstrucción filogeográfica de nuestra especie.

La filogeografía consiste en cartografiar la expansión de las ramas y subramas de los árboles del ADNmt o de cualquier otro marcador. Las reglas que rigen estos estudios son sencillas, para detectar una migración de una región a otra hay que encontrar el tipo de rama original que había en ambas regiones y una posterior ramificación excepcional en la región de destino que no esté presente en la de partida.

La reconstrucción de los movimientos pasados de nuestra especie a través del estudio del ADNmt, requiere asumir que cada vez que se produce una nueva variación aparecerá un nuevo linaje. Este contendrá todas las variantes heredadas más las específicas que se han producido de nuevo. Es decir, los linajes modernos contendrán marcadores específicos discriminantes, y los antiguos marcadores comunes a su vez a todos los linajes derivados. De esta manera podemos reconstruir un árbol que se bifurca en ramas caracterizadas por polimorfismos específicos.

Además, el método permite determinar dónde y cuándo se han producido estos polimorfismos ya que las mutaciones se producen a un ritmo constante pero aleatorio. Así, podemos averiguar el momento y el lugar en el que estas se produjeron estudiando las variaciones de las poblaciones actuales. Es necesario matizar que las estimaciones del tiempo que tienen los linajes son aproximadas y se basan en un cálculo medio de las mutaciones posteriores en los tipos filiales de una rama (Saillard et al., 2000). Por ejemplo, los cambios más antiguos en el ADNmt humano compartidos por todas las poblaciones del mundo, están presentes en algunas poblaciones africanas actuales y tienen una antigüedad de entre 150.000 y 190.000

años. Las poblaciones que presentan los segundos cambios más ancestrales son las asiáticas, con una antigüedad de entre 60.000 y 80.000 años. Esta información puede interpretarse de una forma sencilla: los humanos modernos son originarios de África y algunos de ellos colonizaron el continente asiático hace al menos 80.000 años. Este es el resumen del pensamiento lógico para aplicar esta metodología a la reconstrucción de nuestra historia. Utilizando las variaciones del ADNmt y apoyándonos en herramientas estadísticas e informáticas, podemos reconstruir nuestros orígenes y la dinámica de expansión de nuestra especie.

Antes de hacer una descripción más o menos pormenorizada de la historia del *Homo sapiens* basada en el ADNmt, es necesario comentar un aspecto importante que suele ser fruto de malas interpretaciones. El modelo de estudio basado en variantes polimórficas del ADNmt nos aporta únicamente información de los haplogrupos que generan estas mutaciones, pero no de las poblaciones que portan estos haplogrupos. Por ejemplo, hay una tendencia a equiparar la antigüedad de los haplogrupos con la edad de las poblaciones y asumir que la separación de los haplogrupos refleja migraciones. En realidad, cuando una población migra sus componentes llevan consigo todos los haplogrupos, no sólo uno de ellos, y la edad de los haplogrupos indica cuando ocurrió la mutación puntual que define el haplogrupo, no cuando ocurrió la migración (Simoni et al., 2000a).

Los haplogrupos del ADNmt presentan una distribución bien definida y estructurada geográficamente. El macrohaplogrupo L y sus múltiples variantes se localizan en África (Salas et al., 2002). En el este de África el haplogrupo L3 dio lugar a las variantes M y N que colonizaron y se dispersaron por Eurasia y el nuevo mundo (Oota et al., 2001). Los haplogrupos H, I, J, N1b, T, U, V y W son característicos de las poblaciones descendientes de europeos. Los haplogrupos A, B, C y D se encuentra en Asia y América con las variantes G, Y y Z predominantemente en Siberia. Por otro lado el haplogrupo X se encuentra en bajas frecuencias en el norte de África y centro y oeste de Asia, Europa y el Nuevo Mundo, sin embargo, no se han detectado en poblaciones originarias de Siberia y Este de Asia. Como todos estos haplogrupos comparten un ancestro común, su estudio puede ayudarnos a conocer el origen de las poblaciones así como a reconstruir su rastro migratorio.

I.3.V. Uso potencial del ADNmt en biomedicina

En los últimos años, investigaciones biomédicas están descubriendo la implicación del genoma mitocondrial en una gran cantidad de enfermedades genéticas metabólicas. Debidas a mutaciones puntuales, a la sustitución de unas bases por otras, deleciones o inserciones, afectan finalmente a ciertos genes mitocondriales. La neuropatía óptica hereditaria de Leber, es una de estas enfermedades. El cuadro clínico comprende la degeneración del nervio óptico y la consecuente ceguera alrededor de los 20 años de vida.

Un ejemplo de suma importancia es su implicación en enfermedades neurodegenerativas causadas por deleciones en el ADNmt de células musculares en enfermos con encefalopatías (Serrano, 1995). También su relación con el cáncer ha despertado gran interés a causa de su papel central en la apoptosis, cuya intervención en el desarrollo de tumores es primordial. Al parecer su relación con los procesos funcionales de la mitocondria es crucial.

Algunos estudios publicados recientemente, apuntan que algunas variantes del ADNmt con diferencias para el metabolismo energético, permitieron a los humanos la adaptación a climas fríos, produciendo como resultado el enriquecimiento de linajes específicos (Ruiz-Pesini et al., 2004). También han proliferado los estudios que relacionan este genoma con las cardiopatías (Zaragoza et al., 2010) e incluso con la longevidad (Guo et al., 2010).

I.4. Expansión de *Homo sapiens*

Los avances analíticos, tanto metodológicos como interpretativos, unidos al estudio de gran cantidad de poblaciones humanas de todo el mundo, han permitido postular un modelo más o menos aceptado por la mayor parte de la comunidad científica, sobre el origen y la dispersión de *Homo sapiens* conocido como "Out of Africa". Esta teoría se centra en la interpretación de la diversidad genética humana y postula un origen único africano de nuestra especie entre 150.000 y 200.000 AAP.

Este modelo, de forma resumida y respecto a la información aportada por el ADNmt, propone que las poblaciones humanas se caracterizan por la presencia de variantes mitocondriales que se estructuran en el tiempo y en el espacio. La dispersión de nuestra especie se produjo desde el continente africano hacia Eurasia unos 80.000 AAP a través del estrecho de Omán, siguiendo la ruta del sudeste asiático mediante la

explotación de recursos marinos, hasta colonizar Australia y Nueva Guinea. La colonización de Asia central se habría producido a través de los grandes ríos mediante el modelo conocido como centrípeto, siguiendo las grandes rutas fluviales asiáticas hasta el centro del continente. Esta colonización se pudo llevar a cabo unos 50.000 AAP. El poblamiento de Europa occidental se produjo a partir de dos olas migratorias, la primera caracterizada por el haplogrupo U5 unos 45.000 AAP; y la segunda, por la presencia de variantes mitocondriales HV con una antigüedad de unos 31.000 años. El último continente en ser poblado por *Homo sapiens* fue el americano, y todo parece indicar que lo hicieron grupos procedentes de Asia unos 18.000 AAP.

Los antecedentes de esta teoría se remontan al año 1987 con la publicación en la revista *Nature* de un artículo firmado Rebecca Cann que revolucionaría la concepción que se tenía sobre la historia de *Homo sapiens*. En él se comparaban los ADNmt de 147 individuos procedentes de diferentes poblaciones del mundo. Hasta ese momento la reconstrucción del pasado se realizaba casi exclusivamente mediante el estudio del registro fósil. Los resultados avalaban de forma inequívoca que el origen de la humanidad estaba en el continente africano. Además, demostraban que la diversidad del hombre era muy homogénea al compararla con otras especies como el chimpancé, por lo que concluían que *Homo sapiens* era una especie muy reciente (Cann et al., 1987).

Los autores basaban su propuesta en dos principios. El primero de ellos era asumir que la herencia del ADNmt se realizaba exclusivamente por vía materna. El segundo supuesto, derivado del primero, consideraba que toda la diversidad mitocondrial humana actual debería de proceder de un único linaje mitocondrial que se fue diversificando, mediante la acumulación de mutaciones al azar a lo largo del tiempo, aproximadamente una cada mil generaciones (Forster et al., 1996). Este estudio estableció las bases para el uso del ADNmt como herramienta genética, siendo el inicio de la exploración de la historia biológica de poblaciones de Europa (Simoni et al., 2000b), Asia (Comas et al., 1998), América (Torroni et al., 1993a, 1993b; Bortolini et al., 1997; Alves-Silva et al., 2000) y, por supuesto, África (Chen et al., 2000).

Este trabajo proponía que la totalidad de la diversidad de linajes mitocondriales humanos derivaban de una única variante portada por una mujer procedente de África, a la que se llamó “Eva mitocondrial”. La aceptación de sus postulados implicaba aceptar la teoría de “Out of Africa” y rechazar la hipótesis vigente hasta el momento llamada “Teoría multiregional”, que aunque apoya también un origen único y africano

de *Homo sapiens* propone una antigüedad mucho mayor. Además considera que toda la humanidad procede de *Homo erectus* que salió de África aproximadamente 1.200.000 AAP. Estos se distribuyeron por Euroasia diferenciándose y gracias a un flujo genético entre poblaciones, dieron lugar a *Homo sapiens* en diferentes lugares del mundo. La diversidad genética actual no sólo basada en el ADNmt sino también en otros marcadores moleculares, contradice este planteamiento. En la actualidad se acepta que los planteamientos generales de la teoría “Out of Africa” se ajustan a la realidad antropológica, arqueológica, paleoantropológica, histórica y genética del hombre.

I.4.I. África, el origen

En el año 2008 Behar y colaboradores generaron un árbol filogenético del ADNmt que reconstruía la historia de las poblaciones africanas (Behar et al., 2008). En él se presentaba una bifurcación de linajes cercana a la raíz con una antigüedad de unos 200.000 AAP. Esta separación dividía el árbol filogenético en dos ramas. La primera de ellas muy extendida, L1'2'3'4'5'6 (L1'5) según la nomenclatura de los autores. La sub-rama L3 dio lugar a los clados N y M entre 50.000 y 65.000 AAP, a partir de los cuales se generó toda la diversidad genética no africana del mundo (Forster et al., 1996; Watson et al., 1997; Ingman et al., 2000; Saillard et al., 2000). La otra rama denominada L0, es exclusivamente africana y se ramificó en L0abfk unos 140.000 AAP y en L0d unos 100.000 AAP. La rama L0d y una variante de L0abfk llamada L0k, están presentes en el 60% de las poblaciones koisánidas del desierto del Kalahari, en Sudáfrica, de las que entre 1 y 8 haplogrupos se han mantenido hasta la actualidad (Salas et al., 2002). Estos grupos podrían situarse, por sus características culturales y lingüísticas, en la base de la evolución de *Homo sapiens*. Los análisis filogenéticos de las variantes del ADNmt sugieren que la separación entre Koisánidos y el resto de poblaciones africanas tiene una antigüedad mínima de 90.000 años.

Después de esta primera diversificación, los acontecimientos demográficos en África configuraron un espectro de clados de ADNmt con frecuencias variables en el continente, cuya interpretación no es sencilla si no asumimos que el origen del *Homo sapiens* anatómicamente moderno debe encontrarse en algún lugar del sudeste africano y se produjo aproximadamente 200.000 AAP. Estas primeras poblaciones de *Homo sapiens* sufrieron una separación muy antigua que facilitó la existencia de una profunda estructura genética, asociada a importantes fluctuaciones climatológicas y a fuertes procesos de aislamiento durante periodos de entre 100.000 y 50.000 años. En

la actualidad esta estructura genética ancestral todavía es detectable, especialmente entre los clados presentes en los Koisánidos y en el resto de poblaciones africanas. Por otro lado, es llamativa la alta representación del clado L1e en el centro y oeste africano entre los Biaka o Pigmeos del Oeste, y la alta proporción de los linajes L1e y L1f en el este de África.

Poblaciones portadoras de L1'5 quedaron recluidas a la región de origen, mientras que otras caracterizadas por las variantes L0 se desplazaron hacia el sur. Ambos grupos evolucionaron de forma independiente, diferenciándose respecto a su composición en los haplogrupos del ADNmt. Entre 100.000 y 120.000 AAP aparecieron las variantes L0abf en las poblaciones del sur, que regresaron a regiones septentrionales del continente formando un *pool* genético compuesto por los recién llegados del sur y las poblaciones que permanecieron en el este del continente. Otro acontecimiento importante lo representó la expansión y colonización de la región noreste del continente, y la salida de África por el estrecho de Omán. Los investigadores consideran que en este periodo estaban presentes como mínimo 40 variantes mitocondriales en las poblaciones africanas, siendo insólito que únicamente una de ellas saliera de África y se convirtiera en la raíz de todos los linajes no africanos del mundo. Aunque hubieran sido varias, es improbable que desaparecieran por efecto de la deriva genética a no ser que asumamos la existencia de una estructuración genética. Si además tenemos en cuenta que probablemente los primeros grupos de exploradores fueran de tamaños relativamente pequeños, sería aceptable considerar que no todos los haplogrupos de ADNmt estaban representados. A partir de este momento y entre 80.000 y 60.000 AAP, se produjo una re-expansión que colonizó gran parte del continente africano produciéndose una sustitución del haplogrupo L1, que pasó a ser minoritario a excepción de los ya mencionados Koisánicos y Biaka, por las nuevas variantes L2 y L3. Este período es coincidente con la aparición del comportamiento moderno de *Homo sapiens*, datándose en la cueva de Blombos (Sudáfrica) muestras geométricas sobre hueso, consideradas primeras expresiones artísticas en 70.000 AAP (Henshilwood et al., 2002).

Este modelo de expansión justificaría gran parte de la realidad genética actual africana, aunque es necesario considerar otros dos sucesos fundamentales. En primer lugar, la expansión de los grupos de habla Bantú y la agricultura transformaron grandes regiones y poblaciones de África (Salas et al., 2002). En segundo lugar la recolonización del norte de África por poblaciones eurasiáticas que dispersaron las

variantes M1 y U6 llegando a regiones meridionales del Sahara y terminando de definir el panorama mitocondrial (Rando et al., 1998; Richards and Macaulay, 2001).

I.4.II. El éxodo africano

Los ciclos interglaciares marcaron la salida del hombre del continente ya que abrían o cerraban el acceso a determinadas rutas. El hombre pudo ver únicamente dos de estos fenómenos, que permitieron su salida de África a través de dos posibles vías. La primera de ellas por el norte hacia el Sahara, Oriente Próximo y Europa, la llamada ruta septentrional; y la ruta meridional, hacia el sur, el mar Rojo, Omán, Yemen y la India (Quintana-Murci et al., 1999).

Durante el periodo Eemiano tuvo lugar el primero de ellos, aproximadamente 125.000 AAP, y coincidiendo con el primer intento de expansión del hombre fuera de África. Los primeros restos fósiles extra-africanos, con una antigüedad entre 120.000 y 90.000 AAP, se encontraron en Oriente Próximo (De Laet 1994). En este periodo la ruta septentrional permaneció abierta y se produjo la primera salida del continente africano, aunque la gran mayoría de investigadores considera que esta primera migración no tuvo éxito, ya que 90.000 AAP se produjo un breve e intenso periodo de frío que habría bloqueado el paso, evitando así la progresión del hombre anatómicamente moderno más allá de Oriente Próximo (Siddall et al., 2003)

La ruta, el número de olas migratorias y el periodo de las primeras migraciones fuera de África han generado un intenso debate (Templeton 2002; Chiaroni et al., 2009; DeGiorgio et al., 2009; Amos and Hoffman, 2010; Petraglia et al., 2010), aunque el hecho de que todos los individuos extra-africanos procedan de un único linaje materno, apoya inequívocamente una única salida desde África. Teniendo en cuenta los fenómenos de deriva génica y efecto fundador y en base a la dinámica poblacional, se considera que los primeros humanos que salieron del continente portaban quince linajes distintos (existen quince linajes con una antigüedad superior a 80.000 años), que podrían haber ido desapareciendo por el efecto de estos fenómenos. Todas aquellas líneas sin descendencia o con descendencia masculina se perderían. Los datos de cromosoma Y apoyan esta hipótesis de una sola salida de los importantes efectos de las fuerzas evolutivas, puesto que todos los varones de *Homo sapiens* comparten el haplogrupo originario M168 (Underhill et al., 2001; Garrigan and Hammer, 2006; Campbell and Tishkoff, 2008).

Sin embargo, la ruta meridional es la salida más probable a través del estrecho de Bab el Mandeb o Puerta de las Lamentaciones, 80.000 AAP durante el período glacial. La separación de los linajes mitocondriales africanos y no africanos habría tenido lugar unos 52.000 AAP, con un margen de error de unos 28.000 años (Macaulay et al., 1999; Chen et al., 2000; Herrnstadt et al., 2002a, 2002b). El paso de unos 17 Km debió hacerse necesariamente con balsas, y la expansión de los primeros exploradores por el continente euroasiático se hizo siguiendo la costa y mediante el modelo llamado “barre playas” aprovechando los recursos marinos (Rohling et al., 1998; Fenton et al., 2000; Walter et al., 2000). Esto está apoyado por la bajada en el nivel del mar de más de 80 metros que se produjo entre 60.000 y 80.000 AAP. El yacimiento de Adbur en Eritrea, muy cerca de la Puerta de las Catástrofes y de unos 100.000 años de antigüedad, confirma esta hipótesis (Walter et al., 2000). La estrategia de expansión asociada a la explotación de los recursos implicó la llegada del hombre a Indonesia en solamente 10.000 años desde su salida de África. Esta hipótesis de expansión está confirmada por el hallazgo de restos líticos de la cultura Kota Tampan en las cenizas de la explosión del volcán Toba producida unos 74.000 AAP, lo que confirmaría la presencia humana en el sudeste asiático coincidiendo con la expansión costera del hombre anatómicamente moderno (Majid 1998). Igualmente la antigüedad de los linajes mitocondriales de los aborígenes australianos de unos 80.000 años, confirmaría esta hipótesis (Redd and Stoneking, 1999).

Los datos genéticos apoyan este modelo expansivo, el linaje L3 con una antigüedad estimada de unos 83.000 años, dio lugar a dos variantes M y N; la primera es exclusivamente asiática, más antigua y heterogénea, y está distribuida en la India, Asia oriental y meridional. La variante N es algo más reciente, y se extendió preferentemente por Europa. Todavía se pone en duda si los linajes M y N son realmente originarios de África, ya que están presentes en poblaciones africanas y, o bien se produjeron durante el éxodo (Forster et al., 2001) o en un lugar mucho más oriental cerca de la India (Kivisild et al., 2003, 2004). Probablemente el origen del haplogrupo N esté en el golfo Pérsico, formado en aquella época por lagos de agua dulce poco profundos, alimentados por los ríos Tigris y Éufrates. Este linaje dio lugar al haplogrupo R unos 60.000 AAP. Los descendientes de estas poblaciones no aparecen en Europa y Oriente Próximo hasta 45.000 y 50.000 AAP, por lo que debieron mantener cierta estabilidad y poca movilidad. Esta fecha coincide con un inter-estadio de calentamiento global que favoreció la aparición de un corredor verde que habría conectado el golfo Pérsico y Siria, permitiendo a las poblaciones recorrer el Creciente

Fértil hacia Oriente Próximo. El origen en el sudeste asiático de la variante N de los proto-europeos se constata con el hecho de que cada variante de N en Europa tiene una variante en la India. El haplogrupo M se ha detectado con cierta frecuencia en Asia, aunque sus portadores son hablantes de lenguas afroasiáticas sus variantes son muy semejantes a las encontradas en la India (Kivisild et al., 1999). Por lo tanto es probable que la presencia de este linaje en África haya sido causada por una recolonización llevada a cabo en los últimos 20.000 años (Parra et al., 1999; Schurr and Wallace, 2002; Hill et al., 2006) aunque otros autores sitúan su origen en el continente africano, concretamente en Etiopía (Quintana-Murci et al., 1999).

La ausencia de las variantes M y N en África del norte y Oriente Próximo, y su abundante presencia en la India permiten descartar la ruta septentrional y aceptar la meridional por la desembocadura del Mar Rojo (Stringer and McKie, 1996; Lahr and Foley, 1998; Underhill et al., 2001).

I.4.III. El sudeste asiático y Australia

Asumiendo que todas las poblaciones no africanas tienen un mismo origen, se esperaría que sus antepasados tuvieran un referente mitocondrial único y de la misma antigüedad. De manera que todos los linajes deberían remontarse a un solo punto en el tiempo y en el espacio. El estudio genético de las poblaciones aborígenes y originarias de las primeras expansiones de *Homo sapiens* por el sudeste asiático, debería aportar datos para la mejor comprensión de la colonización más ajustada a la realidad.

Las conclusiones del estudio genético del pueblo Semang o Negritos de las islas de Malaca, muestran que este pueblo indígena tiene exclusivamente líneas mitocondriales situadas en la raíz de los haplogrupos M y N. Presentan la variante R21 descendiente de R que a su vez deriva de N, y la variante M de la que son portadores no aparece en ninguna otra parte del mundo y tiene una antigüedad de 60.000 años. Sus antepasados probablemente fueron los primeros buscadores de conchas extra africanas. El estudio de otros pueblos como los Jarawa y Onge, que habitan las regiones más meridionales de las islas de Andamán frente al golfo de Bengala, revelan que algunos de sus linajes son el M2 y M4 asignados al haplogrupo M, lo que confirmaría sus estrechas relaciones con los primeros pobladores no africanos que colonizaron el sur de Asia (Redd and Stoneking 1999; Karafet et al., 1999). Igualmente el haplogrupo F, derivado del R, se encuentra con cierta frecuencia en los indígenas

Malayos que habitan Indochina y la península de Malaca, y en algunos pueblos aborígenes de las grandes Andamán, siendo característico de poblaciones asiáticas (Endicott et al., 2003; Thangaraj et al., 2005; Reich et al., 2009).

La ruta de expansión por el sudeste asiático llegó a la península de Malaca, Nueva Guinea y Australia, ya que los pueblos aborígenes australianos y neoguineanos pertenecen al clan N y M. De hecho, las dataciones sitúan los primeros asentamientos en la Tierra de Arnhem en el norte de Australia entre 60.000 y 50.000 AAP (Roberts et al., 1990). Por otra parte el polémico esqueleto del lago Mungo (Barbetti and Allen, 1972; Thorne et al., 1999; Brown 2000; Thorne and Curnoe, 2000; Adcock et al., 2001; Bowler et al., 2003; Smith et al., 2003; Grün 2006) se ha datado en 62.000 años de antigüedad, por lo que la llegada de los primeros colonizadores a Australia tuvo que realizarse entre 70.000 y 65.000 AAP. Este periodo coincide con una cota baja del nivel del mar, lo que habría favorecido la expansión de los humanos anatómicamente modernos. Las líneas ancestrales del ADNmt de los aborígenes australianos se han datado en unos 68.000 AAP (Redd and Stoneking, 1999).

La colonización de Australia debió ser independiente a la de Nueva Guinea, que probablemente se realizó a través de Sulawesi y las Malucas, donde las distancias interinsulares son más cortas y no están afectadas por los cambios del nivel del mar. La información genética muestra que australianos y neoguineanos no tienen haplogrupos en común, estando los últimos tan alejados de los primeros como de cualquier otra población no africana. Además, la antigüedad de los linajes neoguineanos se remonta a 77.000 años, muy anterior a la de los australianos (Redd and Stoneking, 1999).

Se detecta una transición gradual en las proporciones de los dos linajes N y M en el sudeste asiático (Yao et al., 2002), dando información sobre su posible origen conforme nos desplazamos en dirección oeste-este. El pueblo Hadramaut en Yemen presenta una proporción de 5:1 en su relación N:M, y en Pakistán entre el mar Rojo y la desembocadura del Indo se sigue manteniendo esta relación aunque en una proporción algo menor. Más hacia el este, el Rajastán y Gujerat en el noroeste de la India, la proporción se iguala 1:1, llegando a invertirse a la altura de Bengala y Bangladesh (Kivisild et al., 2003).

La ruta meridional alternativa que siguieron los primeros humanos al salir de África unos 120.000 AAP, les permitió cruzar Asia por el sur y progresar hacia el este aprovechando los recursos disponibles.

I.4.IV. El centro del continente asiático

En su desplazamiento bordeando las costas los humanos encontraron las desembocaduras de los grandes ríos, lo que tuvo que plantear la posibilidad de atravesarlos o ascender por ellos explotando los recursos. De cualquier manera se les planteó la alternativa de una expansión centrípeta a través de los ríos y dirigida hacia el centro del continente, eludiendo así las grandes barreras geográficas que representaban las cadenas montañosas y los desiertos. Los estudios genéticos sobre la colonización de Asia Central postulan al menos cuatro posibles rutas, tres desde la costa indo-pacífica (por el oeste, el sur y el este) y otra a través del noroeste por Rusia (Derenko et al., 2007). La primera ruta indo-pacífica permitiría el acceso desde el oeste a Sinkiang y Mongolia aprovechando el curso de los ríos Tarim y Dzungaria; la segunda ruta ascendería por el río Indo, bordeando las estribaciones del Himalaya por su vertiente oeste hacia el centro del continente. Probablemente los grupos que eligieron esta vía llegaron en un período benigno a la región de Altái unos 43.000 AAP. Una tercera alternativa implicaría remontar ríos como el Mekong, el Brahmaputra, el Yangtsé y/o el Saluén, alcanzando la meseta de Quinghai unos 60.000 AAP. Como última alternativa y sin excluir las anteriores, cabe considerar el acceso a Asia Central a través de una vía más septentrional desde la Rusia asiática hasta el norte de Mongolia (Goebel et al., 1993).

De esta manera las poblaciones irían adentrándose por vías fluviales diferentes a la vez que se irían diferenciando en sus expansiones, debido al azar y de la deriva genética esbozando una situación compleja ya que la composición haplotípica de los grupos que remontaron cada una de las rutas debió de ser diferente. Además durante estas colonizaciones continentales las poblaciones habrían estado sometidas a los mismos fenómenos evolutivos en el proceso de ascensión de los grandes cauces fluviales (Zakharov 2000). Como consecuencia, el continente asiático se convirtió en un mosaico de poblaciones con una misma raíz genética pero ligeramente diferenciadas (Heyer et al., 2009) en función de su particular historia demográfica y evolutiva (Comas et al., 2004). En este contexto es difícil asignar una fecha concreta a la colonización de Asia central aunque la información arqueológica, por ejemplo el yacimiento de U'st-Karakol y Kara-Bom, certifica la presencia de tecnología lítica en la

región septentrional del continente asiático unos 43.000 AAP. Por tanto se podría considerar que el inicio de la exploración y de los asentamientos del hombre moderno en Asia central se dio 50.000 AAP (Heyer et al., 2009). De hecho, las estimas de coalescencia basadas en la diversidad genética de las poblaciones originarias de la región son acordes a esa datación. Para Asia Central la temporalidad asignada es de 49.000 años (Atkinson et al., 2008).

I.4.V. Europa Occidental

La colonización de Europa por los humanos anatómicamente modernos no se produjo hasta 50.000 AAP, y la información arqueológica indica que la colonización de Europa occidental se llevó a cabo en dos olas culturales sucesivas; la auriñaciense, con una antigüedad estimada de unos 46.000 años, y la gravettiense mucho más moderna datada entre 30.000 y 21.000 AAP (Gamble 1999; Davies 2001). Aunque las expansiones culturales no han de coincidir necesariamente con las migratorias humanas, la información genética apoya la existencia de al menos dos migraciones diferentes que se corresponderían con las nuevas tradiciones culturales introducidas por los humanos modernos. Es importante matizar que en Europa los estudios genéticos presentan un problema añadido, ya que diversos acontecimientos han afectado a la distribución de los marcadores genéticos: el efecto del UMG (Torroni et al., 2001; Roostalu et al., 2007; Holt and Formicola, 2008; Qin et al., 2010), la expansión del neolítico (Currat and Excofier, 2005; Balaresque et al., 2010), las migraciones y movimientos históricos (King et al., 2008; Zalloua et al., 2008; Pereira et al., 2010) que implicaron la entrada de poblaciones procedentes del centro del continente y del norte de África, etc.

Se cree que la cultura auriñaciense del Paleolítico Superior asociada al hombre anatómicamente moderno, se inició en Bulgaria al menos 50.000 AAP (Olszewsky and Dibble, 1994; Otte 2004), extendiéndose hacia el oeste y alcanzando la cornisa cantábrica en la Península Ibérica unos 38.000 AAP (Davies 2001). Estudios filogenéticos llevados a cabo en poblaciones europeas detectan un único haplogrupo de esa antigüedad en Europa, la variante U5 (Richards et al., 2000), pero los estudios realizados sobre la distribución de este haplogrupo lo ubican exclusivamente en Europa y Oriente Medio, estando totalmente ausente en Asia Oriental. Las condiciones climatológicas 51.000 AAP permitieron la apertura de un corredor desde el golfo Pérsico hasta Europa occidental, que coincidiría además con la expansión de la mencionada cultura auriñaciense, y apoyando la hipótesis de que los primeros

Europeos podrían proceder de alguna zona del Creciente Fértil (Refrew and Boyle, 2000).

La segunda ola migratoria coincidiría con el haplogrupo HV con una antigüedad de 33.000 años en Europa, y tendría su origen en Asia Meridional con una profundidad temporal de unos 40.000 AAP (Metspalu et al., 1999). Debe tenerse en cuenta la posibilidad de que estas poblaciones, aunque con un origen similar a las primeras colonizadoras de Europa occidental, siguiesen una ruta alternativa, atravesando el Cáucaso por su extremo más occidental.

I.4.VI. La convivencia con neandertales.

Es importante considerar el momento en que se encontraron *Homo neandertalensis* y *Homo sapiens*, dos grupos humanos procedentes de un mismo tronco evolutivo y muy similares anatómicamente, aunque con otras muchas diferencias, y separadas por unos cuantos cientos de miles de años de evolución independiente. El origen del neandertal está relacionado con *Homo heidelbergensis* que procede de *Homo ergaster*, y un amplio registro fósil confirma su presencia en Eurasia. El periodo de tiempo en que el Neandertal estuvo presente en Eurasia oscila entre 250.000 y 28.000 AAP. Teniendo en cuenta que el hombre anatómicamente moderno colonizó Europa occidental unos 45.000 AAP, la co-existencia en la misma región geográfica de ambas especies fue al menos de 20.000 años. Aunque estas especies se separaron aproximadamente 370.000 AAP (Noonan et al., 2006), desde el punto de divergencia hubo una franja temporal en la que se produjeron diferencias genéticas y anatómicas causadas por las distintas estrategias adaptativas.

En estudios llevados a cabo en secuencias de ADNmt neandertal comparadas con poblaciones europeas, no se han detectado variables lo suficientemente diferenciadas como para defender un rastro de mezcla genética entre ambas especies. A partir de esta información se asume que no hubo mestizaje entre ambas (Green et al., 2008; Briggs et al., 2009), aunque deben considerarse otras posibilidades, como que los cruzamientos entre humanos y neandertales si los hubo, fuesen entre varones neandertales y hembras humanas, y desde esta perspectiva el estudio del ADNmt no sería concluyente a la hora de determinar el mestizaje. Otra posibilidad es que la frecuencia de los cruzamientos entre mujeres neandertales y varones humanos fuese muy baja, y por cuestiones probabilísticas estos linajes desapareciesen o simplemente no se hayan detectado.

Gracias al desarrollo tecnológico de los últimos años, los especialistas han podido recuperar material genético procedente de especies extintas. Los resultados al estudiar el ADN de neandertales, apuntan a que podría haber habido un aporte de hasta el 20% de ADNn de los neandertales al hombre (Noonan et al., 2006). Estas dudas quedarán sin resolver hasta la finalización del Proyecto Genoma Neandertal en desarrollo en la actualidad.

I.4.VII. Consideraciones climatológicas

Tanto la expansión de *Homo sapiens* como la actual distribución de los linajes del ADNmt estuvieron íntimamente asociadas a la geografía y a la climatología, considerándose como punto de referencia el UMG que tuvo lugar aproximadamente 22.000 AAP, y el periodo inmediatamente posterior conocido como Bølling/Allerød entre 15.000 y 12.000 AAP.

Durante el UMG las regiones continentales septentrionales quedaron cubiertas de hielo, circunstancia que modeló la orografía, varió el nivel del mar y ocasionó un cambio en la distribución de las especies, entre las que estaba incluido *Homo sapiens*. En este periodo de profundas oscilaciones térmicas tuvieron lugar tres sucesos simultáneos en el tiempo. En primer lugar una expansión de *Homo sapiens* en Eurasia que coincidió con una difusión demográfica asociada a la aparición y expansión de la cultura Gravetiense, mostrada claramente por las representaciones *network* en forma de estrella de las variantes del ADNmt (Forster et al., 2001). En segundo lugar, la desaparición de los neandertales posibles descendientes de los *Homo erectus* en Europa. Y por último, la colonización América del nuevo continente a través de Beringia claro ejemplo de la importancia del clima como regulador de la dinámica del hombre. En aquella época los continentes asiático y americano estaban prácticamente unidos gracias a que el nivel del mar estaba unos 120 metros por debajo de su nivel actual, emergiendo el puente terrestre entre ambos continentes.

La presión climatológica provocó el desplazamiento de los humanos a regiones más meridionales en busca de climas más cálidos, lo que implicó el refugio de las poblaciones humanas en puntos concretos a la espera de una mejora climática y una nueva expansión. La dinámica de este suceso implicó que desaparecieran las poblaciones del norte, concentrándose en las regiones del sur y generando de esta manera una redistribución (Xue et al., 2006). Los grupos que sobrevivieron en los refugios estuvieron sometidos de forma muy estricta a las fuerzas evolutivas, lo que

ocasionó una diferenciación genética con la desaparición de algunos linajes mitocondriales y la aparición de nuevas variantes (Richards et al., 2000). Este fue el punto de partida para una nueva recolonización post-glaciar que al menos en Europa puede ser cuantificada. Actualmente se considera que entre el 20 y el 34% de los linajes de ADNmt actuales presentes en este continente tienen su origen en un periodo anterior al UMG, por lo que gran parte de las variantes mitocondriales actuales son linajes locales, derivados de haplogrupos ya existentes (Torroni et al., 2001). El 55% de la riqueza mitocondrial europea apareció inmediatamente después del UMG, reflejando las dispersiones post-glaciares procedentes de los refugios. En cuanto a la elevada diversidad de los linajes europeos, es importante tener en cuenta que la importantísima revolución cultural y demográfica que supuso el periodo neolítico unos 8.000 AAP, implicó únicamente la entrada de un 15% de líneas modernas (Richards et al., 2000).

Europa y Asia fueron las grandes áreas geográficas que se vieron afectadas durante el UMG. Su posible efecto sobre la expansión del hombre en América se revisará en profundidad en el apartado I.5. El poblamiento de América. Europa sufrió de forma diferencial el efecto de las oscilaciones del clima debido a su orografía, la zona oriental se vio menos afectada que la occidental, por lo que los tres grandes refugios glaciares se distribuyeron preferentemente hacia el oeste. Se han detectado refugios en la Península Ibérica (Flores et al., 2004; Cherni et al., 2009), Italia y especialmente la región de los Balcanes, en la que en la actual Ucrania se concentró la densidad demográfica más alta del continente (Richards et al., 2000; Torroni et al., 2001).

En la Península Ibérica hubo una importante reducción de la diversidad genética mitocondrial (Pereira et al., 2005), manteniéndose mayoritariamente variantes de los haplogrupos H y V, cuyo análisis representa un interesante ejemplo para aproximarnos a los acontecimientos que se dieron en esta región durante el UMG (Roostalu et al., 2007). El haplogrupo preV versión preglaciar del linaje V, tiene origen oriental y llegó a Europa occidental en torno a 26.400 AAP, y los rastros en las poblaciones parecen indicar que es originario de algún lugar entre el sudeste de Francia y España. Se calcula una antigüedad de 16.300 años y la hipótesis más probable es que se originó en Euskadi en el periodo post-glacial, extendiéndose por Europa durante el proceso de expansión posterior al UMG (Pereira et al., 2005; Roostalu et al., 2007).

En el caso de Italia (Barbujani et al., 1995), la presencia de los Alpes en la región norte, confirió a la península Itálica unas características orográficas y climatológicas

que influyeron en su composición haplotípica. Estuvo ocupada ininterrumpidamente por poblaciones mediterráneas, siendo utilizada como refugio de poblaciones europeas procedentes del septentrión del continente. Más de la tercera parte de los linajes de ADNmt de esta región tienen una antigüedad superior al UMG. Respecto al refugio de los Balcanes es necesario subrayar que se caracterizó por espectaculares expansiones de linajes aparecidos durante el UMG, y por la presencia de una proporción superior al 30% de líneas originarias del periodo preglacial (Soares et al., 2010).

En Asia la situación fue ligeramente distinta por tratarse de una extensión de terreno muy superior a Europa, y a las implicaciones que eso tenía en cuanto a la diversidad de ecosistemas y de oportunidades. Durante el UMG los desiertos de Asia central se proyectaron hacia el sur ocupando un extenso terreno, distinguiéndose dos grandes regiones: Eurasia central y septentrional, y el este y sudeste del continente (Mukherjee et al., 2001). Probablemente la presencia humana no llegó nunca a desaparecer en ninguna de las dos regiones, pero si debió tener grandes oscilaciones demográficas. De hecho hay indicios de que la región del Tíbet o la vertiente meridional de Mongolia permanecieron ininterrumpidamente habitadas durante la época glacial, con grupos de cazadores-recolectores que explotaban mediante la caza grandes rebaños de herbívoros presentes en la zona (Shi et al., 2008; Qin et al., 2010). El registro arqueológico confirma que las poblaciones que habitaron las regiones más norteñas de Mongolia sobrevivieron sobre el *permafrost*, y que su estrategia fue la de contraerse y expandirse en función de la presión que ejercían las temperaturas (Zerjal et al., 2002). Desde el punto de vista demográfico la región del sur y sureste de Asia, que captó parte de las poblaciones humanas procedentes de Asia septentrional, resulta más interesante ya que abandonaron sus tradicionales regiones de asentamiento y tuvieron un modelo que los especialistas llaman dispersión centrípeta (Hill et al., 2006; Zhong et al., 2010). Este modelo conlleva la llegada de poblaciones de cazadores-recolectores mongoloides que probablemente utilizaron las mismas rutas fluviales que miles de años antes sus antecesores utilizaron para colonizar el centro y la región norte del continente. La teoría de la recolonización centrípeta no está avalada por todos los investigadores, y de hecho gran número de arqueólogos la rechazan activamente, pero es complejo profundizar más en el tema. En cualquier caso parece lógico que las poblaciones de la región septentrional de Asia se replegaran hacia las costas y el sur y que en este desplazamiento, se pusieran en contacto con los pueblos ancestrales descendientes de las primeras migraciones humanas en la región

(Quintana-Murci et al., 2004). La co-existencia actual de gran número de linajes pre y post glaciales apoyan la hipótesis de que este nuevo contacto entre poblaciones no produjo la extinción de ninguno de ellos, sino que más bien implicó una co-existencia entre los grupos humanos que se tradujo en mestizaje (Atkinson et al., 2008).

I.4.VIII. La velocidad de expansión

La velocidad media de expansión de los humanos es uno de los parámetros más interesantes a la hora de interpretar los datos moleculares, y está relacionado con múltiples factores como la orografía, la climatología, el desarrollo tecnológico, el acceso a los recursos, etc. Es necesario asumir como válidas algunas fechas estimadas a partir de la información genética y/o arqueológica, pero es llamativo que las velocidades de expansión estén en un rango temporal bastante parecido, en función del continente o la región analizada. Es necesario tener en cuenta que las fechas que se utilizan están basadas en la bibliografía propuesta por determinados autores y que no siempre están en concordancia con la aceptada por otros especialistas.

Si se inicia la revisión en África con los haplogrupos L2/L3 datados entre 60.000 AAP en el oeste y 30.000 AAP en el este del continente (Watson et al., 1997), la velocidad de desplazamiento sería de unos 200 metros por año. Pero si se hace el mismo planteamiento suponiendo que la salida de África se realizó hace 58.000 años y que el límite temporal inferior para la colonización del sudeste oriental asiático es de 31.000 AAP, y asumiendo una distancia de unos 8.000 Km, la velocidad de desplazamiento sería en torno a 300 metros por año. Por último, y dando como válida la fecha de colonización del continente americano entre 20.000 y 25.000 AAP, y que el yacimiento arqueológico de Meadowcroft (Pennsylvania, USA) está bien datado en 18.000 AAP (Adovasio, 1993), y que el de Monte Verde (Chile) tiene 14.000 (Dillehay and Collins, 1988) la velocidad de desplazamiento es en torno a 1.000 metros por año. En todos los casos la velocidad de colonización se mantiene constante y estable lo que confirma cierta regularidad en la expansión de nuestra especie.

I.5. El poblamiento de América

Desde la llegada de los europeos al nuevo continente, el origen de los aborígenes americanos ha sido fuente de innumerables hipótesis, algunas de ellas disparatadas.

Algunos autores relacionaban los pueblos americanos con las tribus perdidas de Israel o con hipotéticos descendientes de la Atlántida, un mítico continente situado en el Océano Atlántico. Otros investigadores intentaron establecer una relación entre las culturas americanas y las mediterráneas utilizando hipotéticas semejanzas culturales. Según estos planteamientos las culturas americanas serían descendientes de pueblos procedentes del entorno mediterráneo como etruscos o egipcios.

Al margen del origen de los pueblos nativos también se han postulado innumerables teorías sobre la ruta que utilizaron los primeros pobladores de América. Se ha postulado una ruta meridional por la que poblaciones australianas colonizaron el nuevo continente a través de la Antártida, dispersándose inicialmente por Tierra del Fuego para acabar expandiéndose por toda América. Para otros autores existe una clara relación entre las culturas de la Polinesia y las andinas, lo que demostraría contactos a través del Océano Pacífico. Para ello habría que asumir un desplazamiento a saltos entre los archipiélagos del Pacífico hasta llegar al poniente sudamericano, lo que implícitamente supone admitir una técnica de navegación muy sofisticada. Recientemente se ha propuesto incluso una colonización de origen europeo, dadas las supuestas semejanzas entre material lítico americano y la cultura solutrense desarrollada en Europa occidental (Figura I.4).



Figura I.4. Algunas de las rutas propuestas para la colonización de América.

Dejando al margen estas teorías pseudocientíficas, desde el siglo XVI y basándose en la sutil observación e interpretación de los rasgos fenotípicos de los aborígenes americanos, algunos autores propusieron que su origen era asiático y que colonizaron el nuevo continente atravesando el estrecho de Bering, la menor distancia que separa Asia de América. El origen asiático y único para la población americana fue propuesto tempranamente por el jesuita español José de Acosta, quien en su “Historia Natural y Moral de las Indias” introdujo este concepto (Acosta 1589). Acosta fundamentaba su teoría sobre la base de caracteres morfológicos existentes en la mayoría de los aborígenes, como el color de la piel y la forma y el color del cabello. Años más tarde, el explorador alemán Alexander Von Humboldt recuperaba esta idea después de reflejar en sus obras las observaciones realizadas en sus viajes por el Nuevo Mundo (Humboldt 1810).

Es evidente que el nuevo continente fue la última región continental en ser poblada por los humanos anatómicamente modernos y su colonización representa un interesante capítulo de nuestra historia. Las teorías sobre su colonización y la dispersión de los primeros grupos humanos han sido el centro de multitud de opiniones contrapuestas. Desde las primeras teorías científicas de Florentino Ameghino sobre el surgimiento del hombre moderno en las llanuras patagónicas a finales del siglo XIX, hasta los análisis moleculares en secuencias completas de ADN mitocondrial en la actualidad, los pasos que dieron los pioneros en el Nuevo Continente han sido sin lugar a dudas ampliamente estudiados.

Florentino Ameghino que propuso en 1921 el nacimiento de la especie humana en el sur del continente americano a mediados de la era Terciaria (Ameghino 1921), basó su tesis en algunos restos arqueológicos de supuesta antigüedad hallados en la provincia de Buenos Aires (Monte Hermoso), y llamó al precursor de nuestra especie *Homunculus patagonicus*.

El antropólogo checo Alex Hrdlicka encabezó en 1925 un movimiento que defendía el poblamiento reciente de América a través del estrecho de Bering. Sus investigaciones se basaban en el asombroso parecido que presentan las poblaciones esquimales y las poblaciones del Nordeste Asiático, ambas muy bien conocidas y estudiadas por él mismo (Hrdlicka 1925). La comparación con otros grupos amerindios con menor apariencia asiática forzó a los investigadores de este movimiento a proponer poblaciones de origen cada vez más distantes, llegando a sugerirse relaciones con grupos de Siberia, Mongolia, Tibet, Japón e incluso Malasia.

En 1943 el antropólogo francés Paul Rivet, fundador del Museo del Hombre de París en 1937, defendía que la población indígena americana era el resultado de cierto número de migraciones, procedentes de Asia, Australia y Melanesia. Rivet tomó como base para su teoría múltiples datos como estatura, color de piel, datos culturales y lingüísticos, tanto de los asiáticos, como de los americanos y oceánicos, con el objetivo de encontrar semejanzas que le permitieran reconstruir las posibles rutas migratorias que habrían seguido los primeros hombres americanos (Rivet 1943).

América, según todos los indicios científicos fue colonizada durante el Pleistoceno tardío aproximadamente entre 10.000 a 30.000 años AAP (Fagan 2004). Sin embargo, en este complejo proceso de las migraciones humanas el momento preciso, el número de migraciones y las rutas de llegada siguen siendo muy controvertidos ya que las evidencias lingüísticas, arqueológicas y genéticas todavía no han dado respuestas coherentes.

1.5.1. La cultura Clovis: presencia humana en el Nuevo Mundo

Para muchos autores la cultura Clovis es la más antigua del continente americano y marca la presencia humana en este continente; no pudo haber humanos en América antes de la aparición de las puntas Clovis, puesto que el hielo bloqueaba el camino. Algunas dataciones señalan que 13.000 AAP, se abrió un amplio pasillo entre los hielos que cubrían Norteamérica y que discurría de noreste a sureste, iniciándose en el actual estrecho de Bering. La antigüedad de este corredor coincidía con la datación de material lítico correspondiente a la cultura denominada Clovis. Esta cultura se basaba en la identificación de puntas de flecha y material lítico distribuidos en varios puntos de Norteamérica y datados en unos 13.000 AAP (Adovasio and Pedler, 2004). Originalmente se hallaron asociadas a restos de mamuts, lo que transformó la visión de los convencionales cazadores en especialistas que colonizaron América al perseguir grandes manadas de estos animales después del UMG. Hay una gran cantidad de obras que abordan la problemática de la cultura Clovis (Boldurian 1991; Meltzer and Michael, 1987; Seebach 2002; Hamilton and Buchanan, 2007; Haynes et al., 2007; Mann 2007; Waters and Stafford, 2007).

Esta teoría que durante muchos años fue el centro del debate en el poblamiento del nuevo continente, se vino abajo al demostrarse falso uno de los principales pilares en la que se sostenía: no había otros asentamientos humanos con datación anterior a la cultura Clovis. En la figura 1.5 puede observarse un mapa con el resumen de los sitios arqueológicos más antiguos del continente. Al margen de los posibles errores en la

datación y la polémica asociada a algunos de ellos, muchos tienen una antigüedad superior al margen propuesto para la cultura Clovis. Además están distribuidos por todo el continente, lo que sugiere la hipótesis de que la presencia de *Homo sapiens* en América fue anterior al desarrollo de la cultura Clovis.

Tradicionalmente uno de los yacimientos más sobresaliente lo constituyó Monte Verde (Chile), que durante muchos años fue considerado la evidencia arqueológica más antigua del continente con una datación supuestamente superior a 43.000 AAP. Esta prueba demostraría una arcaica presencia de nuestra especie en América. Pero en 1997 se confirmó que la presencia humana en este tenía una antigüedad de 14.500 AAP, datación anterior a la de los yacimientos Clovis de entre 12.700 y 13.200 AAP (Waters and Stafford, 2007). Esta fue una de las evidencias que obligó a buscar alternativas a la hipótesis Clovis (Dillehay 2004). La datación para Monte Verde implicaba que el poblamiento de América durante el Pleistoceno tardío había tenido lugar antes de que el corredor quedara libre de hielo, por lo que la ruta costera del Pacífico se convirtió en una alternativa posible y exitosa en la explotación de recursos durante su colonización (Dixon 2001; Dillehay et al., 2008; Goebel et al., 2008).

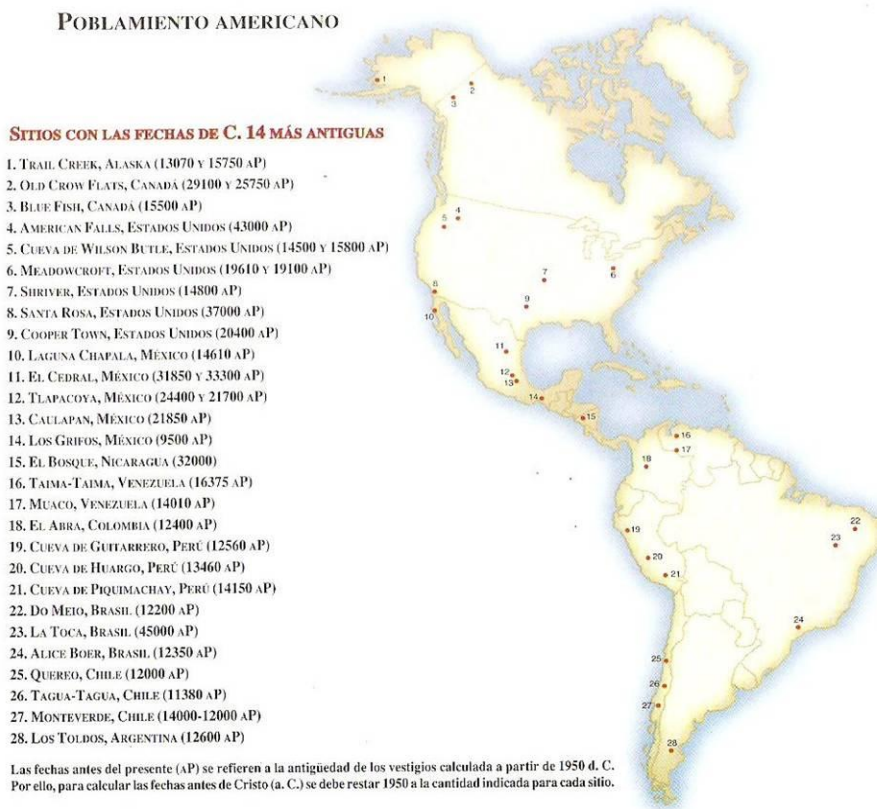


Figura I.5. Mapa resumen donde se muestran la ubicación y antigüedad de los yacimientos arqueológicos más antiguos de América (Solanes Carraro and Vela Ramírez. 2005).

La polémica sobre la importancia de la cultura Clovis continúa en la actualidad con la modelización de los gradientes espaciales de distribución en Norteamérica. Según algunos autores el patrón de dispersión indica un rápido avance con origen en el norte, la velocidad del frente de colonización sugeriría que no había competencia demográfica, ecológica o territorial de otras poblaciones humanas preexistentes (Hamilton and Buchanan, 2007). Otros trabajos no detectan pautas espaciales en la ocupación y defienden que la brevedad del período Clovis indica que esta cultura representa la difusión de una tecnología sobre la preexistente población más que una expansión cultural en sí misma (Waters and Stafford, 2007).

Se han reportado algunos yacimientos pre-Clovis en Norteamérica, aunque su datación y su origen cultural siguen siendo controvertidos sobre todo por la falta de restos humanos o herramientas datadas (Falk 2004). Por este motivo, la problemática se ha trasladado al continente asiático; recientes trabajos en Siberia sugieren que no había asentamientos humanos previos al UMG, por lo que no existía una potencial cultura que diese origen a una rápida colonización al abrirse el corredor de hielo en Beringia (Goebel et al., 2003, 2008; Hoffecker et al., 1993).

Algunos trabajos con diferente enfoque demuestran una presencia humana pre-Clovis. Por ejemplo los estudios centrados en el análisis de ADN de coprolitos y pequeñas herramientas (Gilbert et al., 2009), dan una datación 1.000 años inferior a las anteriormente mencionadas de 13.000 AAP, apoyando nuevamente un poblamiento pre-Clovis del continente americano (Goldberg et al., 2009; Poinar et al., 2009).

I.5.II. Información osteológica

Una prueba directa de la presencia de los humanos en el continente americano lo constituye la información osteológica. Los restos humanos más antiguos detectados en el nuevo continente y representados por cráneos completos o casi completos, presentan una antigüedad inferior a 12.000 AAP. En la tabla I.1 se presenta un listado resumido de estos cráneos acompañado de su origen, antigüedad y referencia bibliográfica.

El cráneo más antiguo del continente corresponde al Peñón III y procede de México, con una datación de 10.755 AAP. El estudio de estos cráneos antiguos se centra en determinar si su morfología corresponde a los actuales grupos aborígenes o se trata de un sustrato morfológico diferente. La mayor parte de los trabajos, tienden a atribuir a estos cráneos una morfología diferente a los rasgos asiáticos que caracterizan a la mayor parte de los grupos indígenas actuales de América. Para muchos autores se

trata de los representantes de una colonización temprana del continente americano, emparentada con grupos del sudeste asiático y Australia y sin continuidad con los actuales aborígenes americanos. Para estos autores sería una prueba de que el nuevo continente fue poblado como mínimo por dos oleadas migratorias, distantes en el tiempo y con morfologías distintas (Jantz and Owsley, 2001; González-José et al., 2003, 2005; Neves and Hubbe, 2005).

Especimen	Procedencia	Cronología	Bibliografía
Spirit Cave	Nevada, USA	9.415 AAP	Dansie, 1997; Jantz and Owsley, 1997
Wizards Beach	Nevada, USA	9.225 AAP	Dansie, 1997; Owsley and Jantz, 1999
Browns Valley	Minnesota, USA	8.900 AAP	Myster and O'Connell, 1997; Jenks, 1937
Peñón III	Texcoco, México	10.755 AAP	González et al., 2001
Metro Balderas	México D.F., México	10.500 AAP	Arce et al., 2003
Cueva del	Huapalcalco, México	10.500 AAP	Pompa y Padilla and Serrano Carreto, 2001
Chimalhuacan	Chimalhuacan, México	10.500 AAP	González-José et al., 2003
Santana do	Santana do Riacho,	9.000 AAP	Neves et al., 2003
Lagoa Santa	Cerca Grande, Brasil	9.330 AAP	Neves et al., 2004

Tabla I.1. Cuadro resumen de los restos de cráneos completos encontrados en América.

En el año 1989 Walter A. Neves, de la Universidad de Saõ Paulo (Brasil) y Héctor M. Pucciarelli de la Universidad Nacional de La Plata (Argentina), propusieron la teoría de los dos componentes migratorios o del componente no mongoloide, desarrollándola mediante el análisis de series de materiales procedentes de Lagoa Santa (Brasil), fechado en 9000 AAP, Tequendama (Colombia) con una antigüedad aproximada de 7500 AAP, Palli-Aike (Chile) fechado en al menos 7.830 años, así como material de Norteamérica y el noroeste asiático (Neves and Pucciarelli, 1989, 1990, 1991). Esta propuesta fue apoyada posteriormente y ampliada por otros autores por observaciones realizadas en restos provenientes del continente americano (Steele and Powell, 1999; Powell and Neves, 1999; Lahr 1995; González-José et al., 2003).

El punto principal de esta teoría es que al ser comparados con series craneológicas de amerindios actuales, estos restos presentan una variabilidad muy elevada, apartándose significativamente del patrón mongoloide y del grupo formado por los amerindios actuales. Los autores proponen que los Paleoamerindios habrían ocupado tempranamente el continente, y posteriormente hicieron aparición los grupos de morfología mongoloide bien diferenciada que probablemente reemplazaron o se mezclaron con los primeros (González-José et al., 2003, 2005, 2007, 2008; Neves et al., 2005; Pucciarelli et al., 2006, 2008). Sin embargo, aunque esta diferenciación no

conlleva necesariamente dos migraciones diferenciadas e independientes (Powell and Neves, 1999), puesto que otros fenómenos evolutivos como deriva génica, presión selectiva, etc. podrían haber dado lugar a la aparición de las poblaciones Amerindias a partir de las Paleoamericanas, los estudios más recientes apuntan a que la separación de ambos grupos se produjo en el continente asiático (Hubbe et al., 2010). Es importante tener en cuenta también que en estos trabajos no se incluyen análisis de muestras de individuos distribuidos en las actuales regiones ocupadas por grupos Na-Dene o Esquimo-Aleutianos.

Aunque este planteamiento explica satisfactoriamente la variabilidad morfológica y los recientes hallazgos arqueológicos, debería complementarse mediante el estudio del ADN de las muestras paleoamericanas, analizando si estos grupos hicieron alguna contribución a los actuales amerindios y en qué grado. Si esto fuera válido, además sería esperable encontrar poblaciones mestizas entre ambos sustratos genéticos. La diversidad genética de los grupos aborígenes actuales no demuestra la existencia de un sustrato paleoamericano, aunque esto puede deberse a varios motivos. A un error de muestreo y a la falta de estudios sistemáticos, o a que aquellos primeros pobladores fueron desplazados sin mestizarse por posteriores migraciones de las que proceden los grupos indígenas actuales.

I.5.III. Familias lingüísticas y poblamiento

En 1986 Joseph Greenberg y colaboradores publicaban un artículo que reunía pruebas lingüísticas, morfológicas y moleculares (Greenberg et al., 1986). Estas tres evidencias apoyaban una división de las poblaciones del Nuevo Mundo en tres grupos: Amerindio 11.000 AAP, Na-Dene 9.000 AAP y Esquimo-Aleutianos 4.000 AAP, defendiendo como hipótesis más plausible que cada una de ellas representara una oleada independiente procedente de Asia. Greenberg defendía esta agrupación basándose en el estudio de palabras y elementos gramaticales (Greenberg 1987).

Según esta teoría la primera oleada dio lugar a los Amerindios, el estrato más importante demográficamente y distribuidos por todo el continente; la segunda a los Na-Dene localizados en Alaska, la costa norte del Pacífico y una pequeña área de Norteamérica; y la tercera a los Esquimo-Aleutianos que poblaron el sur del Ártico. Sin embargo, las posibles correspondencias entre grupos lingüísticos y genética de la población americana actual han sido estudiadas en muchas regiones del continente con resultados ambiguos, no pudiéndose definir con exactitud esta propuesta (Torrioni

et al., 1992; Bonatto and Salzano, 1997a, 1997b; Fagundes et al., 2002; Hunley and Long, 2005).

A pesar de la amplia aceptación inicial de la propuesta de las tres oleadas migratorias esta teoría fue también rechazada y ampliamente discutida por varios autores desde el momento de su publicación (Greenberg et al., 1986). Algunos autores que la aceptaban inicialmente rechazaron posteriormente este supuesto (Bonatto and Salzano, 1997a, 1997b; Cavalli-Sforza et al., 1994; Torroni et al., 1993a, 1993b; Zegura et al., 2004) y otros autores se decantan por trabajar en base a otras clasificaciones lingüísticas como la de Campbell y Mithun (1979). Después de casi 30 años, esta polémica continua abierta (Bolnick et al., 2004; Schroeder et al., 2007).

En el año 2002 se propuso la distinción del continente en cinco áreas: oeste de Norteamérica hacia las Montañas Rocosas, este de Norteamérica desde las montañas hacia el Atlántico, Mesoamérica (comprendiendo el sur de México y Centroamérica), oeste de Sudamérica y el este de Sudamérica (Nichols 2002). Estas regiones formarían una red de conexiones lingüísticas con el centro en el este del golfo de México y/o el este del Caribe, desde donde se dispersaría en varias direcciones. Mesoamérica y el oeste de Norte y Sudamérica formarían una red fuertemente conectada, con el este de Norte y Sudamérica más distantes. Lo importante es que el factor principal de estas relaciones lingüísticas no habría sido la migración externa sino la diversificación autóctona. Algunos de los puntos de vista de Nichols fueron fuertemente criticados por sus colegas (Adler 2000), especialmente Nettle (1999) que discutió la asunción de la autora sobre el aumento de la diversidad lingüística, dando por inválida su teoría y argumentando que la gran diversidad de las lenguas amerindias era debida a la reciente colonización del continente.

I.5.IV. Estudios y evidencias moleculares del poblamiento de América

El interés de los científicos por conocer a fondo el origen de nuestra especie, condujo la mayoría de los primeros trabajos hacia África, Europa y Asia. Todas las investigaciones apuntaban a que el Nuevo Mundo fue el último continente en ser poblado por los humanos anatómicamente modernos, y se acepta que la llegada de las poblaciones fundadoras se realizó de Asia a América a través del estrecho de Beringia durante el Pleistoceno. De todas formas, existen múltiples hipótesis en cuanto al momento exacto, el número de migraciones y el tamaño de estas poblaciones ancestrales.

Con el fin de resolver los primeros procesos de expansión de nuestra especie en América, se han realizado numerosos estudios basados en la información genética aportada por grupos aborígenes de este continente. Así, durante el siglo XX diversas poblaciones indígenas de América fueron caracterizadas para una serie de marcadores clásicos como grupos sanguíneos, proteínas séricas o polimorfismos enzimáticos. Todos ellos confirmaron el origen asiático de los nativos americanos (Crawford 1998; O'Rourke 2006) y mostraron que las frecuencias de los marcadores estaban estructuradas geográficamente en América. Pero no fue hasta 1985, cuando el genetista Douglas C. Wallace realizó una importante contribución al analizar 74 individuos nativos del sudoeste de Estados Unidos (Wallace et al., 1985). Los resultados del ADNmt ponían de manifiesto relaciones con grupos asiáticos. El trascendental estudio de Wallace fue el primero de una serie de trabajos en el continente americano cuya finalidad fue, y sigue siendo, el conocimiento detallado del poblamiento de América para construir la historia biológica de las poblaciones humanas. En 1990 el trabajo de Theodore G. Schurr y sus colaboradores apoyó también el origen asiático de los americanos al descubrir ciertas mutaciones en el ADNmt americano, que describían cuatro linajes cuya procedencia era inequívocamente asiática (Schurr et al., 1990).

Uno de los ejemplos más completos y de mayor utilidad en la investigación molecular de las poblaciones americanas, fue realizado por Satoshi Horai (Horai et al., 1993). En 1993 realizaron un estudio del poblamiento del continente americano, analizando la variación del ADNmt de 72 individuos nativos de América pertenecientes a 16 poblaciones (nueve procedentes de Chile, cuatro de Colombia, una de Brasil, una Maya e indígenas Apache). La comparación de las secuencias obtenidas, de 482 pares de bases, generó como resultado 43 tipos diferentes de ADNmt. Los análisis filogenéticos revelaron que podían ser clasificados en cuatro grandes grupos: I, II, III y IV con una nomenclatura propuesta por el autor. Estos grupos fueron determinados como exclusivos americanos, con una relación muy próxima con los tipos asiáticos identificados hasta entonces. Estos hallazgos fueron la base para postular el poblamiento del Nuevo Mundo a partir de cuatro grandes linajes maternos y asociarlos a varias migraciones. El autor en base a sus resultados fechó la primera migración entre 14.000 y 21.000 AAP.

De forma independiente, otros investigadores (Ginther et al., 1993) llegaron a una conclusión similar a la de Horai, al reconocer cuatro haplogrupos característicos del ADNmt de poblaciones amerindias. En estudios realizados de forma paralela en 1993,

Antonio Torroni (Torroni et al., 1993a, 1993b, 1994a, 1994b) reconoció los haplotipos I, III y IV de Horai, dando sin embargo una nomenclatura distinta refiriéndose a los haplotipos reconocidos como los haplogrupos B, A y C respectivamente (Bailliet et al., 1994). Los siguientes trabajos siguieron la línea de Torroni, apoyando sus postulados e indagando en los haplogrupos fundadores, en las afinidades entre poblaciones de distintas regiones e integrando distinta información generada por otro tipo de análisis (Green et al., 2000; Goicoechea et al., 2001; Jorde et al., 2000; Mesa et al., 2000; Silva-Júnior et al., 2002). A partir de esta información, se publicaron las normas para establecer el orden jerárquico de haplogrupos y sub-haplogrupos (Richards et al., 1998).

El tiempo de diferenciación de estos haplogrupos, unos 21.000 años, apoyaban una procedencia única de estas secuencias sugiriendo un origen común. Evidencias adicionales sugerían una coincidencia de este período de diferenciación con el proceso de colonización del Nuevo Mundo, apoyando la hipótesis de una única y temprana entrada de poblaciones asiáticas en América (Merriwether et al., 1995; Forster et al., 1996; Bonatto and Salzano, 1997a, 1997b; Stone and Stoneking, 1998). Sin embargo, algunos estudios concluyeron que el haplogrupo B presentaba un patrón de diversidad y distribución distinto al resto, por lo que se hipotetizó que podía corresponder a migraciones independientes (Starikovskaya et al., 1998; Torroni et al., 1993a, 1993b). Otros trabajos proponían más de una oleada migratoria (Greenberg et al., 1986; Horai et al., 1993; Torroni et al., 1992, 1994b); aunque la fecha de entrada en el continente americano sigue siendo en la actualidad motivo de acalorados debates (González-José et al., 2003; Dillehay 2003).

La mayoría de los grupos indígenas americanos presentan, por lo tanto, cuatro haplogrupos fundadores, inicialmente nombrados A, B, C, D (Torroni et al., 1993a) y posteriormente caracterizados como A2, B2, C1 y D1, y cinco linajes menores nombrados C4C, D2a, D3, D4h3 y X2a (Achilli et al., 2008; Bandelt et al., 2003; Brown et al., 1998; Fagundes et al., 2008a; Perego et al., 2009; Schurr et al., 1990; Tamm et al., 2007; Torroni et al., 1992, 1993a, 1993b). Los haplogrupos A-D están presentes en el resto del mundo y son frecuentes en Asia (Merriwether et al., 1996; Schurr 2004).

Más recientemente se detectó la presencia de otra variante mitocondrial en América, se trataba del haplogrupo X (Brown et al., 1998). La detección del haplogrupo X implicaba un aumento en la complejidad de la interpretación de los datos procedentes del ADNmt. Se encontraba en el continente americano en una frecuencia baja (Smith et al., 1999) y únicamente en Norte América (Dornelles et al., 2005), y era raro en el

oeste de Eurasia y casi ausente en Siberia (Derenko et al., 2001, 2003; Reidla et al., 2003). Algunos autores incluso han apuntado que el haplogrupo X en Nativos Americanos es menos diverso y tiene un tiempo de coalescencia menor que los haplogrupos A, B, C y D. Así, la frecuencia y el patrón de distribución del haplogrupo X podría estar sugiriendo una migración independiente a América desde Asia o incluso desde Europa (Brown et al., 1998). Estudios más detallados de la presencia de este haplogrupo en América han permitido clasificarlo como un subclado denominado X2a del haplogrupo X. Este último está presente en diferentes poblaciones del mundo, desde África hasta el nuevo continente (Reidla et al., 2003). De cualquier manera el haplogrupo X2a no tiene una relación cercana con ninguna variante de este haplogrupo en Asia ni en el Viejo Mundo. Su posición filogenética sugiere una antigua separación del resto de clados X2, probablemente al principio de su expansión y dispersión desde el Oriente Próximo.

Algunos autores describían un patrón no aleatorio de los linajes A y B en las poblaciones del Nuevo Mundo. El linaje A muestra una clara tendencia a disminuir su frecuencia de norte a sur en el continente, mientras que el linaje B lo hace a la inversa (Merriwether et al., 1995). Aunque las frecuencias de los linajes C y D varían ampliamente en el norte, el centro y el sur de América, exhiben una tendencia general a aumentar hacia el sur (Lalueza et al., 1997).

Los estudios de poblaciones extintas de Tierra del Fuego en la Patagonia argentina (Selknam, Yámana, Kaweskar y Aonikenk), revelan una completa ausencia de los linajes A y B fundadores presentes en amerindios (Lalueza et al., 1997). En contraste, los haplogrupos C y D son detectados en todos los individuos estudiados con frecuencias aproximadas de 38% y 60% respectivamente. Estos resultados unidos a la disminución del haplogrupo A en la mayoría de latitudes del sur del continente americano, y del B por encima de los 55° N en América y Asia, argumentan que los primeros pobladores entraron en América entre 21.000 y 14.000 AAP con ambos linajes ausentes. Es importante tener en cuenta la posibilidad de que estas últimas conclusiones se deban a un error de muestreo ya que el tamaño muestral fue de 60.

Observando la distribución geográfica de los haplogrupos mitocondriales en el continente americano, considerando los estudios de población viva y los resultantes de las extintas, fue posible trazar clinas opuestas en las frecuencias relativas de los linajes del ADNmt. El haplogrupo A dibujaba un marcado descenso en la frecuencia con latitudes sur, mientras que el haplogrupo B como ya se ha comentado, estaba confinado al área central del continente, tal vez debido a sucesivas migraciones sin

este haplogrupo. Ambos estaban ausentes de los grupos de Tierra del Fuego. Teniendo en cuenta la ausencia de presión selectiva, el gradiente geográfico podría ser debido a deriva genética o a tendencias evolutivas diferenciadas. Estas interpretaciones no son excluyentes pero de este modo, esta distribución podría estar reflejando también una migración independiente. A pesar de todo, los fósiles paleoindios del estudio presentan el patrón típicamente mongoloide, lo que sugeriría más de una entrada migratoria desde el continente asiático. Además la presencia exclusiva de los linajes C y D en el extremo del continente, nos lleva a una población ancestral más antigua que la colonizó posteriormente, con los cuatro linajes típicos amerindios en el norte, centro y parcialmente sur de América. La ausencia de los haplogrupos A y B se puede explicar o porque los ancestros no los tenían, o porque los hubieran perdido por deriva genética. Sin ninguna duda, 10.000 años de aislamiento genético habrían contribuido a evitar la llegada de los haplogrupos A y B del *pool* génico a las poblaciones de Tierra del Fuego.

Los estudios más recientes sobre diversidad genética en América muestran que estos cinco haplogrupos (A2, B2, C1, D1 y X2a) estaban ya presentes en América en tiempos precolombinos (Malhi and Smith, 2002; Stone and Stoneking, 1998), y tienen todos ellos profundas raíces en el continente asiático.

Aunque se han propuesto innumerables posibilidades en la colonización de América, en los últimos años se han realizado importantes trabajos que han aportado nuevas visiones y enfoques del poblamiento. Estos nuevos trabajos intentan resolver un problema que se arrastraba hasta ese momento; la posibilidad de que los estudios restringidos a la región control del ADNmt sean muy limitados. Por este motivo empezó a utilizarse también la región codificante completa (Ingman et al., 2000; Macaulay et al., 2005), aunque la tendencia en los últimos estudios es trabajar con secuencias completas de la molécula de ADNmt (Fagundes et al., 2008a; Perego et al., 2009, 2010; Tamm et al., 2007).

En el año 2007 el equipo dirigido por Erika Tamm publicó un estudio de casi 300 ADNmt completos de poblaciones asiáticas y americanas, haciendo una fuerte crítica a los estudios basados únicamente en la región control dada su baja variabilidad (Tamm et al., 2007). Planteaban dos posibles hipótesis sobre la llegada de los primeros pobladores al Nuevo Continente; en primer lugar la que llamaban *Beringian Incubation Model* (BIM), considerando que las poblaciones quedaron aisladas en Beringia antes del máximo glacial, con lo que dispersaron nuevos tipos de ADNmt no presentes en Asia generados en ese período. Y el *Direct Colonization Model* (DCM) considerando

colonización directa y que las poblaciones no llegaron a Beringia hasta que no pudieron pasar, asumiendo que entonces se dispersaron los tipos de ADNmt existentes en Asia. Los autores planteaban además dos premisas: si el grado de diversidad era creciente en el continente, el poblamiento tuvo que ser gradual; y en cambio si existía una distribución uniforme de los haplogrupos, entonces tuvo que darse una colonización rápida.

Los resultados obtenidos mostraban variantes americanas del haplogrupo C (C1b, C1c, C1d) con tiempos de coalescencia muy similares entre ellos y en torno a 13.900 ± 2.700 AAP, quedando en el rango propuesto por otros autores (Eshelman et al., 2003; Schurr and Sherry, 2004). Detectaban además de la presencia de los haplogrupos A2, B2, D1, y el haplogrupo C1a presente en Asia y ausente en América, derivado de la misma población ancestral que C1b, C1c, C1d. Encontraban también D4h3 y C4c, con una distribución este último poco estudiada en el continente (Malhi et al., 2010). Además reportaban el haplogrupo D2 ausente en América y D3 ausente en América excepto en el norte.

Los autores explican los resultados sugiriendo que tras la llegada inicial de los primeros grupos a Beringia entre 20.000 y 25.000 AAP, los ancestros de los nativos americanos pasaron por un período de incubación en Beringia de aproximadamente 15.000 años, después del cual unos se distribuyeron rápidamente por América y otros, portadores del haplogrupo derivado C1a, regresaron hacia el oeste. Estos resultados apoyaban su hipótesis BIM. Este período de aislamiento posterior a la migración fue lo suficientemente largo como para dar lugar a nuevas variantes, y que los haplogrupos fundadores acumularan mutaciones, diferenciándose de los presentes en Asia. Además proponían que durante esa pausa hubo flujo génico bidireccional entre Siberia y el Ártico norteamericano. Además concluían que la distribución de los nuevos linajes de norte a sur era uniforme por lo que la ocupación del continente fue rápida. Mencionaban también el uniforme patrón de distribución y diversidad del marcador autosómico D19S1120 para reforzar las conclusiones obtenidas en su estudio (Schroeder et al., 2007).

En febrero del año 2008 Andrew Kitchen y su equipo, propusieron un modelo de tres fases integrando datos genéticos, arqueológicos, geológicos y paleoecológicos (Kitchen et al., 2008; Mulligan et al., 2008). Basándose en el análisis de datos nucleares y mitocondriales e incorporándolos para aumentar la relevancia, trabajaron con modelos coalescentes de historia demográfica. Esta teoría planteaba una sola migración en tres fases diferenciadas. Una expansión inicial entre 43.000 y 36.000

AAP de grupos asiáticos que colonizaron la región helada de Beringia, situada entre el noreste asiático y el noroeste americano. En la segunda fase las poblaciones permanecieron aisladas en esta región durante un período aproximado de 20.000 años, sufriendo drásticas disminuciones poblacionales e importantes procesos de deriva genética. En el último estadio estas poblaciones, ligeramente diferenciadas de los grupos asiáticos originarios, se extenderían unos 18.000 AAP hacia el nuevo continente dando lugar a las poblaciones amerindias.

Los esquemas bayesianos con los que trabajaron generaron representaciones dinámicas de los cambios en el tamaño de las poblaciones a lo largo del tiempo, indicando que los amerindios pasaron por dos fases de crecimiento entre 40.000 y 15.000 AAP separadas por un largo período de estabilidad poblacional. Los análisis de coalescencia de aislamiento con migración, utilizando datos de poblaciones hermanas para estimar una fecha de divergencia y tamaños poblacionales fundadores, sugerían una expansión amerindia iniciada 15.000 AAP.

Los resultados apoyaban un modelo para el poblamiento del Nuevo Mundo en que los antecesores de los amerindios divergieron del *pool* genético asiático previo en 40.000 AAP, experimentando una expansión gradual de población a medida que se movían hacia Beringia. Después de un largo período de pequeños cambios en el tamaño de las poblaciones, los amerindios se expandieron rápidamente en América 15.000 AAP tanto a través de un corredor en el hielo como por la costa. Esta rápida colonización del Nuevo Mundo habría sido lograda por un grupo fundador con una población efectiva de entre 1.000 y 5.400 individuos.

El modelo de Kitchen presentaba un escenario detallado para el tiempo y la escala de la migración inicial hacia América, refinando las estimas del poblamiento y unificando una teoría para trabajar con datos futuros y métodos analíticos. Pero a pesar de la robustez de esta teoría, no estuvo exenta de severas críticas que pusieron en duda este extenso periodo de permanencia de los grupos pre-amerindios en Beringia (Fagundes et al., 2008b). Las últimas investigaciones que abordan el poblamiento de América y que intentaron resolver estas dudas se centraban en la secuenciación de ADNmt completos para aportar mayor profundidad interpretativa a los resultados, en marzo del mismo año 2008, Nelson Fagundes y sus colaboradores publicaron un trabajo utilizando 86 genomas mitocondriales completos. Decían demostrar que todos los haplogrupos de los nativos mitocondriales, incluyendo el haplogrupo X2a, fueron parte de una única migración fundadora, refutando de este modo los modelos de múltiples migraciones (Fagundes et al., 2008a). La historia demográfica detallada de

las secuencias de ADN mitocondrial estimada con un método bayesiano de coalescencia, indicaba un modelo complejo para el poblamiento de América, en el que la diferenciación inicial de las poblaciones asiáticas terminó con un cuello de botella moderado en Beringia durante el último máximo glacial, entre 23.000 y 19.000 AAP. Los autores proponen que hacia el final del último máximo glacial, empezó una fuerte expansión poblacional, iniciada hace aproximadamente 18.000 AAP y terminada aproximadamente 15.000 AAP. Estos resultados volvían a apoyar una ocupación preClovis del Nuevo Mundo, sugiriendo además un rápido asentamiento del continente a lo largo de una ruta costera.

En este mismo sentido, otros autores han perfilado incluso las rutas por las cuales se llevó a cabo la colonización. Así Ugo Perego y Alessandro Achilli de la Universidad de Pavia publicaron un estudio de 69 genomas mitocondriales completos, llegando a la conclusión de que el nuevo continente se pobló en un periodo comprendido entre 15.000 y 17.000 AAP, y que la colonización se realizó simultáneamente siguiendo dos rutas distintas e independientes. Cada uno de los grupos era portador de un haplogrupo raro distintivo (Perego et al., 2009). De acuerdo a datos ambientales y paleoecológicos, existió un camino representado por el corredor entre dos placas de hielo que se abrió 15.000 AAP (Goebel et al., 2008) o que nunca estuvo cerrado del todo (Fagan 2004). A través de este corredor, en el que habría varios refugios (Loehr et al., 2006), la variante haplotípica X2a se habría movido desde Beringia directamente al este de las Montañas Rocosas quedando restringido a Norteamérica. Esta hipótesis implicaría que la expansión del haplogrupo X2a en América habría ocurrido en la región de las grandes llanuras donde el corredor glacial terminaba, estando en completa concordancia con la distribución y extensión de la diversidad de X2a observada en poblaciones modernas de nativos americanos

La otra variante que marca la línea de expansión es D4h3 que se extendió siguiendo la ruta de la costa del Pacífico ya sin hielo, alcanzando rápidamente Tierra del Fuego. Es muy raro en Norteamérica y frecuente en Sudamérica con variaciones, ya que el frente de expansión debía contener todos los haplogrupos incluido D4h3. En un esfuerzo por aportar más detalles a las investigaciones se incrementó el número de secuencias completas hasta 276, obteniéndose tiempos similares de llegada al nuevo continente e indicando al menos un origen dual para los Paleoamericanos, que darían lugar a prácticamente todas las poblaciones modernas. Los autores concluyen que este origen dual de los primeros americanos es una novedad sorprendente desde el punto de vista genético, y hace plausible un escenario en el que durante un periodo de tiempo

bastante corto se produjeran varias entradas en América a partir de orígenes distintos desde Beringia. Probablemente estos haplogrupos no entraron solos durante el proceso de colonización, por lo que los otros deberían tener patrones de distribución semejantes.

Esto hace plausible que durante estos períodos de hielo y sequía hubiera varias migraciones a pequeña escala de ida y/o vuelta. Incluso algunas zonas debieron de ser pobladas de forma interrumpida, lo que causaría diferentes grupos de poblaciones y un sustrato genético dinámico continuamente re-creado no solamente por deriva sino también por el flujo génico, lo que tendría consecuencias en las hipótesis lingüísticas, ya que esta situación hace aún más improbable la teoría de que una única familia lingüística llegó al continente (Gibbons 1996; Greenberg 1987). Desde este punto de vista para algunos autores la hipótesis amerindia de Greenberg no puede ser ni aceptada ni rechazada (Kaufman and Golla, 2000).

De todas formas, el popular modelo de tres oleadas que considera un único origen es demasiado simplista, los datos del ADNmt sugieren que la población residente de Beringia nunca fue reemplazada del todo, pero tampoco se desarrolló completamente aislada hasta el holoceno medio. Las migraciones del Pleistoceno hacia América debieron ocurrir en una estrecha ventana de no más de 2.000 años, con una sucesión temporal de movimientos por la costa y el corredor libre de hielo (Perego et al., 2009).

Otra pregunta clásica en el poblamiento de América y también frecuente motivo de grandes controversias, es el tamaño de la población fundadora. Los resultados iniciales muestran la existencia de pocos haplogrupos fundadores para el ADNmt y el cromosoma Y, sugiriendo un cuello de botella (Schurr 2004), aunque esta interpretación no está apoyada por los estudios de ADNmt (Bonatto and Salzano, 1997b). Recientes análisis de varios loci genómicos incluyendo ADNmt sugieren que América pudo ser poblada por un efectivo mínimo de 80 individuos (Hey 2005), aunque otros autores apuntan a valores moderados de 500 individuos (Fagundes et al., 2007, 2008a).

I.6. Regiones culturales: Oasisamérica, Aridoamérica y Mesoamérica

La composición indígena de la actual República Mexicana, base de esta tesis, es el resultado de un complejo proceso histórico, cultural y demográfico. Algunas de las etnias estudiadas son las herederas culturales de las civilizaciones más interesantes

que florecieron en América a lo largo de su historia. Entender la importancia que representan estos grupos e interpretar correctamente la información genética que nos aportan, requieren una revisión detallada de la composición, clasificación, historia y cultura de la región que abarca este trabajo. En los siguientes apartados se realizará una descripción de los aspectos culturales, históricos y lingüísticos en los que se desarrollaron, y pueden englobarse y clasificarse, los siete grupos indígenas estudiados.

En la época prehispánica México presentaba un amplio abanico de culturas con desarrollos marcados por las condiciones que ofrecía el territorio. Durante miles de años sus pobladores fueron grupos nómadas de cazadores-recolectores que explotaron los ecosistemas. La combinación de algunas de sus prácticas culturales unidas a las características de las regiones pobladas, propició la adopción de la agricultura como medio principal de subsistencia y base de la economía. Asociada a este cambio de estrategia se encuentran las primeras evidencias de los asentamientos en forma de instrumentos líticos (la etapa lítica se detalla en I. 7. Cronología del México antiguo). Esta época, también llamada *precerámica*, destaca por la fabricación de cerámica que representa el fin de la trashumancia cazadora-recolectora, y da paso a la sedentarización asociada a la agricultura.

La agricultura fue el eje fundamental en la evolución de las poblaciones prehispánicas de México, y el desencadenante del cambio al tipo de vida sedentaria que en último término dio lugar al desarrollo de distintas culturas y, como consecuencia, a la posibilidad de definir diferentes áreas culturales. Estos cambios de estrategia y las diferentes expresiones culturales de los grupos aborígenes de la región llevaron a Kirchhoff a definir dos áreas en el México antiguo: Mesoamérica y Aridoamérica (Kirchhoff 1943, 1952; Campbell et al., 1986; Pailles and Whitecotton, 1995). El desarrollo de las culturas agrícolas de Mesoamérica fue el hito histórico que marca la separación de esta superárea cultural con respecto a Aridoamérica, ocupada por pueblos cazadores-recolectores nómadas y con un paisaje sumamente variado, desde planicies a montañas escarpadas, costas y mesetas, con vegetación y fauna igualmente diversas.

Existe un acuerdo generalizado sobre la ubicación de Mesoamérica (Figura I.6) en la mitad sur de México y una parte de América Central que incluye Guatemala, El Salvador, y el occidente de Nicaragua, Honduras y Costa Rica. No es una entidad estática, en especial la frontera norte, ya que se contrajo y expandió en el transcurso

de su historia, alcanzando su límite máximo con el desarrollo de ciudades relacionadas con Teotihuacán en la planicie central mexicana. Se trata de un área sumamente diversa tanto en variabilidad climatológica como en recursos accesibles a las culturas que en ella se desarrollaron. Si bien, el medio contribuyó a la diversificación inicial de los primeros habitantes que comenzaron a especializarse en ciertas actividades económicas acordes a los recursos disponibles de su entorno, los pueblos se vieron integrados en un proceso civilizador único que adquirió características regionales. Por ello se divide Mesoamérica en áreas culturales, como se detallará más adelante en este apartado.

Las poblaciones ubicadas en Aridoamérica que convivieron con las mesoamericanas, mantuvieron su carácter nómada durante toda la época pre-hispánica y resistieron las condiciones ambientales del norte de México. En esta región coincidían gran variedad de climas, desde el seco y árido en verano hasta un frío intenso durante el invierno, lo que provocaba la escasez de cultivos y la necesidad de actividades como la caza, la pesca, y la recolección.



Figura I.6. Localización y distribución de las tres áreas culturales e históricas del México antiguo.

La gran área cultural denominada Oasisamérica abarcaba el norte de México y el suroeste de Estados Unidos, la América del Norte precolombina, desde Utah al sur de Chihuahua en México; y desde la costa del golfo de California en Sonora, hasta el valle del río Bravo. Oasisamérica fue la última de las superáreas culturales del México precolombino en desarrollarse, y es el resultado de un lento proceso de introducción de las técnicas agrícolas debido entre otros factores a la carencia de agua. Los pobladores de esta área eran sedentarios agrícolas aunque las condiciones climatológicas no eran favorables, por lo que complementaban su subsistencia con la recolección y la caza. Construyeron grandes asentamientos en Chihuahua como Casas Grandes o Paquimé, con sistemas hidráulicos modernos y muy innovadores cuyo apogeo fue entre 1.205 y 1.261 ADP.

La región que abarca este estudio se divide en dos grandes áreas, cada una de ellas relativamente bien definida y relacionada con las estrategias de explotación y supervivencia. Este es un aspecto *a priori* crucial ya que el desarrollo demográfico de las poblaciones, su movilidad y contactos con otras culturas está íntimamente relacionado con los sistemas de explotación. Por este motivo la tesis abarca casi en su totalidad poblaciones del ámbito mesoamericano como indicia su título, pero se han incluido en la medida de lo posible, comunidades de otras regiones culturales que indirectamente expresan estrategias productivas y demográficas distintas a la mesoamericana. La composición genética de los grupos étnicos estudiados debería ser en gran medida, un reflejo de estos comportamientos diferenciales según las áreas culturales. La mayor parte de los grupos indígenas estudiados en esta tesis se distribuyen en Mesoamérica, excepto la población Pima que procede de Oasisamérica.

Pero Mesoamérica geográfica y culturalmente es y ha sido un área muy compleja, en la actualidad se definen diez grandes regiones culturales con identidad propia (Figura I.7). Uno de los planteamientos iniciales de la tesis que influyó directamente en las estrategias de muestreo fue representar al menos con una población, un área extramesoamericana, y además por otra parte, representar el máximo número de estas regiones o áreas culturales dentro de Mesoamérica.

A continuación se describen las áreas culturales de Mesoamérica haciendo especial incidencia en aquellos aspectos que pueden ser relevantes en la historia de los grupos que las habitan.



Figura I.7. Mapa de las regiones culturales de México. Los grupos indígenas estudiados, excluyendo a la región de Centroamérica que escapa al alcance del trabajo, representa a cinco de las siete áreas.

I.6.I. Occidente de México

El Occidente es un área de notable diversidad ambiental, el rico litoral del Pacífico, eficazmente aprovechado por algunos asentamientos, permitía además la comunicación con otras regiones de las que parece proceden algunas tradiciones cerámicas así como la tecnología metalúrgica, de la que aparecieron las evidencias más antiguas en Tomatlán (Jalisco). Fue una zona que favoreció los intercambios de prácticas culturales así como la circulación de materiales y la migración de distintos grupos étnicos desde la periferia hacia el centro mesoamericano.

El grupo Tarasco o Purépecha desarrolló en esta región del poniente mexicano una de las culturas más importantes del Posclásico Tardío entre el 1.200 y el 1.521 ADP, que no solamente dominaba el actual estado de Michoacán y las áreas adyacentes, sino que además desafió la hegemonía Mexica y detuvo su expansión hacia el occidente. La capital estaba establecida en Pátzcuaro (Michoacán) pero posteriormente fue Tzintzuntzan (Michoacán) la más importante del reino Tarasco, con un centro ceremonial de cierta magnitud y que llegó a albergar a más de 40.000 habitantes. La

llegada de los españoles marcó el final de su hegemonía, aunque los descendientes de los Tarascos siguen habitando aquella región y conservan costumbres y prácticas de raíces prehispánicas. Esta área está representada por las muestras del grupo indígena Huichol.

I.6.II. Norte

Inicialmente poblada por grupos sedentarios durante la época en que Mesoamérica alcanzó su mayor expansión hasta el norte, las evidencias más antiguas prueban la domesticación temprana de plantas facilitada por unas condiciones climáticas propicias, la canalización de las corrientes de agua superficial y el almacenamiento del agua de lluvia. Sin embargo, a la llegada de los españoles y tras una época prolongada de sequía, el área norte estaba ocupada por grupos nómadas de cazadores-recolectores a los que llamaban Chichimecas, especialmente característicos por su particular fortaleza y que resistieron con más éxito la invasión de su territorio.

Destaca en la zona noroccidental el desarrollo de la cultura Chalchihuites tras la llegada y colonización de grupos del centro mesoamericano, ya que tras el apogeo de Teotihuacán y el crecimiento de la población, migraron hacia el norte. Esta cultura presenta importantes elementos defensivos en sus construcciones y en las estructuras de los asentamientos. Cuando la agricultura, excesivamente dependiente del buen clima, y el sistema social, sufrieron un colapso en el Norte es posible que algunos miembros de esa cultura, los llamados Toltecas-Chichimecas fundaran la ciudad de Tula junto con los Nonoalcas. El grupo Mayo es el representante de esta región cultural.

I.6.III. Centro de México

Su extraordinaria ubicación y riqueza dio lugar a un diverso conjunto de culturas locales, desde los primeros pobladores de México hasta el desarrollo y pleno dominio de la agricultura. En varias cuevas en la zona de Tehuacán (Puebla) hay evidencias de asentamientos de grupos nómadas, así como pruebas del proceso de domesticación de ciertas especies vegetales como la calabaza, el maíz o el frijol.

Son evidentes las relaciones de esta área con la cultura Olmeca, situada hacia el sureste, así como el desarrollo de lo que sería Teotihuacán (Estado de México) y Cholula (Puebla), ciudades de gran vigor cultural e influencia muy poderosa. En esta

región fueron muy importantes los movimientos migratorios de la población, entre los que destacan las invasiones Toltecas desde el norte que llevaron a la fundación de Tula (Hidalgo) gran urbe con una amplísima zona de influencia, y posteriormente a Tenochtitlán (Distrito Federal), sede de la Triple Alianza, la ciudad de mayores dimensiones de su época fundada por los Mexicas. El grupo Otomí representa esta región del altiplano central mexicano.

I.6.IV. Golfo de México

Las favorables condiciones de esta área permitieron el desarrollo de varios grupos de gran importancia como Olmecas, Huastecos o Totonacos. El desarrollo de Tajín (Veracruz) estuvo facilitado por las condiciones derivadas del declive de Teotihuacán (Estado de México), llegando a tener entre 20.000 y 30.000 habitantes. Los Huastecos que habitaron de forma continuada la sierra Huasteca, son un grupo hablante de una lengua de filiación Maya que se habría separado del tronco principal alrededor del 900 AAP, alcanzando momentos de esplendor dominando grandes extensiones de territorio dividido en señoríos. La región llamada Totonacapan fue la que ocuparon de manera más o menos continuada los Totonacos, grupo emigrado desde el centro de México, con ciudades destacadas por su especial desarrollo y gran tamaño como Cempoala (Veracruz), la primera ciudad que encontraron los españoles a su llegada a México. Los grupos Totonacos y Téenek llegaron a dominar la región, aunque tanto Toltecas como Mexicas hacían continuas incursiones. En este punto hay que hacer una aclaración con respecto al grupo indígena presentado en esta tesis, como son Nahuas de la Huasteca. La clasificación de este grupo es un tema muy complejo y de importantes connotaciones poblacionales. Los grupos Tennek se distribuyen en la región de la Huasteca, entre los estados de San Luís Potosí e Hidalgo. Este grupo, al que se le suponen ancestría Maya pues su lengua y su cultura tiene muchos puntos en común como ya se ha mencionado, no tiene en principio nada que ver con los Nahuas de la Huasteca. Estos últimos son un grupo indígena de filiación cultural Nahuatl y lengua náhuatl. Se trataría de grupos relacionados con los Nahuas que habitan el centro del altiplano mexicano y que muchos autores han intentado relacionar con Aztecas, suponiendo que son sus herederos culturales.

Esta área, al margen de los ya mencionados Nahuas de la Huasteca, está representada por Tepehuas, situados en la región meridional de la Huasteca, entre los estados de Hidalgo y Veracruz.

I.6.V. Guerrero

Tradicionalmente se consideraba Guerrero como una región perteneciente al área Occidental, pero los descubrimientos más recientes han reorientado la división de las áreas culturales mesoamericanas, apareciendo actualmente como una zona independiente y privilegiada por su costa pacífica, en la que se halla el origen del cultivo del maíz. Área de gran influencia Olmeca por sus importantes relaciones comerciales, constituyó para muchas otras culturas como Teotihuacán, Tula y la Triple Alianza (Tenochtitlán, Texcoco y Tlacopan), la fuente de determinadas materias primas. Albergó también la llamada cultura Mezcala de gran importancia por la extensión de objetos de su cultura, altamente apreciados por algunas culturas circundantes y otras muy lejanas, simultáneas y posteriores en el tiempo.

I.6.VI. Oaxaca

La región oaxaqueña fue desde la época mesoamericana una de las más diversas, y una de las áreas en las que se dio la transición de cazadores-recolectores a agricultores sedentarios. En los valles centrales de Oaxaca se originó la civilización Zapoteca, una de las más antiguas que estableció el calendario de 260 días, usado posteriormente por la mayoría de los pueblos mesoamericanos, así como un sistema de escritura propio diferente al Olmeca y al Maya. La hegemonía de San José Mogote, centro ceremonial en la región del Valle, pasó a manos de Monte Albán, la capital clásica de los Zapotecos, que se convirtió en el paradigma de esta civilización y fue la primera gran ciudad que llegó a controlar una amplísima región. Su abandono dio lugar a una serie de centros regionales que se disputaron la hegemonía política.

La Mixteca es también una zona ocupada desde tiempos prehispánicos y aunque las capitales no alcanzaron nunca la magnitud de sus vecinas zapotecas, unificaron las ciudades-estado llegando a ocupar los valles centrales de Oaxaca, donde establecieron relaciones de dominio o alianza con los Zapotecos.

I.6.VII. Costa Sur

Por su amplitud y su prolongada ocupación, la Costa Sur (que comprende el centro y sur de Chiapas y Guatemala) fue una región escenario de múltiples tradiciones culturales, algunas de origen local y otras pertenecientes a la esfera mesoamericana. En esta zona altamente influenciada por las culturas Mayas, se encuentran algunas de las primeras manifestaciones de sedentarización y escritura, así como pruebas de una

eficiente explotación de los recursos marinos. Mantuvieron firmes relaciones comerciales con los Toltecas y con la Triple Alianza, dado lo especial de algunos productos como el cacao.

I.6.VIII. Centroamérica

El área mesoamericana conocida como Centroamérica ocupa una región de clima tropical que los pueblos desarrollados explotaron de forma altamente eficiente. Los primeros contactos con el área nuclear mesoamericana los detectamos por la presencia de influencia Olmeca, y posteriormente recibieron mayor influjo de las culturas del altiplano colombiano. Ejemplo de ello es el desarrollo temprano de la metalurgia en Centroamérica con respecto al resto de los pueblos mesoamericanos. Esta es un área que aunque definida como mesoamericana, queda un poco al margen de la región geográfica cubierta por el trabajo.

I.6.IX. Maya

El área Maya es una de las más amplias de Mesoamérica, y desarrolló una de las tradiciones de mayor duración de la época prehispánica con la transformación de pequeños asentamientos en ciudades de gran tamaño y complejidad, dando lugar a un abanico muy rico de manifestaciones culturales. Estas grandes capitales como Calakmul (Campeche) o Tikal (Guatemala), mantenían relaciones por medio del intercambio de bienes o mediante la alianza de grupos gobernantes con otras ciudades importantes como Palenque (Tabasco). Aunque la gran proliferación de ciudades dio lugar a centros urbanos de la grandeza de Chichen Itzá (Yucatán), acabó siendo el origen del declive Maya por la combinación de diversos factores como guerras internas, cambios climáticos y las diferencias y rivalidades entre distintos grupos de población migrante procedentes del norte de Mesoamérica. Esta región está representada por el grupo indígena Maya.

I.7. Cronología del México antiguo

Una vez descrita la diversidad cultural es necesario profundizar en las raíces históricas de la región. Aunque la tesis no pretende abordar la problemática de la cronología antigua de México sí que es preciso revisarla para comprender con profundidad temporal, cómo los grupos indígenas actuales se han configurado y han heredado las tradiciones etnohistóricas antiguas.

Las primeras evidencias de los asentamientos humanos, ya fueran temporales o de larga duración, se hallan en forma de instrumentos líticos. Este periodo no se considera estrictamente como una fase histórica, y por este motivo es tratado más como una introducción que como un subapartado con identidad. Está caracterizada por una industria lítica arcaica de cantos tallados, lascas musteroideas, bifaces, etc. y representa los primeros estadios del desarrollo cultural de México. Se relaciona inicialmente con los grupos humanos de cazadores-recolectores y la frecuencia de sus hallazgos es muy escasa. Posteriormente, y con la aparición de una agricultura incipiente que se convierte en la base de la economía, y la continuación y especialización de la caza, la tecnología lítica avanza, caracterizándose por diversos tipos de puntas de lanza foliáceas finamente elaboradas.

La etapa lítica se divide en cuatro períodos

- **Arqueolítico**, de 30.000 a 9.500 AAP. Asociada a las primeras evidencias de presencia humana en México. Los instrumentos encontrados denotan cierta destreza tecnológica con un uso orientado a la recolección y procesamiento de vegetales y tratamiento de pequeños animales. Ausencia casi total de puntas de proyectil. Los yacimientos más importantes se encuentran en la Baja California Sur (Laguna Chapala), San Luis Potosí (El Cedral) y el Estado de México (Tlapacoya).
- **Cenolítico inferior**, de 9.500 a 7.000 AAP. Período asociado a un fuerte cambio climático y a cambios asociados a la organización de los grupos. Aumento en la variedad y eficiencia de los instrumentos. Se detecta una continuidad en la importancia de la recolección y un aumento en la caza como método de obtención de recursos animales. Aparecen varios tipos de puntas de proyectil. Destacan los yacimientos de Sonora (El Plomo y Sásabe), Tamaulipas (Lerma y Cueva del Diablo) y Oaxaca (Guilá Naquitz y Cueva Blanca).
- **Cenolítico superior**, de 7.000 a 2.500 AAP. La organización y las pautas de subsistencia continúan sin cambios, aunque se explotan también los recursos costeros. Avances tecnológicos en la fabricación de instrumentos, aparición del pulido y con él, metates y hachas. Objetos de cestería y cordelería. Los yacimientos más importantes se encuentran en Hidalgo (Tecocolote), Estado de México (Chicoloapan), Puebla (Texcal y Coxcatlán).
- **Protoneolítico**, de 5.000 a 2.500 AAP. Solapado temporalmente con la etapa anterior destaca por tener lugar en este periodo el cultivo de plantas (maíz, frijol y

calabaza) y el dominio en la fabricación de instrumentos líticos asociados al procesamiento vegetal, aparecen muelas y morteros. Continúa la cestería y cordelería (redes, textiles y adornos), y aparecen objetos de madera y hueso.

La complejidad del desarrollo de cada uno de los grupos humanos del México antiguo ha llevado al planteamiento de cronologías, modelos de aproximación historiográfica elaboradas por los especialistas para la interpretación de los datos sobre el desarrollo de los pueblos mesoamericanos hasta la llegada de los españoles a América. En base a estas interpretaciones, han surgido cronologías que abordan la historia desde distintos enfoques: enfatizando la continuidad de los procesos culturales e históricos, la evolución de los estados en la región, las innovaciones de urbanización de las sociedades, etc. (Piña Chan 1971, 1997; Duverger 2007). Pero la cronología más utilizada a pesar de las controversias, divide la historia precolombina de esta región del mundo en tres grandes períodos: el Preclásico, el Clásico y el Posclásico (López Austin and López Luján, 1996; López Luján and Manzanilla, 1989). En la figura I.8 se muestra un esquema con los detalles más importantes de esta cronología. El cuadro se divide por regiones culturales y a su vez por periodos cronológicos, y además en cada uno de ellos se hace referencia a los yacimientos más importantes o representativos.

I.7.I. Período Preclásico Temprano

En este período, que engloba aproximadamente desde el 2.500 al 1.200 AAP, se establecieron las bases para el desarrollo de las grandes culturas mesoamericanas, consolidándose dos de sus rasgos esenciales: el sedentarismo y la agricultura como base de la economía de subsistencia. La sedentarización fue fruto de diversos factores entre los que destacan el crecimiento de la población y la necesidad de almacenar las excedencias del cultivo, que llevaron a fijar los asentamientos inicialmente estacionales. Este cambio conllevó además importantes cambios en la organización social. La adopción de la residencia fija estuvo asociada tanto a la domesticación y cultivo de ciertas plantas, como a la explotación eficiente de medios especialmente ricos.

El establecimiento del límite inicial de este período es motivo de controversias ya que se basa en la fabricación de objetos de barro asociados a grupos sedentarios y con control de domesticación de plantas, incluyéndose también en esta etapa zonas de ocupación más temprana. El cultivo de plantas propició cambios que dieron lugar a los

instrumentos de molienda, herramientas, cestería y cerámica, entre otros. La cerámica del Preclásico destaca por su variedad de formas y principalmente por su elaborada decoración que podría estar relacionada con el incipiente desarrollo de los códigos simbólicos mesoamericanos. La elaboración de figurillas es de especial relevancia en esta etapa, reflejando la complejidad social de estas sociedades. Aunque eran esencialmente agrícolas, se siguió recurriendo a la caza, la recolección y la explotación de los recursos de río, lagos y mares, tendencia que tuvo gran relevancia en esta primera etapa de desarrollo.

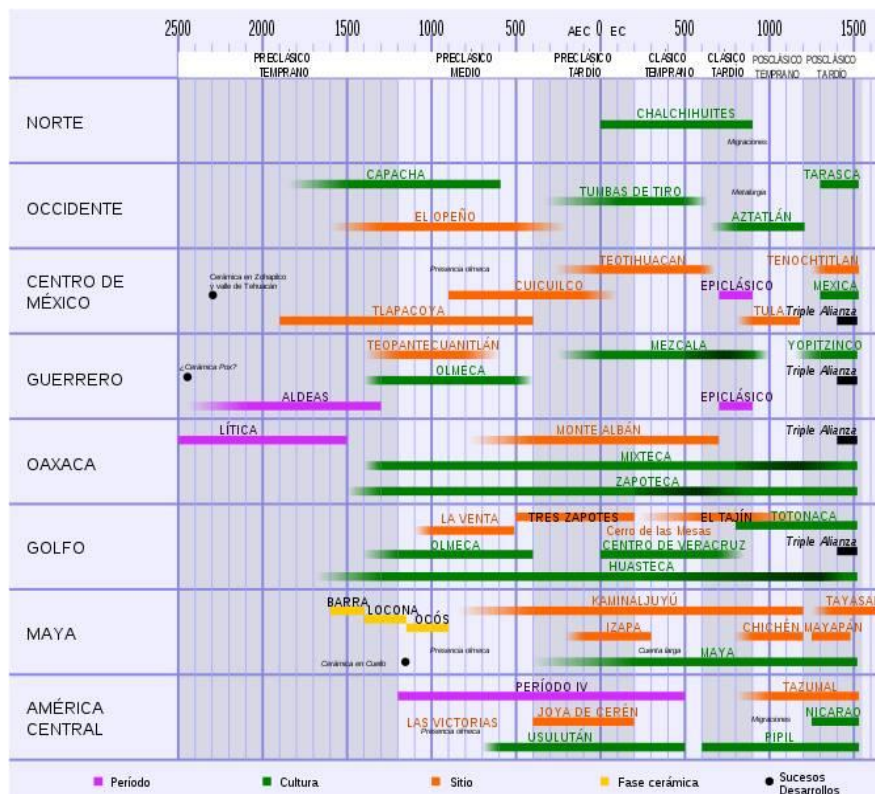


Figura I.8. Esquema que representan las áreas históricas del México antiguo por regiones culturales. Se dan los yacimientos más representativos para cada una de ellas.

Los primeros asentamientos del Preclásico Temprano eran escasos y de dimensiones pequeñas, con casas sencillas, poca población y una organización social simple. Al final de esta etapa algunas de estas aldeas alcanzarían mayores dimensiones dominando en cierto modo sobre las otras, como Tlatilco y Coapexco (Estado de México), Chalcatzingo (Morelos), San José Mogote (Oaxaca) y Chiapa de Corzo (Chiapas).

I.7.II. Período Preclásico Medio

Este período abarca desde el 1.200 al 400 AAP y en él se consolidan algunos cambios que se vieron en la etapa temprana, como por ejemplo el incremento de población derivado de las favorables condiciones asociadas a la agricultura y la sedentarización, que además se ven potenciados por el desarrollo de técnicas para el incremento de la productividad como los sistemas de irrigación o las represas. El aumento de los asentamientos conllevó cierta categorización de los mismos, de forma que algunos tuvieron edificios públicos como Cholula (Puebla), Tlapacoya (Estado de México) y San José Mogote (Oaxaca). Este último tuvo su apogeo en esta etapa, logrando una supremacía regional y siendo probablemente uno de los que participaron hacia el 500 AAP en la fundación de Monte Albán, gran centro urbano con las evidencias más tempranas de estratificación social, que desarrolló especializados productos manufacturados de intercambio con otras regiones y donde hallamos las primeras manifestaciones del juego de pelota. La intensificación de las actividades comerciales como el desarrollo de lugares específicos para determinados productos o en la obtención de determinadas materias primas especialmente apreciadas, como la obsidiana o el jade. En la medida que el intercambio de bienes se intensificó, este se convirtió a la larga en un elemento de cohesión del área mesoamericana.

Destaca también Tlatilco (Estado de México) cuya ubicación favoreció el establecimiento de una base de subsistencia combinando agricultura y el aprovechamiento de otros recursos. Tlatilco se distingue por la gran cantidad de objetos encontrados, constituyendo una de las colecciones arqueológicas más importantes para el estudio de las sociedades del Preclásico Medio. Este asentamiento presenta además una gran estructuración, reflejando la estratificación social así como la complejidad de sus ritos funerarios.

En este período se presentó el surgimiento de la cultura Olmeca, centrada en el golfo de México y con elevados niveles de complejidad, que desarrolló un amplísimo abanico de elementos característicos que fueron ejemplo para las sociedades posteriores. Su cultura fue un hito en la historia mesoamericana, y algunas de sus prácticas culturales serían de uso común entre las épocas posteriores. Una de las características fundamentales de los Olmecas es que fueron capaces de trascender su entorno originario para conformar el primer complejo cultural de alcance mesoamericano, probablemente como producto de una amplia red de relaciones tanto políticas como comerciales y de intercambio. El sur de Veracruz y el límite occidental

de Tabasco es considerado el núcleo de la cultura olmeca por la gran cantidad de vestigios que presenta, por ejemplo las cabezas colosales, elementos exclusivos de esta región, localizadas en su mayoría en Tres Zapotes y San Lorenzo (Veracruz), y en La Venta (Tabasco).

I.7.III. Período Preclásico Tardío

La declinación de la cultura Olmeca marca el principio del Preclásico Tardío que se extiende desde el 400 AAP hasta el 200 ADP. En este periodo la consolidación de las tendencias culturales de la etapa anterior, dan paso al surgimiento de poblaciones de gran tamaño que son la base del desarrollo de las ciudades del Clásico. Los rasgos culturales incipientes en la etapa anterior como el calendario, el juego de pelota o la escritura, se consolidan y alcanzan su madurez, convirtiéndose en elementos característicos del área mesoamericana.

En este período, los asentamientos que ya mostraron gran complejidad anteriormente, establecen las bases para el desarrollo de las grandes culturas del Clásico. Alrededor de las capitales regionales, se crearon redes basadas en el intercambio y estructuradas en jerarquías, con complejas relaciones económicas y políticas, lo que consolidó la cohesión entre las regiones mesoamericanas. Los ejemplos más claros son Tikal y El Mirador (Guatemala), Teotihuacan (Estado de México) con la Pirámide del Sol como gran centro ritual, y Monte Albán (Oaxaca). En ellas se acentúan las diferencias en la jerarquización social y surgen las tradiciones regionales como manifestaciones culturales de cierta complejidad, como la de filiación Zapoteca en los valles centrales de Oaxaca. Algunas tuvieron su apogeo en este período como el núcleo cívico-ceremonial en Cuicuilco (Distrito Federal), mientras que otras como las de la Huasteca y Mezcala trascendieron hasta el Clásico.

El desarrollo de la escritura, más complejo en este período, se convierte en una forma de expresión reservada probablemente a miembros destacados de las clases dominantes. Esta práctica firmemente establecida fue un medio eficaz para registrar y reforzar actos rituales y eventos sociales asociados a la vida de pueblos y gobernantes, como las conquistas y otros, que la convertiría en una de las fuentes de información fundamentales de las sociedades mesoamericanas.

I.7.IV. Período Clásico Temprano

Este período, que engloba desde el 200 al 600 ADP, es uno de los de mayor esplendor del área mesoamericana, ya que alcanzaron su madurez las grandes culturas como Teotihuacan (Estado de México), Cholula (Puebla), Monte Albán (Oaxaca), Tikal (Guatemala) o Kalakmul (Campeche). En estos 400 años diversas regiones de Mesoamérica fueron el escenario del surgimiento de ciudades que alcanzaron no solamente unas dimensiones y complejidad que las convirtieron en centros dominantes de sus regiones inmediatas, sino que establecieron relaciones entre ellas que abarcaban desde el ámbito económico al político. En esta época la zona norte fue colonizada por grupos procedentes del interior de Mesoamérica. El gran aumento demográfico apuntado en los períodos anteriores, era ya una realidad constatable por la gran cantidad de pequeños asentamientos alrededor de las grandes ciudades.

En el centro de México, el desarrollo de Teotihuacan se vio favorecido por su estratégica ubicación que le permitía el acceso privilegiado a los recursos naturales, así como a los depósitos de obsidiana, además de facilitar la comunicación e intercambio con otras regiones. Durante este período se convirtió en la ciudad más importante de Mesoamérica, a lo que contribuyeron factores como sus dimensiones, su gran población, así como la variedad y calidad de sus expresiones culturales. Llegó a albergar más de 100.000 habitantes, concentrando a la práctica totalidad de la población del centro de México. En el proceso de declive intervinieron probablemente revueltas internas combinadas con otros factores, como el debilitamiento de la estructura social así como el surgimiento de otras ciudades que le disputaron su hegemonía. Hasta la llegada de los españoles siguió siendo una gran ciudad, pero ya nunca más de las características e influencia de esta época.

I.7.V. Período Clásico Tardío

Determinado por la caída de Teotihuacan o el declive de Monte Albán y el esplendor de las ciudades-estado de la región Maya, el Clásico Tardío abarca desde el 600 al 900 ADP, y se caracteriza por el ocaso de asentamientos dominantes y el surgimiento de numerosos nuevos centros de menores dimensiones en el centro de México. Este último evento es el centro del llamado periodo Epiclásico, 700 al 900 ADP, que se distingue por la proliferación de estilos regionales derivados de la ausencia de un poblado hegemónico, así como por el surgimiento de nuevos centros de poder cuyo

esplendor sería relativamente corto. En esta época tuvieron lugar las primeras migraciones hacia Centroamérica tal vez debido al florecimiento cultural de las grandes ciudades como Tikal (Guatemala), Palenque (Chiapas) y Calakmul (Campeche), entre otras. Este desarrollo hacia el final de este período se trasladó hacia la península de Yucatán, en la que el control era ejercido por Uxmal (Yucatán).

El fin del clásico se halla marcado por varios hechos, entre ellos la caída de los centros del Epiclásico, el abandono de la mayor parte del norte mesoamericano, el colapso de parte de las ciudades Mayas, y un importante conjunto de migraciones que daría como resultado la fundación de Tula.

Durante el período Clásico, entre el 292 y el 909 ADP, el territorio Maya vivió su época de máximo esplendor. Esta cultura desarrolló ciertos elementos que los distinguen, como el uso del arco falso en la arquitectura y la utilización de un sofisticado sistema de escritura y cómputo del tiempo. Las inscripciones jeroglíficas además de proporcionar información, reflejaban eventos de la vida social como matrimonios entre la realeza de las distintas ciudades, el establecimiento de alianzas políticas y hacían constante alusión a las guerras, tomas de prisioneros y a eventos de carácter militar. Este habría sido uno de los factores principales que produjeron el colapso de estas culturas. Los grandes asentamientos eran núcleos urbanos con construcciones de grandes dimensiones en los que se desarrollaban actividades cívico-ceremoniales y extensas áreas de habitación. Entre las grandes ciudades Mayas destaca Tikal (Guatemala) que llegó a albergar 50.000 habitantes, aunque fueron también importantes Copán (Guatemala) y Palenque (Chiapas).

I.7.VI. Período Posclásico Temprano

Esta época se define por una combinación de distintos rasgos, entre las que destaca la caída de las grandes ciudades del Clásico, la supremacía Tolteca, los movimientos migratorios y la existencia de datos históricos. En este periodo desde el 900 hasta el 1.200 ADP, el área mesoamericana sufrió un cambio radical, prácticamente ninguna de las grandes ciudades del Clásico mantenía su hegemonía aunque algunas mantenían cierto vigor, y las regiones Maya y del norte se encontraban prácticamente abandonadas. Aunque las guerras eran una práctica habitual en los periodos anteriores aquí aparecen intensificadas y asociadas al sacrificio humano, por lo que se generalizó el uso de elementos como el *tzompantli* en el que se exhibían los cráneos

de los sacrificados. La metalurgia es la innovación más significativa de este periodo, con las evidencias más tempranas de su uso en el occidente mexicano.

En este periodo los movimientos de grupos de población tuvieron gran influencia en la conformación política de las distintas zonas mesoamericanas, de hecho en buena parte estas migraciones se derivaron del fuerte proceso de recomposición de los centros del Clásico y el abandono paulatino de la mayor parte de las regiones del norte. Para la identificación de estos movimientos y a diferencia de lo ocurrido en otras épocas, los códices y las crónicas recopiladas tras la llegada de los españoles fueron especialmente útiles. Los flujos migratorios con mayor impacto fueron los procedentes del norte, los conocidos con el nombre genérico de Chichimecas. Entre estos se encontraban varios grupos de habla náhuatl que en algunas regiones provocaron movimientos de la población autóctona hacia Centroamérica.

El límite del Posclásico Temprano está asociado al auge de Tula (Hidalgo), la capital Tolteca. Sin ser tan grande ni poderosa como las capitales predecesoras, fue el centro de un sistema de relaciones comerciales y políticas que alcanzaban lugares tan distantes como Chichén Itzá (Yucatán), para entonces la de más importancia y dimensiones de la región Maya. La ausencia de centros hegemónicos en varias regiones mesoamericanas aparecieron o se consolidaron tradiciones locales, dando lugar al desarrollo de los centros urbanos de la Huasteca (Hidalgo); de Veracruz, donde llegaron los Totonacas hacia el 800 ADP; o de la Mixteca de Oaxaca, de la que se conoce en detalle su historia a partir del 940 APD gracias a los códices prehispánicos conservados.

La cultura Tolteca surgió en una época inestable derivada en gran medida por la llegada de diversos grupos migrantes, los Nonoalcas procedentes de la costa del golfo de México, y los Toltecas-Chichimecas originarios del norte mesoamericano. Estos grupos fundaron Tula (Hidalgo), centro urbano que impuso su hegemonía sobre la región del Centro de México, llegando a tener hasta 60.000 habitantes, siendo capaz de trascender a su tiempo como modelo para las sociedades mesoamericanas que la sucedieron. Es posible que la expansión de Tula se debiera a cierto control militar en algunos territorios, e incluso se cree que podría haber conquistado regiones lejanas por migración de grupos de este origen hacia regiones distantes. Causa de esto sería la presencia de grupos Nahuas en Centroamérica, la Huasteca (Hidalgo) o Chichen Itzá (Yucatán). En este último caso, ambas urbes comparten tal cantidad de elementos (arquitectura, escultura, representaciones de guerreros con la serpiente emplumada,

etc.) que se cree que los gobernantes compartían el mismo linaje. Es importante tener en cuenta que la influencia Tolteca estaba también motivada por las redes de intercambio y comercio de materias primas de las que disponía (maíz, frijol y amaranto), y su decadencia no hizo que perdiera el gran prestigio que poseía. La emigración de los habitantes hacia el 1.100 ADP fue la primera señal que condujo a la caída de esta gran cultura.

I.7.VII. Período Posclásico Tardío

Época también muy esplendorosa, se extiende desde el 1.200 al 1.521 APD y se caracteriza por el desarrollo de dos de las culturas más complejas de la historia mesoamericana: los Tarascos y los Mexicas o Aztecas. En esta época el norte estaba dominado por los grupos nómadas Chichimecas y en el aspecto militar, el México prehispánico estaba dominado por la llamada Triple Alianza, liderada por la capital del grupo Mexica, Tenochtitlán (Distrito Federal). Multitud de materias primas y productos llegaban al centro de México, en base al intercambio comercial o como tributos. Existían puertos de comercio de gran fama como Tulum (Quintana Roo). Además de los Tarascos existían otras regiones independientes habitadas por grupos que gozaban de cierta autonomía como los Huastecos en Tamuín (San Luís Potosí) y los Totonacos en Cempoala (Veracruz). En Oaxaca los Zapotecos seguían manteniendo una presencia importante aunque los Mixtecos estaban en pleno proceso expansivo. Las tierras Mayas se encontraban abandonadas, con Yucatán fragmentado después del apogeo y declive de Mayapán (Yucatán).

El Posclásico Tardío finaliza con la llegada de los españoles a Mesoamérica, territorio que durante miles de años fue escenario de uno de los desarrollos de mayor esplendor del mundo antiguo.

A finales del Posclásico Tardío, gran parte del territorio mesoamericano estaba bajo el dominio de la Triple Alianza: Tlacopan con Texcoco y encabezada por México-Tenochtitlán, la capital de los Mexicas. Según su propia historia los Mexicas procedentes del legendario Aztlán junto con otros dos grupos (por eso recibían el nombre de Aztecas) llegaron al lago de Texcoco, donde fundaron la ciudad. El acelerado desarrollo de la misma y el crecimiento de la población hicieron que llegara a albergar entre 150.000 y 200.000 personas. El sometimiento de más de 400 pueblos y señoríos les facilitaba materias primas y una amplia variedad de bienes, pues tenían establecido un importante sistema de tributos que cada provincia tributaria debía

satisfacer. La supremacía Mexica estaba basada en la elevada capacidad militar así como la eficaz estrategia de demostración pública de su poder, incluyendo todo tipo de ceremonias con sacrificios humanos, por ejemplo. Pero la Triple Alianza se encontraba en estado casi permanente de guerra debido al control de las zonas conquistadas y la distancia que las separaba de ellas, manteniéndose cierto número de señoríos independientes con capacidad militar como Metztlán (Hidalgo), Tlaxcala (Tlaxcala) o Cholula (Puebla). Mención especial merecen los Tarascos que con sus guarniciones preservaron las fronteras de sus dominios, resistiendo a la Triple Alianza en Michoacán, y cuya capital estaba asentada en Tzintzuntzan (Michoacán).

I.7.VIII. Período Colonial

Con la llegada de los españoles las sociedades mesoamericanas sufrieron un proceso de transformación, y aunque preservaron parte de sus costumbres y cultura tuvieron que adaptarse al orden impuesto por los invasores. Este período se extiende desde 1.521 a 1.700 ADP.

En 1.517 y 1.518 se realizaron los primeros viajes de reconocimiento a territorio mexicano, alguno de ellos tierra adentro. Tras la conquista y sometimiento de los Mexicas o Aztecas, el imperio más poderoso de su tiempo, Tenochtitlán se convirtió en los cimientos sobre los que se construyó la ciudad de México, base de la conquista de los pueblos prehispánicos. Los españoles llevaron a cabo una política de poblamiento, por lo que una de las empresas prioritarias fue la evangelización, para lo que fundaron multitud de conjuntos conventuales franciscanos, dominicos y agustinos.

Un número indeterminado de indígenas estimado en varios millones, vivía en América en el momento de la llegada de los españoles. Historiadores y demógrafos coinciden en el hecho de que las enfermedades infecciosas inexistentes en América portadas por los colonizadores europeos (principalmente la viruela), fueron un factor decisivo en el colapso que afectó masivamente a la población indígena americana que no poseía defensas contra las mismas. Entre 1.492 y 1.620 la población indígena experimentó una catástrofe demográfica, estimándose que entre un 75 y un 95% de la población se vio diezmada por las epidemias. También se mencionan como causas de la debacle de estos imperios, la muerte de indígenas causada en la guerra y violencia de la invasión, así como las condiciones de explotación a las que eran sometidos. Este período histórico de contacto entre el Viejo Mundo y América conllevó la invasión y colonización del continente americano por los habitantes de las potencias coloniales,

con millones de emigrantes europeos y africanos durante el siglo XVII, que se establecieron permanentemente en América y mantuvieron considerables grados de mestizaje con la población indígena en algunas regiones.

I.8. Clasificación de las lenguas indígenas

México es una de las áreas con mayor diversidad del planeta, tanto en términos biológicos como culturales, y la suma de las lenguas vivas de México lo sitúa como uno de los cinco países más diversos del mundo. La variación lingüística presenta patrones de distribución espacial que reflejan la dinámica cultural de la época prehispánica y la llegada de los españoles, concentrándose la mayor parte de la población hablante de lengua indígena en la región centro y sur del país. La complejidad es mayor dentro del ámbito cultural mesoamericano, delimitado de manera aproximada por las cuencas de los ríos Lerma-Santiago y Pánuco, mientras que en las áreas hacia el norte la variación lingüística es relativamente baja. Hacia el sur, la región de Oaxaca (incluyendo las zonas adyacentes del sur de Veracruz, sur de Puebla y oriente de Guerrero) destaca como área donde se dispara la diversidad cultural, con una densidad de lenguas y familias comparable a las zonas más complejas del mundo. Aunque los factores que propiciaron esta diversificación extraordinaria no están claros, procesos paralelos en la historia temprana de esas áreas debieron desempeñar un papel relevante.

El interés de abordar el complejo problema de la clasificación lingüística de Mesoamérica se centra en la posibilidad de que este parámetro cultural sea indicativo e incluso responsable, de una subestructura genética en la región. Han sido mucho los autores que han abordado esta relación genética-lengua en diferentes partes del mundo y utilizando diferentes marcadores moleculares (Cavalli-Sforza et al., 1988, 1992; Barbujani and Sokal, 1990; Barbujani 1997; Rosser et al., 2000). Es lícito por lo tanto, ahondar en la rica diversidad lingüística de Mesoamérica y relacionarla, en la medida de lo posible y mediante la aplicación de los métodos estadísticos pertinentes, con la diversidad genética mitocondrial.

Lorenzo Hervás y Panduro podría ser considerado el inicio de la investigación de las lenguas indígenas, y es el autor de uno de los primeros compendios de las mismas. Su *Catálogo de las lenguas de las naciones conocidas y numeración, división y clases de éstas según la diversidad de sus idiomas y dialectos* (1800-1805), constituye el

primer estudio sobre las lenguas indígenas mesoamericanas. En él no se concentra en la descripción de las lenguas sino en la enumeración, localización y clasificación de estas. En esta clasificación genealógica sugiere por primera vez la agrupación de algunas lenguas mesoamericanas en el tronco lingüístico Yuto-nahua (Hervás y Panduro, 1800-1805).

En 1816 Adelung y Vater publicaron *Mithridates*, un estudio general sobre las lenguas del mundo. Esta obra es en realidad un compendio de conocimientos que contiene gran cantidad de datos históricos y geográficos, bibliografías, esbozos gramaticales y listas de vocabularios (Adelung and Vater, 1816). Ese mismo año se publicó el libro de Franz Bopp sobre lenguas indoeuropeas, considerado el primer trabajo lingüístico de tipo comparativo que fue la tendencia del siglo XIX (Bopp 1816).

El trabajo de Wilhelm von Humboldt tuvo un papel relevante en la lingüística general, siendo probablemente responsable del uso del náhuatl como ejemplo de lengua en que las palabras son equivalentes a oraciones (Humboldt 1960). En este trabajo se encuentran referencias a otras lenguas mesoamericanas como el mixteco, el huasteco o el yucateco.

Sin embargo no fue hasta que Francisco Pimentel publicó su obra en los años 1874 y 1875, que se discutió ampliamente el criterio de clasificación de las lenguas tomando una posición intermedia entre el criterio de similitud tipológica y el basado en las semejanzas a través de listas de vocabulario. Esta obra permaneció durante muchos años como una referencia obligada (Pimentel 1874-1875).

En el siglo XIX se hicieron los primeros estudios de lenguas indígenas a la vez que se realizaba en Europa para las lenguas indoeuropeas. Esta tarea fue emprendida por Manuel Orozco y Berra, intelectual mexicano de la segunda mitad de siglo, que aunque siempre negó cualquier competencia en el problema de clasificación de las lenguas, fue tan citado como Pimentel (Orozco y Berra 1864). Para hacer su clasificación se basó en documentación histórica e información contemporánea, e hizo un esfuerzo para diferenciar las lenguas de los dialectos. Sin embargo su principal contribución fue su delimitación de la distribución geográfica de las lenguas. El mapa incluido en su estudio (Figura I.9) constituyó un hito en la representación cartográfica de las lenguas mesoamericanas.

Algunas de las hipótesis clasificatorias de Orozco y Berra fueron retomadas posteriormente por Morris Swadesh a principios del siglo XX, que estimaba el número

de idiomas hablados en el territorio mexicano en 140. La clasificación propuesta por este autor para el análisis de la diferenciación gradual por distancia y aislamiento de los grupos, se basa en la denominada glotocronología, técnica utilizada para definir la genealogía de las lenguas. Una de sus principales herramientas es lo que se denominó lista Swadesh de 100 palabras básicas utilizada no sin controversia, para determinar el grado de separación temporal entre dos o más lenguas (Swadesh 1959).



Figura I.9. Representación cartográfica de las lenguas mesoamericanas según Manuel Orozco y Berra (Orozco y Berra 1864)

La primera clasificación de lenguas indígenas utilizada en el XI Censo General de Población y Vivienda en el año 1990, fue elaborada por Leonardo Manrique Castañeda, lingüista e investigador del Instituto Nacional de Antropología e Historia, con amplia experiencia en la historia, clasificación y ubicación de las lenguas indígenas mexicanas. La clasificación considera las semejanzas y diferencias existentes entre las lenguas indígenas, y a partir de su parentesco lingüístico establece un orden en familias y estas a su vez en grupos (Manrique Castañeda 1988a, 1998b, 1995).

Ethnologue es una publicación impresa y virtual del Summer Institute of Linguistics International, y posee una de las base de datos más extensas con estadísticas de 6.809 idiomas, incluyendo número de hablantes, ubicación, dialectos, afiliación lingüística, etc. A pesar de que actualmente es uno de los inventarios de idiomas

existentes más detallados, la falta de rigor por sus implicaciones religiosas ha sido muy discutida. En esta clasificación definen la familia lingüística como el grupo de lenguas de las que se puede mostrar que tienen una relación genética, considerando veinte familias lingüísticas en la República Mexicana. Igualmente definen tronco lingüístico como el grupo de lenguas que se relacionan genéticamente, pero debido al tiempo transcurrido desde que se produjo la separación es más difícil probarlo (Gordon 2005; Lewis 2009).

En esta tesis se utilizará el *Catálogo de las lenguas indígenas nacionales: Variantes lingüísticas de México con sus autodenominaciones y referencias geoestadísticas* que fue publicado en el Diario Oficial de la Federación el 14 de Enero de 2008 del Instituto Nacional de Lenguas Indígenas (INALI 2007, 2008). Los estudios e investigaciones consultados para la elaboración de este Catálogo corresponden a publicaciones editadas a partir de la década de 1950, aún cuando la discusión sobre la catalogación de las lenguas indígenas mexicanas sobre criterios genealógicos se remonta al último cuarto del siglo XVIII. De las obras consideradas algunas de las más reconocidas son las de Morris Swadesh (1959, 1960, 1963), Lyle Campbell (1979 y 1997), Leonardo Manrique (1988a, 1998b, 1994-1995, 1995), Jorge Suárez (1995) y Raymond Gordon (2005).

El catálogo hace una clasificación de las lenguas indígenas de México desarrollada considerando la diversidad lingüística y basándose en resultados y procedimientos de trabajos de genealogía lingüística, dialectología y sociolingüística, enfocadas a las lenguas indígenas. La clasificación establece las siguientes categorías: familia lingüística, agrupación lingüística, variante lingüística, lengua o idioma. En la figura I.10 se muestra la distribución de todas las familias lingüísticas en el territorio mexicano. Para más detalle sobre la distribución de las lenguas de México y el número de hablantes de cada una de ellas, consultar las figuras VII.1, VII.2, VII.3 del apartado VII. Anexos.

La categoría familia lingüística es la agrupación más global de los niveles de catalogación aplicados, y se define como el conjunto de lenguas cuyas semejanzas estructurales y léxicas se deben a un origen histórico común. Cada una de estas familias está representada por al menos una lengua. Dichas familias ordenadas por su ubicación geográfica de norte a sur son: Álgica, Yuto-nahua, Cochimí-yumana, Seri, Oto-mangue, Maya, Totonaco-tepehua, Tarasca, Mixe-zoque, Chontal de Oaxaca y Huave.

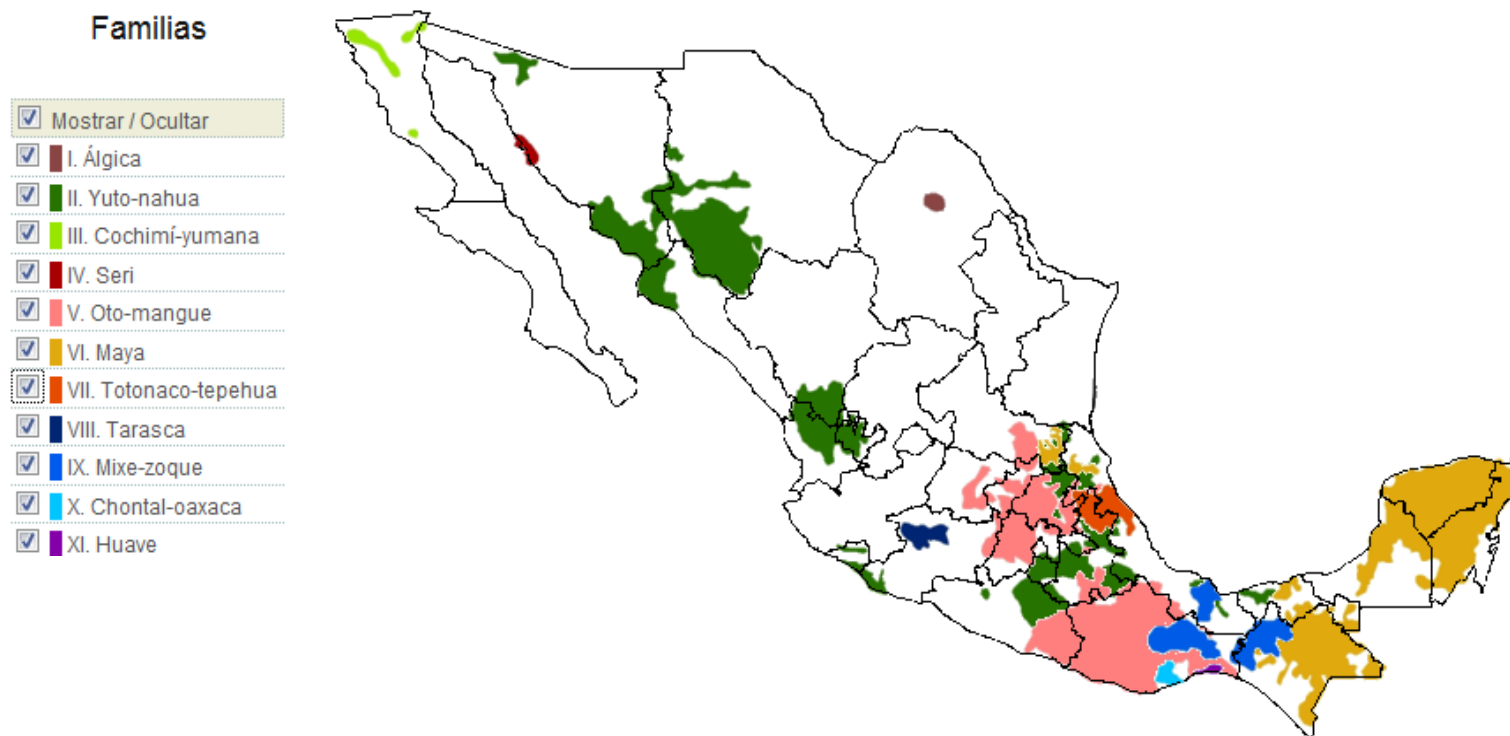


Figura I.10. Distribución de las familias lingüísticas en México (INALI 2008).

La categoría agrupación lingüística ocupa el lugar intermedio en los niveles de catalogación aplicados, y se define como el conjunto de variantes lingüísticas comprendidas bajo el nombre dado históricamente a un pueblo indígena. De acuerdo con esta definición, las agrupaciones lingüísticas catalogadas están relacionadas con un pueblo indígena, pudiendo estar formadas por una o más variantes lingüísticas. Las 68 categorías de agrupación lingüística pueden ser consultadas en la bibliografía (INALI 2007, 2008).

La categoría variante lingüística es la que alcanza el mayor grado de detalle de los niveles de catalogación aplicados en el catálogo, y se define como una forma de habla con diferencias estructurales y léxicas respecto a otras variantes de la misma agrupación lingüística, implicando una identidad distinta de la de otros usuarios de otras variantes. El INALI considera que las variantes lingüísticas deben ser tratadas como lenguas. Esta categoría es empleada habitualmente por la población hablante de lengua indígena para hacer referencia a formas de hablar que contrastan entre comunidades o regiones asociadas con un mismo pueblo indígena. Las variantes lingüísticas representan uno de los indicadores más importantes de la enorme diversidad lingüística y cultural de México. En el catálogo se incluyen 475 autodenominaciones correspondientes a 364 variantes lingüísticas; la diferencia entre ambos valores radica en los casos en que para una variante lingüística existe más de una autodenominación. Estas categorías pueden ser consultadas en la bibliografía (INALI 2007, 2008).

La población hablante de lengua indígena en México no es conocida con precisión, aunque el XIII Censo General de Población y Vivienda del Instituto Nacional de Estadística y Geografía (INEGI 2000) señala que la población de 5 años y más que habla alguna lengua indígena en México, es de 6.044.547 personas. Teniendo en cuenta que el total de población para los Estados Unidos Mexicanos es de 97.483.412, el porcentaje de hablantes de lengua indígena en la población total es del 6,2%. En la figura I.11 se observa la fluctuación de hablantes de lengua indígena desde 1895, año del primer censo de la República Mexicana hasta el último disponible. Es destacable la clara disminución durante el siglo XX de los hablantes de lengua indígena, y la estabilización de los últimos 20 años. Cabe mencionar que en el presente año 2010 se está realizando el XIV Censo General de Población y Vivienda.

**Porcentaje de la población hablante de lengua indígena,
1895 a 2005**

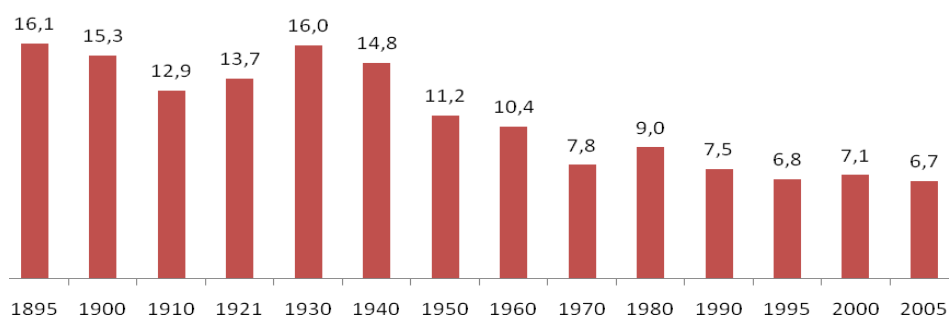


Figura I.11. Porcentaje de población hablante de lengua indígena desde 1895 a 2005. De 1895 a 1921 los datos se refieren a población hablante de lengua indígena total, mientras que de 1930 a 2005 se refieren a población de 5 años o más (INEGI 2005).

A continuación se describirán las familias y agrupaciones en que quedan incluidas cada una de las lenguas de los grupos estudiados en esta tesis, englobando los grupos indígenas en la clasificación mencionada. Es importante tener en cuenta que la clasificación con la que se trabajará en esta tesis es equivalente a las establecidas por Manrique Castañeda (Manrique Castañeda 1988a, 1998b, 1995) y Ethnologue (Gordon 2005; Lewis 2009). El muestreo en algunos casos se hizo fragmentado en distintos municipios y a su vez en distintas comunidades, aspecto detallado en el caso de los Otomíes ya que las áreas de distribución incluyen dos variantes lingüísticas.

Familia Yuto-nahua

Agrupaciones lingüísticas	Total	%	% de la PHLIN
Total	1.648.481	100	27,3
Huichol	30.686	1,9	0,5
Mayo	35.490	2,2	0,6
Náhuatl	1.448.936	87,9	24,0
Pima	741	0,04	0,01

Tabla I.2. Población de 5 años y más hablante de alguna lengua indígena (PHLIN) por agrupación lingüística de la familia Yuto-nahua para el año 2000 (6.044.547 hablantes). Estimación del INALI con base en el XII Censo General de Población y Vivienda (INEGI 2000, INALI 2008).

La tabla I.2 muestra el número de población mayor de 5 años hablante de alguna lengua indígena de la familia Yuto-nahua. El total de indígenas que hablan alguna

variante de esta lengua es de 1.648.481, lo que representa la tercera familia lingüística más hablada por la población indígena mexicana, con un 27,3%. Cuatro de las poblaciones indígenas estudiadas pertenecen a esta familia, siendo el grupo más numeroso Nahuas de la Huasteca (24%) y los menos representados Pimas, con un porcentaje respecto a la población hablante de alguna lengua indígena (PHLIN) de tan solo el 0,01% y el bajísimo número absoluto de 471.

Familia Otomangue

Agrupaciones lingüísticas	Total	%	% de la PHLIN
Total	1.917.015	100	31,7
Otomí	291.722	15,2	4,8

Tabla I.3. Población de 5 años y más hablante de alguna lengua indígena (PHLIN) por agrupación lingüística de la familia Otomangue para el año 2000 (6.044.547 hablantes). Estimación del INALI con base en el XII Censo General de Población y Vivienda (INEGI 2000, INALI 2008).

En la tabla I.3 se dan algunos datos referentes a los hablantes de alguna variante de la familia Otomangue. El otomí es la lengua más hablada por la población indígena de México, con 291.722 hablantes, y representa un 4,8% del total de PHLIN.

Para este grupo indígena se organizó un muestreo en dos zonas distintas del estado de Hidalgo como se detalla más adelante, en el apartado II.4. Toma de muestras en las comunidades indígenas. Con el objetivo de cubrir áreas de distribución de dos variantes lingüísticas diferentes, se trabajó en la sierra Otomí-tepehua donde se habla la variante llamada *hña-hñu*, y en el valle del Mezquital donde se habla la variante denominada *ñuju*.

Familia Maya

Agrupaciones lingüísticas	Total	%	% de la PHLIN
Total	1.788.000	100	29,6
Maya	795.719	44,5	13,2

Tabla I.4. Población de 5 años y más hablante de alguna lengua indígena (PHLIN) por agrupación lingüística de la familia maya para el año 2000 (6.044.547 hablantes). Estimación del INALI con base en el XII Censo General de Población y Vivienda (INEGI 2000, INALI 2008).

Casi 1.788.000 indígenas de México hablan alguna variante de Maya (Tabla I.4) lo que representa un 29,6% del PHLIN. La población Maya está representada por 795.719 individuos, lo que representa un 13,2% del total.

Familia Totonaco-tepehua

Agrupaciones lingüísticas	Total	%	% de la PHLIN
Total	249.469	100	4,1
Tepehua	9.435	3,8	0,2

Tabla I.5. Población de 5 años y más hablante de alguna lengua indígena (PHLIN) por agrupación lingüística de la familia Totonaco-Tepehua para el año 2000 (6.044.547 hablantes). Estimación del INALI con base en el XII Censo General de Población y Vivienda (INEGI 2000, INALI 2008).

La lengua Tepehua representa una infrecuente variante dentro de la familia Totonaco-tepehua. En la actualidad solo hablan esta lengua 9.435 personas, representando un 0,2% del total de la familia lingüística (Tabla I.5).

De la totalidad de grupos representados en la tesis y por su excepcionalidad, destacan el grupo Pima (741 hablantes) y el Tepehua (9.435 representantes).

I.9. Estudios moleculares en la región de estudio

A pesar de que algunas de las poblaciones indígenas estudiadas son inéditas existen estudios moleculares que han intentado explicar, caracterizar y reconstruir el pasado de Mesoamérica. En este apartado se abordará específicamente la información molecular que en la actualidad se tiene de esta excepcional región americana.

El continente americano es poseedor de una rica historia que ha dado lugar a la extraordinaria diversidad genética actual. Mesoamérica constituye un claro ejemplo, la expansión de los grupos amerindios en esta región ejerció una influencia en sus pautas de variabilidad genética. Esta expansión puede generar un patrón de diversidad caracterizado por el exceso de alelos raros producido por el crecimiento rápido de la población, viéndose intensificado si hay un elevado nivel de flujo genético entre poblaciones, y este fenómeno debería poder rastrearse (Fuselli et al., 2003). Los patrones de estructuración genética de las poblaciones actuales son una importante fuente de información en lo referente a historia demográfica reciente. La variación

genética entre poblaciones ha mostrado que los grupos que viven en un mismo continente son relativamente homogéneos (Bamshad et al., 2004). Sin embargo las poblaciones nativas americanas muestran una variabilidad interpoblacional considerable que indica por ejemplo, diferencias entre poblaciones en el norte y sur de América (Bortolini et al., 2003; Mao et al., 2007; Wang et al., 2007).

Las civilizaciones precolombinas de gran parte de México y Centroamérica que formaban Mesoamérica, participaron del mismo universo de creencias y ritos, compartían cierto tipo de vida (sedentarismo) así como organización política y social, y además se ha descrito cierta homogeneidad cultural (Duverger 2007). Todos estos aspectos debieron influir potenciando su homogenización genética y disminuyendo las diferencias interpoblacionales. Sin embargo, también se observa una elevada heterogeneidad lingüística relacionada con factores demográficos y biológicos lo que implicaría cierta diferenciación genética (Wang et al., 2008).

Mesoamérica, fue una región cultural y lingüística en la que las poblaciones nativas ancestrales compartieron lenguajes, tradiciones, costumbres e historia (López Austin and López Luján, 1996) pero no era una región geográfica o una unidad sociopolítica (Suárez 1995), era un área ocupada por poblaciones que compartían unas características culturales como definió Kirchhoff (Kirchhoff 1952). Esto sugiere que Mesoamérica representó un cuello de botella geográfico restringiendo posiblemente el flujo genético entre el norte y el sur del continente, y concentrando en consecuencia gran cantidad de diversidad en esa región. De hecho en la actualidad, el territorio de México que cubre Mesoamérica en su gran mayoría, tiene una de las diversidades lingüísticas y étnicas más ricas de todo el continente. México tiene once familias lingüísticas divididas en 68 grupos lingüísticos que incluyen 291 lenguas vivas y 7 extintas (INALI 2007). La población indígena es de 10,2 millones, representando el 9,6% de la población total del país, en cuyas comunidades más del 30% de la población habla una lengua indígena (INEGI, 2005). Alrededor del 80% de esta población está concentrada en los estados de Chiapas, Oaxaca, Guerrero, Hidalgo, Yucatán, Campeche, Veracruz y San Luís Potosí.

Estudios recientes en poblaciones indígenas mexicanas basados en marcadores moleculares muestran una relativa homogeneidad (Barrot et al., 2005; González-Martín et al., 2008; Gorostiza et al., 2007; Martínez-Cortés et al., 2010; Sandoval et al., 2009) y una elevada diversidad genética (Salzano 2002).

Las poblaciones indígenas actuales son, además de las herederas de una cultura y unas tradiciones únicas en América, el resultado del contacto entre indígenas nativos y poblaciones extranjeras, generándose en principio una reducción demográfica y posteriormente un intenso mestizaje. Un punto de convergencia para generar un patrón de mestizaje conocido como *modelo tri-híbrido*. Con la reducción drástica de la población nativa, se utilizaron gran número de esclavos africanos para reemplazar la mano de obra indígena. Los testimonios históricos documentan la esclavitud africana durante los siglos XVI y XVIII, y aunque los registros de este período son incompletos, se estima que el número de esclavos fue de entre 200.000 y 500.000, aunque probablemente el número real sea superior (Beltrán, 1944). En la actualidad, el número de individuos con herencia africana es desconocido, aunque algunos autores estiman que puede ser superior al 75% de la población (Muhammad, 1995). De hecho, en esta región, europeos, africanos e indígenas se mezclaron con gran intensidad, y muy pocos grupos se mantuvieron aislados. Por ejemplo, Lisker (1981), Lisker et al., (1996) y Zavala et al., (1982) en México, Salzano et al., (1988) en Sudamérica y Crawford (1992) en el continente americano, demostraron que casi todos los grupos indígenas presentan algún grado de mestizaje con europeos y/o africanos. Aunque no es el cometido específico de este trabajo, implícitamente se podrá estimar el componente europeo y/o africano que existe en los grupos indígenas actuales.

Existen diferentes estrategias para abordar la historia biológica del continente. La primera de ellas y más lógica, es la de investigar los grupos indígenas actuales asumiendo que el grado de mestizaje es potencialmente bajo. Otra posibilidad es diseñar investigaciones sobre la población mestiza aprovechando la parte de su pasado amerindio para proyectar en su historia. Ahora bien, estudios basados en poblaciones mestizas dan poca información sobre la estructura y diversidad de las poblaciones nativas (Guardado-Estrada et al., 2009; Juárez-Cedillo et al., 2008; Martínez-Fierro et al., 2009; Peñaloza-Espinosa et al., 2007; Rangel-Villalobos et al., 2010). Hoy por hoy, la mejor estrategia sigue siendo la investigación sobre grupos nativos, sin embargo los proyectos que se están desarrollando actualmente son en base a estudios sobre GWA (Genome Wide Association), que aportarán potencialmente un nuevo punto de vista aunque en la actualidad se centren en enfoques biomédicos (Engelman et al., 2010).

Una alternativa al estudio de la población viva actual consiste en la obtención y análisis de ADN antiguo (ADNa) procedente principalmente de huesos originarios de

yacimientos arqueológicos, aunque no hay muchos estudios fiables en restos de yacimientos mesoamericanos. Uno de los más fiables es el realizado en huesos de Xcaret (Quintana Roo) de la cultura Maya del Clásico Tardío-Posclásico (González-Olivier et al., 2001). En este trabajo los autores detectaron los linajes A (84%), B (4%) y C (8%) y ausencia del linaje D. Estos resultados contrastan intensamente con los obtenidos en otros restos Mayas en Copán (Honduras), donde faltan los linajes A y B (Merriwether et al., 1994, 1997). Por otro lado, los resultados de Quintana Roo son más similares a los hallados en linajes mitocondriales contemporáneos, tanto de la península de Yucatán como de otras poblaciones nativas de origen mesoamericano. La elevada frecuencia del haplotipo A es consistente con la presencia predominante de este linaje en nativos americanos del Norte y Centroamérica. En particular, el haplogrupo A es uno de los más frecuentes en poblaciones actuales con excepción de los Zapotecos, mientras que los haplogrupos B y C, se encuentran en bajas frecuencias, y el haplotipo D está prácticamente ausente. Desde esta perspectiva, las poblaciones antiguas de Xcaret, parecen ser las más cercanas a los Mayas actuales y a otras poblaciones de origen mesoamericano como Mixes y Mixtecos, que a los restos antiguos de Copán. Otros estudios sin embargo, muestran que las poblaciones del suroeste se caracterizan por elevadas frecuencias de haplogrupo B, con muy bajas frecuencias o la completa ausencia del haplogrupo A; mientras que las poblaciones de Mesoamérica presentan elevadas frecuencias del haplogrupo A con frecuencias moderadas y bajas del B (Malhi et al., 2003).

El asentamiento de Copán como ciudad fronteriza podría haber favorecido el grado de flujo génico con poblaciones no Mayas. Cabe tener en cuenta en este caso también, el posible error de muestreo ($n = 9$), y que las frecuencias podrían variar al aumentar el número de individuos analizados. Estas diferencias podrían reflejar el hecho de que las poblaciones de Copán representan grupos Mayas distintos, pero son necesarios datos adicionales de esta población. Sin embargo, recientes estudios en marcadores autosómicos muestran cierta homogeneidad entre grupos Mayas, tanto de México como de Guatemala, con respecto a poblaciones del resto de Mesoamérica (Ibarra-Rivera et al., 2008).

Peñaloza-Espinosa y colaboradores en un intento por caracterizar el ADNmt del sustrato de Mesoamérica analizaron catorce poblaciones indígenas mexicanas: Mixtecos de dos poblaciones, Otomíes, Purépechas, Tzetzales, Tarahumaras, Huicholes y Nahuas de siete procedencias (Peñaloza-Espinosa et al., 2007). El

resultado obtenido fue heterogéneo, aunque los grupos Tarahumara y Huichol no presentaban haplogrupo A. Se detectó también un haplogrupo L en un individuo de una población Nahua mestizada del golfo de México, apoyando estudios previos de la zona (Aguirre-Beltrán 1989).

En el año 2009 se hizo otro estudio en once poblaciones indígenas mexicanas, incluyendo Triquis, Tarahumaras, Purépechas, Otomíes, Nahuas de cuatro procedencias distintas, Mixtecos, Pimas y Mayas. Este estudio de 477 individuos estaba basado únicamente en el análisis de HVI (Sandoval et al., 2009). Casi todos los individuos pudieron ser clasificados en los haplogrupos americanos (A2, B2, C1, D1). El trabajo intentaba entre otras cosas determinar la presencia de los haplotipos D4h3 y X2a como marcadores de las dos posibles rutas migratorias que colonizaron el continente (Perego et al., 2009), y de hecho en su estudio detectaron dos individuos D4h3. Las filogenias de los haplogrupos presentaban forma de estrella como cabría esperar según los autores, en poblaciones relativamente jóvenes que han experimentado episodios de deriva genética y/o cuellos de botella, seguidos de episodios de expansión poblacional. El haplogrupo X2a no está presente en estas poblaciones, que no se encuentran ubicadas en el norte del continente. La comparación de las secuencias obtenidas con las del resto del continente no permite establecer relaciones geográficas concretas. Sin embargo, no hallaron correlación entre el ADNmt y las cuatro familias lingüísticas representadas por las poblaciones elegidas, sugiriendo que la divergencia genética fue un fenómeno anterior a la diversificación lingüística en México (Sandoval et al., 2009).

En un estudio del año 2010 en varios grupos mesoamericanos, los resultados mostraban que las distancias genéticas dentro y entre grupos deberían correlacionar positivamente con distancias lingüísticas y no con distancias geográficas, pero la realidad detectada es que la variación del ADNmt en las regiones estaba fuertemente correlacionada con la geografía y no con la familia lingüística (Kemp et al., 2010).

Ese mismo año, un análisis mediante marcadores autosómicos STR estudiaba las relaciones genéticas y la estructura poblacional de tres grupos nativos mesoamericanos. El trabajo se hizo con cinco muestras poblacionales: Purépechas (Montaña, Valle y Lago), Triquis y Mayas Yucatecos (Martínez-Cortés et al., 2010). Triquis y Purépechas presentaban mayores diferencias con el resto probablemente por su aislamiento geográfico y cultural, y los primeros, mostraban haber recibido un flujo genético de origen distinto. Téenek y Mayas no presentaban grandes diferencias, lo

que según los autores estaría en concordancia con la hipótesis arqueológica de que los Téenek representan un grupo Maya ancestral. El grupo Téenek se localiza geográficamente en la región Huasteca y por esto a menudo son llamados Huastecos; en esta tesis se utilizará el nombre Téenek para no generar confusión con el resto de grupos de esa zona. En conclusión, aunque los resultados sugerirían que el mestizaje europeo habría aumentado las similitudes entre los grupos nativos mexicanos, la diferenciación e inconsistentes agrupaciones por lengua o geografía remarcaban la importancia del efecto fundador y/o la deriva genética como quedaba reflejado en sus actuales relaciones genéticas.

I.10. Zonas y grupos indígenas estudiados

Hasta ahora se han revisado aspectos culturales relacionados con la historia del México antiguo, con la filogenia de las lenguas y con la distribución de regiones culturales, pero no se ha hablado de la actual distribución y composición indígena de la región. Es decir, de la distribución y densidad actual de los grupos indígenas. Es muy difícil hacer un modelo sobre los parámetros que han podido configurar la distribución indígena actual pero su conocimiento puede aportar alguna información importante a la hora de interpretar los patrones de distribución de los grupos indígenas y de sus haplotipos mitocondriales.

El Gobierno mexicano reconoce 62 lenguas diferentes, aunque algunos lingüistas afirman que se conocen más de 100. Estas discrepancias están basadas básicamente en la consideración por parte de los lingüistas de algunas variantes y dialectos como lenguas distintas. La enorme diversidad etnolingüística de México dificulta la identificación de regiones precisas para cada uno de los grupos etnolingüísticos, ya que los diez grupos con mayor volumen de población concentran el 77% de la población indígena, mientras que el 23% restante pertenece a 52 grupos etnolingüísticos distintos, con volúmenes que van de poco más de 226.000 individuos (Huastecos o Téenek), a algunas decenas (Aguacatecos con 59 personas o Kiliwa con 107).

Tradicionalmente se ha utilizado ser hablante de una lengua indígena como criterio diferenciador de los individuos nativos, y las cifras oficiales de la población indígena mexicana se refieren a "Hablantes de Lengua Indígena". Hasta hace unos años los menores de cinco años no estaban incluidos ya que se consideraba que todavía no

eran hablantes fluidos de sus idiomas, lo que llevaba a subestimar la población indígena. Para compensar esta subestimación, en los últimos años se ha añadido a las cifras la “Población menor de cinco años que vive en hogares donde uno de los padres o abuelos habla una lengua indígena”, y a la “Población de cinco años y más que habita en hogares donde, igualmente, al menos unos de los padres o abuelos habla una lengua indígena”. De esta manera se ha conseguido estimar cifras más precisas.

En términos cuantitativos, esta complejidad se expresa en el hecho de que 60% de la población indígena, aproximadamente 6,02 millones de personas, vive en los municipios indígenas (éstos son cuando la población indígena representa más del 40 por ciento del total municipal), mientras que el resto, cerca de 4,2 millones, habita en municipios donde su representación proporcional es minoritaria. Esta distribución poblacional responde tanto a particularidades históricas locales como a la fuerte migración a las grandes ciudades del país o la de los jornaleros agrícolas. Es importante señalar que la población indígena se encuentra presente en una cuarta parte de las localidades del país, casi en la totalidad de los municipios (sólo 30 de ellos no tienen presencia indígena) y en todas las entidades federativas.

Debe tenerse en cuenta que cada miembro de una comunidad comparte con sus vecinos una serie de elementos culturales, como la lengua, el territorio, los cultos religiosos, las formas de vestir, las creencias, la historia, todo lo cual le permite decir que es parecido a ellos y que comparten una identidad común que lo distingue de otros seres humanos, sean los vecinos de la comunidad más próxima, los habitantes no indígenas de su región o de la ciudad, o los extranjeros. Los pueblos indígenas de México tienen identidades culturales y étnicas muy fuertes, la idea que tienen los miembros de una comunidad de formar una colectividad claramente distinta a las otras con las que conviven y que, por lo tanto, cuenta con sus propias formas de vida, sus propias leyes y formas de justicia, sus propias autoridades políticas y su propio territorio.

La definición de estas regiones está basada en aspectos de carácter histórico y cultural, expresados a través de la comprensión de “espacios tradicionales”, lugares ocupados por los pueblos indígenas y en los que permanecen, que comparten características lingüísticas y/o culturales, expresiones de culto similares, visiones cosmogónicas y relaciones de parentesco entre otras. Se han descrito 25 regiones indígenas que concentran en conjunto 8.000.000 de indígenas, aproximadamente el

77% de la población indígena a nivel nacional, los cuales a su vez representan poco más del 43% respecto a población total en los territorios considerados.

Es importante señalar que la integración de una región no depende únicamente de la distribución, el tipo y la intensidad de las actividades productivas, en términos del peso de los fenómenos y dinámicas demográficas en su configuración, sino de la relación entre la dispersión y la concentración de los asentamientos humanos; fenómenos comúnmente asociados a aspectos económicos, culturales y sociales. En este sentido, los patrones de asentamiento de la población indígena no responden únicamente a procesos económicos, sino también a otros de carácter histórico, geográfico y cultural, lo que explica la dispersión de sus localidades. Baste mencionar que entre los pueblos indígenas que habitan en el Gran Nayar y en la Sierra Tarahumara por ejemplo, predominan patrones de asentamiento de enorme dispersión, vinculados a sus particulares sistemas cosmológicos y de creencias, unidas a las condiciones del carácter geográfico imperante, por lo que la mayoría de su población habita en localidades menores a 100 habitantes (Figura I.12).

En general se reconoce la existencia de las regiones indígenas, ligado a la distribución espacial de los pueblos indígenas identificados por las lenguas y el reconocimiento generalizado de los diferentes pueblos con elementos distintivos, como son Los Altos de Chiapas, la Sierra Tarahumara, las Mixtecas, el Gran Nayar o las Huastecas, entre otras. Es por ello que la mayor parte de las propuestas de delimitación de estas regiones se han basado en la distribución espacial de la población indígena, y las diferencias que presentan entre sí se deben a las dificultades de aplicar el criterio de hablantes de lenguas indígenas casi de manera exclusiva para determinar el carácter indígena.

En este sentido, existen acuerdos sobre la importancia de incluir además del criterio lingüístico, otros elementos como son las formas de organización social, la pertenencia a una comunidad indígena, el origen y la historia, así como las identidades que se asumen de forma individual o colectiva. Sin embargo, uno de los problemas que se presentan al tratar de incorporar consideraciones como las anteriores, es la falta de información precisa para establecer criterios generales, debido entre muchas otras razones, a las dificultades implícitas en las posibilidades de traducir muchos de los conceptos de identidad, históricos o culturales en indicadores medibles que además, permitan la comparación entre las diferentes regiones indígenas de México.

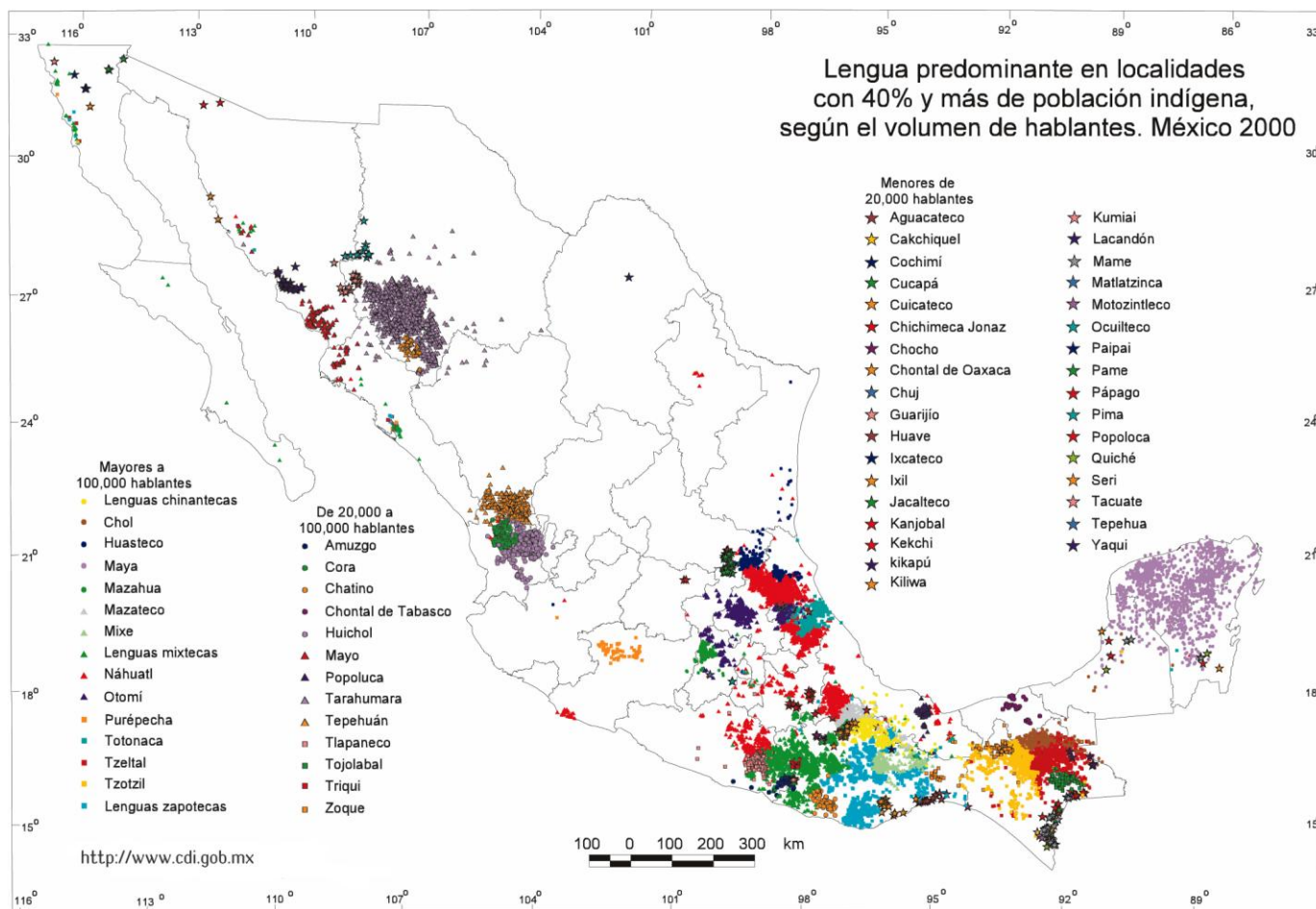


Figura I.12. Mapa de la distribución de los pueblos indígenas de México.

En la propuesta de regionalización de la Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CoNaDePI) utilizada en esta tesis, y basada en el Censo Nacional de Población y Vivienda del año 2000 (censo en el que se basará esta tesis para mantener la lógica del estudio dado que era el disponible en el momento del muestreo de las poblaciones que se incluyen) (INEGI 2000), los criterios metodológicos han tenido como finalidad resolver dos de los principales problemas encontrados al intentar delimitar las regiones indígenas, sobre todo cuando la mayor parte de la información disponible es a nivel municipal.

Por un lado, la enorme heterogeneidad que existe entre los municipios, como los tamaños, la densidad de población total e indígena, las lenguas predominantes en cada uno de ellos. Por el otro, las características geográficas y ambientales, como la morfología y el clima. Debido a esto la mayor parte de las delimitaciones regionales se han ajustado a los límites municipales para facilitar el acceso y acopio de la información. No obstante, las posibilidades que se han abierto en los últimos años para incorporar otros niveles de desagregación, como es la localidad, representan una herramienta alternativa más para mejorar las necesidades de definición de los límites regionales.

En el análisis del mapa representado por la figura I.12 puede apreciarse que el área de distribución de los grupos indígenas no se distribuye uniformemente, más bien todo lo contrario pero se pueden definir más o menos seis regiones. Dos de ellas se distribuyen en el noroeste. El resto de agrupaciones se encuentran en la mitad centro-sur del país. En el centro los asentamientos tienden hacia el Pacífico, mientras que en el sur y a excepción de la península de Yucatán, la población indígena se distribuye uniformemente y casi de forma continua.

La totalidad de las muestras incluidas en el trabajo se han obtenido mediante trabajo directo en las comunidades indígenas y todas ellas se ajustan al protocolo de Helsinki sobre obtención y manipulación de material biológico de origen humano. Igualmente se trabajó de acuerdo a lo dispuesto con el decreto del 14 de Julio de 2008 que regula la salida de muestras de tejido biológico de México con fines de estudios poblacionales. A continuación se realizará una descripción independiente, y siguiendo un orden decreciente de latitud, de cada uno de los grupos indígenas estudiados en la tesis. El objetivo de los siguientes apartados es aportar un marco general en el que ubicar la cultura, geografía y otros aspectos interesantes de cada uno de los grupos, las estrategias de muestreo y otros detalles se desarrollarán en el capítulo II. Material y métodos.

I.10.I. Pimas

Los Pimas se nombran a sí mismos *o'oba* que significa "la gente" o de manera más precisa, aquella que habla la lengua pima, *o'ob no'ok*. El asentamiento más importante se encuentra en la comunidad de Maycoba, municipio de Yécora (Sonora), aunque también tienen una fuerte presencia en Chihuahua, en la sección de Yepachi, perteneciente al municipio de Temósachi, y en la localidad de Mesa Blanca, en el municipio de Madera. Estos son los únicos Pimas que quedan sobre el territorio mexicano (Hope 2006).

Según los historiadores, el pueblo Pima proviene del oeste de Canadá, tierra de la que emigró hacia el sur pasando por California hasta cruzar el río Yuma, donde se dividió y entró a Sonora, donde tras varios conflictos con las comunidades Seris de la zona, logró expulsarlos a la costa del estado (Muñoz Orozco s/f).

En el censo del año 2000 se reportaron 749 hablantes de Pima mayores de cinco años, el total de población en hogares pimas se estima en 1.540 individuos, representando menos del 3% del total de población indígena de la zona. Su relación continuada con los mestizos ha provocado que prácticamente no hallemos población monolingüe (INEGI 2000; Serrano Careto et al., 2006).

La lengua Pima está incluida en el grupo Tepimano dentro de la subfamilia Yuto-nahua del sur, que comprende también las subfamilias Taracahita y Corachol-náhuatl. La familia lingüística Yuto-nahua recibe este nombre a partir de que el yute (Ute) es uno de los idiomas que se hablan en el extremo norte del área ocupada por esta familia (Idaho, Estados Unidos), y de que el náhuatl es el idioma que se habla en el extremo sur de la misma área. La familia Yuto-nahua es una de las que abarca mayor territorio en el continente americano y que comprende uno de los conjuntos de idiomas más numerosos (INALI 2008).

I.10.II. Mayos

La palabra mayo significa "la gente de la ribera". Los Mayos se reconocen a sí mismos como Yoreme. La región Mayo está ubicada en la parte norte del estado de Sinaloa, en el que las comunidades se distribuyen en los municipios El Fuerte, Choix, Guasave, Sinaloa de Leyva y Ahome; y en el sur de Sonora, distribuyéndose en los municipios de Álamos, Quiriego, Navojoa, Etchojoa y Huatabampo. Debido a su proceso histórico, los Mayos han compartido su territorio con mestizos, y en la actualidad prácticamente en todas las comunidades ambos grupos conviven. Los Mayos comparten con los

Yaquis su origen e historia. Los otros grupos indígenas de la zona son el resultado de un proceso de redistribución étnico-regional (Tarahumaras, Guarijíos y Yaquis).

En el censo del año 2000 se reportaron 72.248 hablantes de Mayo mayores de cinco años, el total de población en hogares mayos se estima en 91.261 individuos.

El lenguaje Mayo está incluido también en el grupo Cahita dentro de la subfamilia Taracahita, y a su vez dentro de la familia Yuto-nahua. Dentro de la familia yuto-nahua la lengua Mayo es la más numerosa con el 69,2% del total (INALI 2008).

I.10.III. Huicholes

Llamados Wixárika, los Huicholes son uno de los cuatro grupos que habitan la región conocida como Huicot o Gran Nayar, abarcando territorios de los estados de Durango, Nayarit, Zacatecas y Jalisco, estado en el que son más numerosos. Conviven con los grupos Cora y Tepehuán. No hay un acuerdo entre antropólogos e historiadores sobre su llegada, pero parece que fue posterior a la de los Tepehuanes. No hay información referente a alguna conquista o dominación de los Wixárika por parte de los Tepehuanes o a la inversa. La teoría más aceptada sobre el origen de los Huicholes es la que postula que provienen de la región de San Luis Potosí y que antes de su migración, formaban parte de un grupo Chichimeca llamado Guachichil, cuyo territorio estaba en los actuales estados de Zacatecas, San Luis Potosí y sur de Coahuila, pero también ocupaban partes en el norte de Jalisco (Schaefer and Furst, 1996).

La recolección y consumo ritual del peyote, *Lophophora williamsii*, forma parte de la cultura Huichol. Este cactus es abundante en San Luis Potosí, territorio que fue de dominio central del grupo Chichimeca llamado Guachichiles, que tendrían que haber dado acceso a los Huicholes para su recolección, indicando que los reconocían en cierta manera como parte de su misma etnia y confirmando la historia oral de los Wixárika. Igualmente la similitud entre la lengua Huichol y la de los Guachichiles (ya extinta), es mayor que la Huichol con la de los Cora, sus vecinos actuales (Thomas 1911).

En el censo del año 2000 se reportaron 30.726 hablantes de Huichol mayores de cinco años, principalmente en los municipios de Mezquitic y Bolaños, Jalisco; El Nayar, Tepic y La Yesca, Nayarit, y Mezquital, Durango. El total de población en hogares huicholes se estima en 43.929. La mayoría de los hombres adultos (y en menor grado las mujeres) son bilingües o tienen buenos conocimientos de español (INEGI 2000; Serrano Careto et al., 2006; Neurath 2003).

La lengua Huichol está incluida en el grupo Corachol-náhuatl dentro de la familia Yuto-nahua e incluida en la subfamilia sureña.

I.10.IV. Nahuas

Casi el 30% de los Nahuas de México se encuentran en la Huasteca (hidalguense, potosina, veracruzana o tamaulipeca según su ubicación), y se denominan a sí mismos *macehuale* o *macehualmej*, a su lengua la denominan *méxcatl* aunque también lo denominan náhuatl en muchos lugares de la región. Los Nahuas constituyen el grupo indígena mayoritario en la porción sur de la región Huasteca. Ocupan más de 50 municipios de los estados de San Luis Potosí, Hidalgo y Veracruz, y desde la época prehispánica comparten su territorio con Otomíes, Tepehuas y Totonacos, hacia el sur, y con los Téenek o Huastecos hacia el norte. La relación de los Nahuas con todos ellos es muy estrecha (Valle Esquivel 2006).

Los historiadores distinguen dos incursiones por los Nahuas en la Huasteca, fruto de lo que serían los grupos actuales de habla náhuatl en la región. La primera habría tenido lugar en el período Epiclásico hacia el 800 ADP, mientras que la segunda fue resultado de la expansión Mexica, hacia el 1.400 ADP, asumiendo así que dos oleadas de hablantes de esta lengua llegaron en momentos distintos ocupando territorios ya habitados.

En el censo del año 2000 se reportaron 709.381 hablantes de náhuatl mayores de cinco años, y el total de población en hogares Nahuas de la Huasteca se estima en 1.085.098. Esto representa el 72% de la población hablante de lengua indígena de la Huasteca, y el 27,6% del total de hablantes de lengua náhuatl de México (INEGI 2000; Serrano Careto et al., 2006).

La familia Yuto-nahua es una de las que abarcan mayor territorio en el continente americano y comprende uno de los conjuntos de idiomas más numerosos, siendo la Huasteca es una de las regiones con mayores volúmenes de población Nahua de todo el país. Los once idiomas hablados en México pertenecen a tres de los grupos de la subfamilia sureña, y es en el Corachol-náhuatl dónde se incluye el náhuatl, que también se habla en El Salvador (INALI 2008).

I.10.V. Otomíes

Los Otomíes se nombran a sí mismos *ñähñu* que significa “los que hablan otomí” y este término tiene origen náhuatl (Hekking 1995). Los Otomíes se encuentran dispersos por varios territorios en parte debido a su compleja historia, determinada por

múltiples procesos colonizadores. Con un origen aproximado entre 4.000 y 1.000 AAP (Moreno Alcántara et al., 2006), en tiempos precolombinos la región que ocupan se desarrolló como espacio fronterizo entre los imperios Tarasco y Mexica, sufriendo constantes campañas de conquista infructuosa por parte de estos últimos. Tras la caída de Tula, entre 1.050 y 1250 ADP, los Otomíes se dispersaron por la cuenca de México igual que los grupos Nahuas y los Chichimecas. Cuando la llamada Triple Alianza se apoderó de la zona alrededor del siglo XIII, la distribución de la población sufrió nuevos cambios y algunas zonas céntricas quedaron despobladas. Los Otomíes presentan una afinidad cultural muy marcada con otros grupos con los que han compartido territorio (Barrientos López 2006; Questa Rebolledo and Utrilla Sarmiento, 2006).

En el censo del año 2000 se reportaron 162.798 hablantes de Otomí mayores de cinco años, el total de población en el estado de Hidalgo se estima en 151.743 individuos, representando el 93% del total (INEGI 2000; Serrano Careto et al., 2006).

De acuerdo con la clasificación lingüística, el Otomí, junto con el Mazahua, el Pame, el Ocuilteca, el Chichimeca-jonaz y el Matlatzinca, pertenece a las lenguas Otomianas, las cuales a su vez pertenecen a la rama Otopame de la familia Otomangue. Esta familia lingüística es la más grande y la más diversificada de todo México, y las lenguas que forman parte de esta familia se hablan desde el estado de San Luis Potosí (lengua Pame) hasta el estado de Oaxaca (lengua zapoteca). Actualmente los Otomíes no ocupan un único territorio sino que se encuentran dispersos en varios estados de la República mexicana: Estado de México, Querétaro, Hidalgo, Puebla y Veracruz. El Distrito Federal registra también hablantes de otomí debido a la migración (INALI 2008).

I.10.VI. Tepehuas

Los Tepehuas son un grupo poco conocido y son escasos los estudios que se han dedicado a ellos. La zona en que viven los Tepehuas y en la que encontramos comunidades en las que el Tepehua se habla como lengua materna, está incluida en la Huasteca llamada hidalguense (para diferenciarla de la potosina, la veracruzana y la tamaulipeca). Destaca el municipio de Huehuetla en Hidalgo, Pantepec en Puebla, y Texcatepec, Ixhuatlán de Madero, Tlachichilco y Zontecomatlán, en Veracruz (Hernández Montes and Heiras Rodriguez, 2006).

Según los historiadores la zona Tepehua fue aliada de Tula entre el año 1.000 y el 1.300 aproximadamente, y el área de distribución de los indígenas era más amplia en

sus fronteras norte y oeste. Igualmente es importante destacar que todavía permanece en la memoria colectiva, los movimientos de población a comunidades Totonacas y el hecho de que algunas comunidades adoptaron el idioma Otomí por ser mucho más utilizado en la región (Williams García 1963).

En el censo del año 2000 se reportaron 9.546 hablantes de Tepehua mayores de cinco años, el total de población en hogares Tepehuas se estima en 16.051 individuos, representando menos del 3% del total. Aunque en su gran mayoría hablan español, una parte de la población es todavía monolingüe (INEGI 2000; Serrano Careto et al., 2006).

El nombre de la familia lingüística Totonaco-tepehua proviene de la unión de los nombres de las lenguas que la conforman. Las lenguas de esta familia se hablan en parte de los estados de Hidalgo, Puebla y Veracruz. Anteriormente se consideraba que esta familia lingüística formaba parte de la familia Maya, aunque estudios más recientes han mostrado que no es así (INALI 2008).

I.10.VIII. Mayas

La región Maya es una de las más extensas y pobladas, ya que abarca casi por completo la Península de Yucatán; sin embargo, llama la atención que pese a ser una de las regiones con mayor población indígena, ésta no alcanza el 50 por ciento de la población total. No constituyen un grupo homogéneo, sino un conjunto de etnias con distintas lenguas, costumbres y realidades históricas, pero que comparten rasgos que nos permiten integrarlas en una unidad. Respecto del total de población indígena, los Mayas representan más del 95%, y en la región habita poco más del 97% de la población Maya de México. Se encuentran también hablantes de esta lengua en otros estados del territorio mexicano actual: Tabasco, Chiapas, Veracruz, San Luis Potosí. Es importante recordar que la familia Maya se extiende también hacia Belice, Guatemala y Honduras (Ruz 2006).

Esta área es una de las más amplias de Mesoamérica, y en ella se desarrolló una de las civilizaciones de mayor duración (2.000 AAP a 1.546 ADP) de la época prehispánica con la transformación de pequeños asentamientos en ciudades de gran tamaño y complejidad, que mantenían relaciones por medio del intercambio de bienes o mediante la alianza de grupos gobernantes. El origen del declive estuvo causado por los conflictos generados por pueblos migrantes provenientes del altiplano y del área oaxaqueña.

En el censo del año 2000 se reportaron 1.447.389 hablantes de Maya mayores de cinco años, el total de población en hogares mayas se estima en 1.475.575 individuos. En Yucatán la inmensa mayoría de los indígenas se adscribe a un solo grupo étnico, ya que los mayahablantes representan el 99,6% con 549.532 hablantes de lenguas indígenas mayores de cinco años. En Quintana Roo constituyen el 94,2% con 173.592 hablantes de lenguas indígenas mayores de cinco años (INEGI 2000; Serrano Careto et al., 2006; Ruz 2006).

La lengua Maya forma parte de la subfamilia Yucatecano-maya nuclear de la familia Maya, y aunque es conocida comúnmente como Maya, a menudo se emplea el término Maya yucateco para distinguirla (INALI 2008).

II. MATERIAL Y MÉTODOS

II.1. Muestreo de las poblaciones

Para la realización de esta tesis se analizaron muestras de 7 poblaciones indígenas del México actual, con el objetivo de estudiar su variabilidad mitocondrial y compararlas entre ellas, así como con otras poblaciones indígenas de la región.

En este apartado se describe la metodología desarrollada con el análisis del ADNmt así como las técnicas relacionadas utilizadas en el laboratorio de Genética Forense de GENOMICA S.A.U.

Como base a cualquier estudio poblacional la muestra es esencial, puesto que es el pilar sobre la que se fundamentan el análisis y las conclusiones. En los estudios genéticos, la importancia del muestreo es vital y cualquier error durante el mismo puede invalidar el trabajo en los casos más graves.

Para realizar esta tesis, y en todos los casos, el muestreo se ajustó a la Declaración de Helsinki de la Asociación Médica Mundial sobre obtención y manipulación de material biológico de origen humano. En el momento de la toma de muestras se informó en detalle a cada individuo del objetivo del estudio, resolviendo las dudas y consultas que hubiera. Cada uno de los participantes firmó la autorización pertinente para el uso de su muestra en el estudio.

Igualmente y en relación con lo dispuesto en la Ley de Soberanía Genómica por el decreto del 14 de Julio de 2008 que regula la salida de muestras de tejido biológico de México con fines de estudios poblacionales, se cumplieron todas las especificaciones de la misma, habiendo salido todas las muestras del país con antelación a esa fecha, y habiendo quedado depositadas en el Laboratorio de Genética Humana del Departamento de Zoología y Antropología Física de la Facultad de Biología.

Las muestras utilizadas son excepcionales y únicas en el contexto de investigación en Mesoamérica, conformando la única colección de muestra de esta magnitud e importancia, sobre todo si se atiende a la representatividad de los grupos indígenas.

II.2. Planificación del muestreo

El muestreo fue planificado para ser realizado de forma directa en las comunidades indígenas por considerarse más adecuado y riguroso que otras opciones. Para ello fue necesario un diseño diferenciado para cada una de las distintas regiones teniendo en

cuenta el tamaño poblacional de cada grupo, la accesibilidad geográfica y lingüística (mono o bilingüismo), los contactos disponibles en comunidades indígenas, la presencia de los servicios de salud del Instituto Mexicano del Seguro Social (IMSS), así como la disponibilidad y la estructuración de las comunidades para cada uno de los grupos indígenas elegidos. En todos los casos se asumió la dificultad añadida de hacer los muestreos directamente en las comunidades indígenas lo que complicó tanto la infraestructura necesaria, como el muestreo en sí.

El muestreo se planificó cuidadosamente en cada una de las poblaciones para evitar cometer errores tanto en la toma de muestras como en el completado de las encuestas, punto crítico para la posterior selección de las muestras. Se hizo una encuesta a los individuos (Figura VII.4 del apartado VII. Anexos) que fue rotulada con el mismo código que la muestra correspondiente, vinculándolas. En la encuesta se incluía entre otra información, el origen de los ancestros tanto por vía materna como por vía paterna de las dos generaciones anteriores al individuo muestreado, con el fin de filtrar aquellos individuos emparentados o procedentes de otros lugares. Igualmente se incluyeron en la encuesta preguntas relacionadas con la lengua indígena, el lugar de nacimiento y la presencia de otros familiares en la escuela o en la comunidad.

II.3. Estrategias de muestreo

El muestreo se efectuó siguiendo estrategias distintas para las poblaciones, las campañas se organizaron en función de las características y la infraestructura con la que se contaba. Dependiendo de la comunidad indígena se trabajó con distintos tipos de material biológico, pero en todos los casos se trabajó con guantes y material esterilizado para evitar posibles contaminaciones durante la manipulación y la toma de muestras. Estas fueron almacenadas posteriormente tomando las precauciones necesarias para su correcta conservación.

Nahuas, Otomíes y Tepehuas. En las comunidades del Estado de Hidalgo en las que no estaba presente el IMSS ni se tenían contactos, y en las que era necesario el apoyo institucional ya fuera por su inaccesibilidad o por el monolingüismo de sus integrantes, se recurrió a los estamentos asociados a Gobernación, que brindaron su ayuda en infraestructura, transporte y personal. La toma de muestras se llevó a cabo solicitando la participación de cada individuo “puerta a puerta” con el apoyo de personal de Gobernación del Estado de Hidalgo. Las encuestas se hicieron con la ayuda de un intérprete en todos los casos en los que el individuo era monolingüe. Para

cada individuo se tomaron entre 3 y 5 cabellos arrancados con raíz, que fueron almacenados a temperatura ambiente en tubos tipo eppendorf. Los tubos se rotularon con el código correspondiente al individuo.

Mayas y Pimas. Para la población Maya y Pima se organizó la campaña con la ayuda del Instituto Nacional de Antropología e Historia (INAH) y de la Presidencia Municipal, a la que se entregaron los permisos correspondientes. Se informó previamente del proyecto y de los objetivos del mismo, del tipo de muestras que se iban a tomar y se presentó al personal implicado en la campaña de muestreo. Se solicitó permiso para acudir a la escuela e informar al equipo directivo de la misma. Una vez informadas las familias se trabajó en las escuelas de las comunidades con ayuda de los maestros que actuaron como intérpretes. Para la población Maya se utilizaron dos hisopos de algodón para cada uno de los participantes que fueron almacenados a temperatura ambiente en un sobre de papel. Los sobres fueron rotulados con el código correspondiente al individuo. En las comunidades Pimas se tomaron para cada individuo entre 3 y 5 cabellos arrancados con raíz, que fueron almacenados a temperatura ambiente en tubos tipo eppendorf. Los tubos se rotularon con el código correspondiente al individuo.

Huicholes y Mayos. En estas comunidades se trabajó en coordinación con los centros de salud del IMSS de las cabeceras municipales o de las mismas comunidades cuando los había. La toma de muestras se llevó a cabo solicitando la participación de cada individuo que acudía al centro, e informándole del proyecto con el apoyo del personal del IMSS que en muchos casos actuaron como intérpretes al ser los individuos monolingües. A cada individuo se le extrajeron entre 5 y 10 mL de sangre por punción venosa a tubos con anticoagulante tipo EDTA conservándose en frío, con el correspondiente código. La sangre fue alicuotada en tubos tipo eppendorf y conservada a -80°C. Los tubos se rotularon con el código correspondiente al individuo.

II.4. Toma de muestras en las comunidades indígenas

En base a los muestreos que fueron diseñados especialmente para cada una de las regiones y comunidades a las que se accedió y ya descritos en los apartados anteriores, se especificarán a continuación las zonas en las que finalmente se trabajó

Localización del muestreo	Latitud	Longitud
Estado de Jalisco		
Municipio Mezquitic San Sebastián de Teponahuastlán	22° 04' 51"	-104° 03' 49"
Estado de Quintana Roo		
Municipio Felipe Carrillo Puerto Tihosuco	20° 12' 11"	-88° 22' 14"
Municipio José María de Morelos Sabán	20° 02' 02"	-88° 32' 19"
Estado de Sinaloa		
Municipio El Fuerte Los Capomos	26° 25' 32"	-108° 30' 57"
Estado de Hidalgo		
Municipio Xochiatipán		
Ixtaczoquico	20° 54' 37"	-98° 15' 54"
Texoloc	20° 55' 03"	-98° 15' 14"
Municipio Yahualica		
Atlalco	20° 54' 53"	-98° 19' 41"
Tlalchihualica	20° 58' 13"	-98° 23' 59"
Municipio Atlapexco		
Pahactla	20° 58' 28"	-98° 21' 31"
Tecacahuaco	20° 56' 29"	-98° 20' 21"
Municipio Huejutla de Reyes		
Santa Catarina	21° 05' 56"	-98° 22' 48"
Municipio de Cardonal		
El Buena	20° 38' 02"	-99° 08' 37"
El Deca	20° 36' 44"	-99° 08' 27"
San Andrés Daboxthá	20° 31' 31"	-99° 03' 50"
Municipio Tenango de Doria		
San Nicolás	20° 19' 13"	-98° 11' 16"
Santa Mónica	20° 18' 22"	-98° 13' 18"
Municipio San Bartolo Tutotepec		
Santiago	20° 24' 00"	-98° 13' 05"
Municipio Huehuetla		
Huehuetla	20° 27' 33"	-98° 04' 35"
Estado de Chihuahua		
Municipio de Temósachi		
Yepachi	28° 25' 19"	-108° 22' 39"
Piedras Azules	28° 24' 60"	-108° 18' 59"
Estado de Sonora		
Municipio de Yécora		
Yécora	28° 22' 16"	-108° 55' 32"
Juan Diego de los Pimas	28° 21' 57"	-108° 55' 01"
El Kipur	28° 24' 19"	-108° 35' 45"

Tabla II.1. Coordenadas geográficas de cada una de las comunidades en las que se muestreó (INEGI 2010).

y el número de muestras obtenidas. Se detallarán también las comunidades que accedieron a participar así como las coordenadas geográficas de cada una de ellas (Tabla II.1, Figuras II.1 y II.2).

En todos los casos se reflejan exclusivamente las comunidades indígenas en las que se pudo trabajar y que accedieron a participar en el estudio, no incluyéndose aquellas en las que no se trabajó finalmente, ya fuera por la no autorización del Presidente Municipal, por la negativa de los directores de las escuelas o por otra razón. Igualmente en todas las campañas de muestreo, hubo casos concretos de personas o grupos de personas que no accedieron a participar en el estudio.

II.4.I. Pimas

El muestreo en la región Pima se hizo con ayuda de David Beaumont, fraile franciscano capuchino instalado en Yécora desde el año 1991. Con su ayuda se trabajó en las distintas comunidades Pimas, todas aisladas y en una región abrupta y de difícil acceso, especialmente en el caso de Piedras Azules. En la comunidad Juan Diego de los Pimas el muestreo tuvo que hacerse puerta a puerta y al tratarse de una comunidad muy pequeña se obtuvieron solamente 6 participantes.

Localización geográfica	Municipio	Comunidad	Muestreo
Chihuahua	Temosáchi	Yepachi	46
		Piedras Azules	19
Sonora	Yécora	Yécora	14
		Juan Diego de los Pimas	6
		El Kipur	42
TOTAL			127

Tabla II.2. Localización geográfica del muestreo realizado para el grupo Pima

La distribución del muestreo en las comunidades está reflejada en la tabla II.2. Como puede apreciarse el total fue de 127 muestras; 65 de ellas procedentes de dos comunidades pertenecientes al municipio de Temósáchi, en el estado de Chihuahua, y 62 de tres comunidades de Yécora, estado de Sonora.

II.4.II. Mayos

En el muestreo de la población Mayo se siguió una estrategia similar que con la población Huichol, se hizo en una única comunidad debido a la presencia de un centro de salud del IMSS. Igualmente se trabajó con la ayuda de profesionales de la institución obteniéndose muestras de sangre de 55 participantes voluntarios.

La distribución del muestreo en las comunidades está reflejada en la tabla II.3.

Localización geográfica	Municipio	Comunidad	Muestreo
Sinaloa	El Fuerte	Los Capomos	55
TOTAL			55

Tabla II.3. Localización geográfica del muestreo realizado para el grupo Mayo.

II.4.III. Huicholes

El muestreo de la población Huichol se hizo en una única comunidad debido a la presencia de un centro de salud del IMSS. Se trabajó con la ayuda de profesionales de la institución obteniéndose muestras de sangre de 36 participantes voluntarios como se refleja en la tabla II.4.

Localización geográfica	Municipio	Comunidad	Muestreo
Jalisco	Mezquitic	San Sebastián de Teponahuastlán	36
TOTAL			36

Tabla II.4. Localización geográfica del muestreo realizado para el grupo Huichol.

II.4.IV. Nahuas

El muestreo se hizo en siete comunidades distintas pertenecientes a cuatro municipios del estado de Hidalgo, todos localizados en la llamada Huasteca hidalguense. El número final de individuos participantes fue muy elevado a pesar de que el trabajo se hizo puerta a puerta.

Esta estrategia de disgregación del muestreo permitió obtener siete agrupaciones de individuos pertenecientes a cuatro municipios distintos, teniendo en tres de ellos, individuos procedentes de dos comunidades distintas.

La distribución del muestreo en las comunidades está reflejada en la tabla II.5.

Localización geográfica	Municipio	Comunidad	Muestreo
La Huasteca, Hidalgo	Xochiatipán	Ixtaczoquico	46
		Texoloc	33
	Yahualica	Atlalco	36
		Tlalchihualica	34
	Atlapexco	Pahactla	24
		Tecacahuaco	45
	Huejutla de Reyes	Santa Catarina	34
			TOTAL

Tabla II.5. Localización geográfica del muestreo realizado para el grupo Nahua.

El número de muestras que se obtuvo en cada comunidad fue variable ya que estaba en función de la predisposición de los donantes voluntarios. Así, en Pahactla se obtuvieron 24 muestras mientras que en Ixtaczoquico el número fue de 46 muestras.

II.4.V. Otomíes

Se trabajó en comunidades Otomíes de ambos lados de la Sierra Madre Oriental en el estado de Hidalgo. Por un lado en varias comunidades del municipio de El Cardonal incluido en el valle del Mezquital, y en los municipios de Tenango de Doria y San Bartolo Tutotepec de la llamada Sierra Otomí-Tepehua.

La distribución del muestreo en las comunidades está reflejada en la tabla II.6.

Aunque en ambas localizaciones hablan lengua Otomí tienen variantes dialectales asociadas a cada lugar, y no está claro si se trata de comunidades con orígenes diferentes que han adoptado la misma lengua, o realmente pertenecieron a un mismo grupo.

En este caso el número de muestras recolectadas en cada comunidad fue muy parecido, en todas ellas oscilaron en torno a unas 30 muestras.

Localización geográfica	Municipio	Comunidad	Muestreo
Valle del Mezquital, Hidalgo	El Cardonal	El Buena	31
		El Deca	34
		San Andrés Dabothá	32
Sierra Otomí-tepehua, Hidalgo	Tenango de Doria	San Nicolás	32
	San Bartolo Tutotepec	Santa Mónica	37
		Santiago	34
TOTAL			200

Tabla II.6. Localización geográfica del muestreo realizado para el grupo Otomí.

II.4.VI. Tepehuas

Se trabajó en una comunidad Tepehua de la Sierra Madre Oriental en el estado de Hidalgo, en el municipio de Huehuetla y en la comunidad del mismo nombre de la llamada Sierra Otomí-tepehua. De hecho es la única comunidad del estado de Hidalgo que tiene una representación Tepehua lo suficientemente numerosa como para ser muestreada. El muestreo está reflejado en la tabla II.7.

Localización geográfica	Municipio	Comunidad	Muestreo
Sierra Otomí-Tepehua, Hidalgo	Huehuetla	Huehuetla	66
TOTAL			66

Tabla II.7. Localización geográfica del muestreo realizado para el grupo Tepehua.

II.4.VII. Mayas

En las dos poblaciones Mayas en las que se trabajó, se eligieron estrategias muy especiales debido a la infraestructura de las mismas y por los días que se permaneció en la zona. A la comunidad indígena de Tihosuco se accedió durante un fin de semana, por lo que hubo que descartar la posibilidad de trabajar en la escuela, y se hizo una convocatoria a través del presidente municipal, a la que acudieron 76 personas para participar en el estudio. La distribución del muestreo en las comunidades está reflejada en la tabla II.8.

En Sabán el trabajo pudo organizarse para trabajar en la escuela con la ayuda de los maestros, sobre todo en los cursos inferiores por la corta edad de los estudiantes. En esta comunidad indígena se obtuvieron 104 muestras en un día de trabajo.

Localización geográfica	Municipio	Comunidad	Muestreo
Quintana Roo	Felipe Carrillo Puerto	Tihosuco	76
	José María de Morelos	Sabán	104
TOTAL			180

Tabla II.8. Localización geográfica del muestreo realizado para el grupo Maya.

En la tabla II.9 se ha desglosado, en valores totales y porcentajes, la población de 5 años y más que habla alguna lengua indígena según el XII Censo General de Población y Vivienda (INEGI 2000). El listado se ha hecho para cada estado, municipio y localidad referentes a los puntos de muestreo de los grupos indígenas incluidos en esta tesis. Es importante comentar que los valores de referencia utilizados en esta tesis doctoral, si bien son tomados del XII Censo de Población y Vivienda (INEGI 2000), no son exactos en las regiones remotas y aisladas, por lo que deberían ser tomados como valores puramente orientativos. Como se verá, en algunos casos el tamaño del muestreo es superior al valor poblacional reportado por el INEGI.

La distribución de la población indígena en México no es uniforme, teniendo en cuenta que los estados en los que se ha trabajado con mayor población total son, en orden de mayor a menor: Jalisco (6.300.022), Chihuahua (3.052.907), Sinaloa (2.536.844), Hidalgo (2.235.591), Sonora (2.216.969) y Quintana Roo (874.963). La población indígena se concentra sobre todo en los estados de Hidalgo (339.866) en que es mayoritaria la población hablante de Náhuatl (221.684), seguida de la de Otomí (114.043) y de la de Tepehua (1.842); Quintana Roo con 163.477 hablantes de lengua Maya que es la mayoritaria del estado; en el estado de Jalisco 10.976 hablantes de lengua Huichol siendo la predominante; en Sinaloa la segunda lengua más hablada es el Mayo con 9.077 hablantes. El número total de hablantes de lengua Pima es de 741 en todo el territorio de México, por lo que los números en los estados de Sonora y Chihuahua son muy bajos, con valores de 374 y 346, respectivamente, sin ser en ninguno de los dos casos la lengua predominante.

En cuanto a los municipios de cada región, se eligieron en la medida de lo posible los pequeños, aislados y/o de difícil acceso, con la idea de minimizar el muestreo de individuos no pertenecientes al lugar. Además, en los municipios que presentan mayor

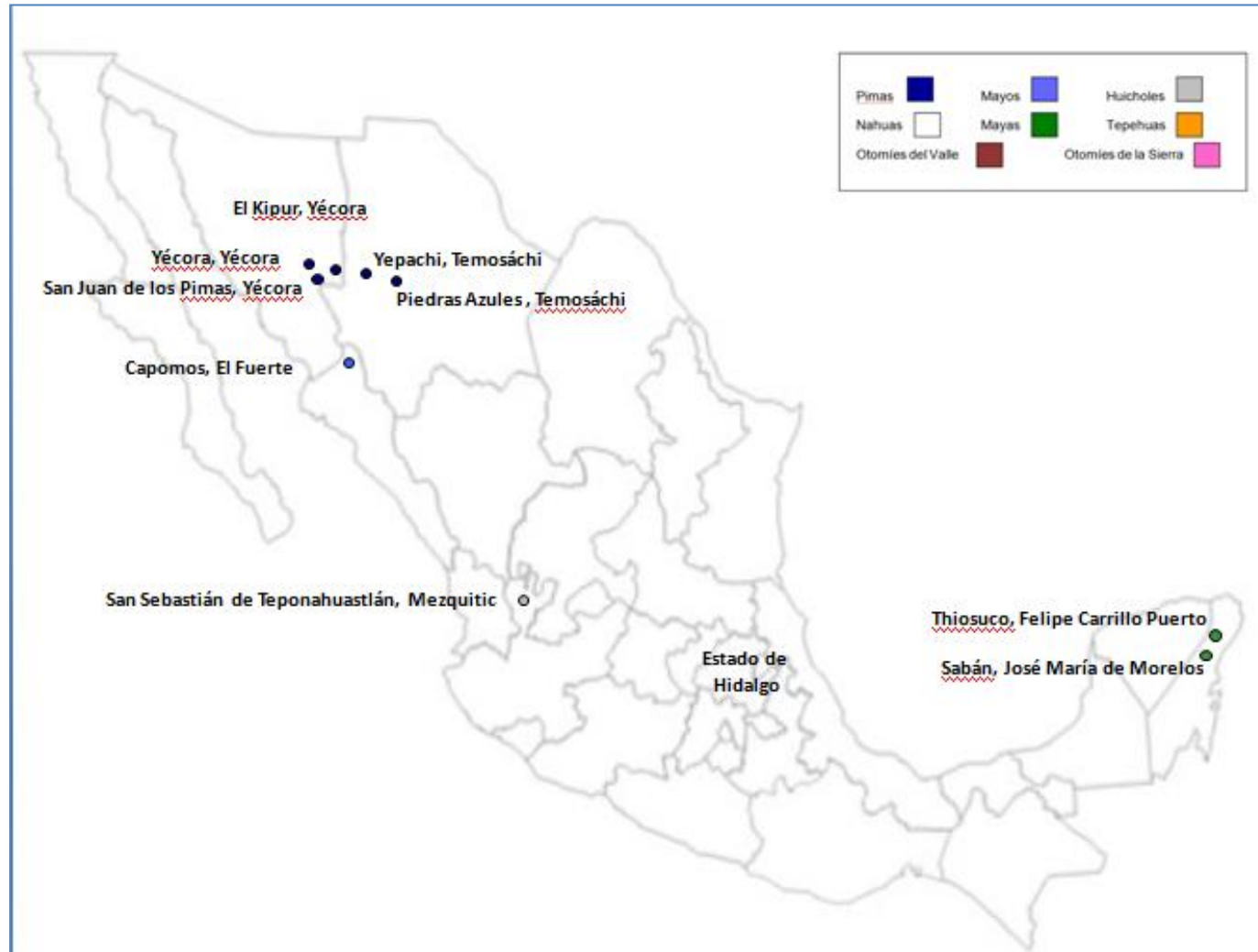


Figura II.1. Localización geográfica de las comunidades indígenas muestreadas en México. En colores se muestran los distintos grupos, el nombre de la comunidad y el municipio al que pertenecen.

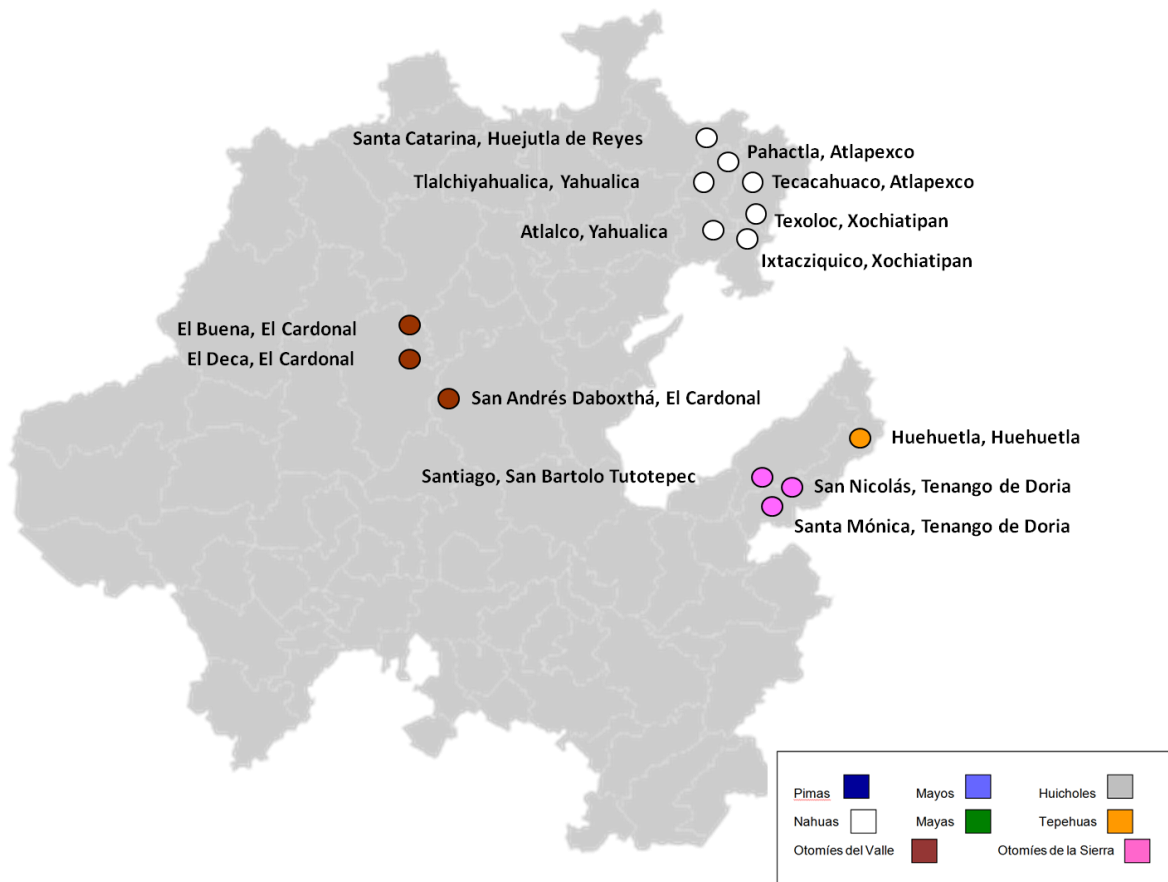


Figura II.2. Localización geográfica de las comunidades indígenas muestreadas en México. En colores se muestran los distintos grupos, el nombre de la comunidad y el municipio al que pertenecen.

población suele haber fenómenos de migración, y como consecuencia puede producirse mestizaje siendo potencialmente un problema que aunque es detectable con los datos de las encuestas, disminuye la proporción de individuos a incluir en el estudio.

Los municipios con menor tamaño de población y mayor porcentaje de hablantes de lengua indígena respecto al total del estado, son los correspondientes a región Pima en Chihuahua y Sonora, Temosáchi y Yécora, con un 43,93% y un 90,11% respectivamente. Estos valores son muy elevados debido al bajo número de hablantes de esta lengua indígena y a su concentración en estas zonas. El siguiente municipio con el porcentaje más elevado de hablantes de lengua indígena es el correspondiente a la región Tepehua, con un valor de 58,22%. Finalmente y de mayor a menor, los municipios de Mezquitic en la región Huichol, El Fuerte de lengua Mayo, Huejutla de Reyes en zona Naha, y Felipe Carrillo Puerto de habla Maya, son los de mayor

concentración de hablantes de lengua indígena con valores de 89,25%, 69,43%, 50,86%, 28,31% y 23,53% respectivamente. Para el resto de municipios los valores oscilan entre el 3,86% de Tenango de Doria en la región Otomí de la Huasteca hidalguense, al 11,06 de la zona Maya de José María de Morelos en Quintana Roo.

Este mismo análisis hecho para las localidades tiene en las comunidades Tepehua, Pimas y las Otomías del municipio de Tenango de Doria, Santa Mónica y San Nicolás, los valores más elevados de porcentaje de hablantes de lengua indígena. Los valores más bajos los detectamos en Santiago (San Bartolo Tutotepec) con un 0,24% de un total de 17 individuos, seguido de Santa Catarina (Huejutla de Reyes), San Sebastián de Teponahuastlán (Mezquitic), Atlalco (Yahualica), El Buena (Cardonal), y Yécora (Yécora) todas ellas por debajo del 5%. El resto de comunidades presentan porcentajes de hablantes de lengua indígena que oscilan entre el 5 y el 11,24%.

II.5. Extracción y cuantificación de ADN

Los métodos de extracción para la obtención de ADN genómico utilizados dependieron de cada tipo de muestra. En todos los casos se trabajó en zonas limpias designadas como áreas de extracción o pre-PCR, utilizando bata, mascarilla y guantes, así como material esterilizado.

Igualmente se trabajó en cabina de flujo laminar y en cabina de gases para evitar posibles contaminaciones en el primer caso, y la inhalación de reactivos orgánicos y volátiles en el segundo. En las ocasiones en que no hubo una cabina de gases disponible se trabajó con mascarilla con filtro de carbono.

II.5.I. Sangre. Se utilizó la extracción orgánica con fenol-cloroformo por su alto rendimiento y sencillez (Sambrook et al., 1989); partiendo de una muestra de 400 μL y tras el lisado de las membranas celulares, pueden obtenerse entre 1.500 y 2.000 $\text{ng}/\mu\text{L}$ de ADN. El protocolo utilizado fue el siguiente:

- Descongelar una alícuota.
- Añadir 1 mL de agua milliQ, mezclar por inversión y centrifugar 5 minutos a 13.200 rpm para sedimentar núcleos. Eliminar el sobrenadante por decantación con

II. Material y Métodos

	Población total/ %	Población LI/ %	LI
México	97.483.412/ 100	6.044.547/ 6,2	
Estado de Jalisco	6.322.022/ 6,49	10.976/ 0,18	Huichol
Municipio Mezquitic	14.614/ 0,23	7.621/ 69,43	Huichol
San Sebastián de Teponahuatlán	256/ 1,75	112/ 1,47	
Estado de Quintana Roo	874.963/ 0,90	163.477/ 2,70	Maya
Municipio Felipe Carrillo Puerto	60.365/ 6,90	38.474/ 23,53	Maya
Tihosuco	4.607/ 7,63	3.714/ 9,65	Maya
Municipio José María de Morelos	31.052/ 3,55	18.088/ 11,06	Maya
Sabán	2.058/ 6,63	1.708/ 9,44	Maya
Estado de Sinaloa	2.536.844/ 2,60	9.077/ 0,15	Mayo
Municipio El Fuerte	89.515/ 3,53	4.617/ 50,86	Mayo
Los Capomos	653/ 0,73	266/ 5,76	Mayo
Estado de Hidalgo	2.235.591/ 2,29	339.866/ 5,62	
		221.684/ 65,23	Náhuatl
		114.043/ 33,56	Otomí
		1.842/ 0,54	Tepehua
Municipio Xochiatipán	16.977/ 0,76	14.072/ 6,35	Náhuatl
Ixtaczoquico	929/ 5,47	838/ 5,96	Náhuatl
Texoloc	943/ 5,55	858/ 6,10	Náhuatl
Municipio Yahualica	20.727/ 0,93	15.380/ 6,94	Náhuatl
Atlalco	628/ 3,03	542/ 3,52	Náhuatl
Tlalchihualica	2.077/ 10,02	1.729/ 11,24	Náhuatl
Municipio Atlapexco	18.029/ 0,80	13.291/ 6,00	Náhuatl
Pahactla	1.121/ 6,22	989/ 7,44	Náhuatl
Tecacahuaco	1.377/ 7,64	1.223/ 9,20	Náhuatl
Municipio Huejutla de Reyes	108.239/ 4,84	62.754/ 28,31	Náhuatl
Santa Catarina	913/ 0,84	732/ 1,17	Náhuatl
Municipio de Cardonal	16.943/ 0,76	8.862/ 7,77	Otomí
El Buena	523/ 3,09	360/ 4,06	Otomí
El Deca	661/ 3,90	505/ 5,70	Otomí
San Andrés Dabothá	689/ 4,07	475/ 5,36	Otomí
Municipio Tenango de Doria	17.175/ 0,78	4.404/ 3,86	Otomí
San Nicolás	1.430/ 8,33	1.258/ 28,56	Otomí
Santa Mónica	1.145/ 6,67	1.1010/ 22,93	Otomí
Municipio San Bartolo Tutotepec	18.650/ 0,83	7.094/ 6,22	Otomí
Santiago	471/ 2,53	17/ 0,24	Otomí
Municipio Huehuetla	25.098/ 1,12	1.649/ 89,52	Tepehua
Huehuetla	2.548/ 10,15	960/ 58,22	Tepehua
Estado de Chihuahua	3.052.907/ 3,13	346/ 0,01	Pima
Municipio de Temosáchi	6.989/ 0,23	152/ 43,93	Pima
Yepachi	649/ 9,29	27/ 17,76	Pima
Piedras Azules	49/ 0,70	16/ 10,52	Pima
Estado de Sonora	2.216.969/ 2,27	374/ 0,01	Pima
Municipio de Yécora	6.069/ 0,27	337/ 90,11	Pima
Yécora	2.763/ 45,53	15/ 4,45	Pima
Juan Diego de los Pimas	108/ 1,78	32/ 9,50	Pima
El Kipur	105/ 1,73	73/ 21,66	Pima

Tabla II.9. Valores totales y porcentajes de población de 5 años y más que habla alguna lengua indígena según el XII Censo General de Población y Vivienda (INEGI 2000). En todos los casos los porcentajes se han calculado respecto a la entidad superior. LI, lengua indígena.

cuidado para no perder el precipitado del fondo. Repetir este paso tantas veces como sea necesario hasta que el color del lavado sea lo suficientemente claro.

- Añadir 500µL de DNA Lysis Buffer, 50µL de SDS (Dodecilsulfato sódico) al 10% y 10 µL de Proteinasa K (20 mg/mL), para producir la lisis y liberar el material celular al medio.
- Incubar los tubos, entre 45 minutos y 2 horas, a 65°C.
- Añadir 500 µL de Fenol: Cloroformo: Alcohol isoamílico (25:24:1), y homogeneizar por inversión las dos fases. Este proceso es una primera purificación del material genético puesto que pasará a la fase acuosa.
- Centrifugar los tubos 5 minutos a 13.200 rpm y pasar la fase acuosa a columnas Microcon-100.
- Tapar las columnas y centrifugar a 4400 rpm durante 15 minutos.
- Vaciar el tubo con los productos de desecho.
- Repetir el lavado con 500µL de agua milliQ.
- Tapar las columnas y centrifugar a 4400 rpm durante 15 minutos.
- Tirar el tubo con los productos de desecho.
- En tubos nuevos, añadir 50µL de agua miliQ sin tocar la membrana e invertir las columnas.
- Centrifugar a 4.400 rpm durante 3 minutos para recoger la muestra purificada.
- El ADN se guarda en la nevera de 4°C cuando el uso va a ser inmediato o almacenado en el arcón de -80°C hasta el momento de su utilización.

Las membranas hidrófilas Ultracel YM de las columnas de filtrado por centrifugación de Microcon se caracterizan por su *Límite de Peso Molecular Nominal* (NMWL) que les confiere la capacidad de retener moléculas de un peso molecular determinado. Las características de baja adsorción de estas membranas, junto con las velocidades de centrifugación utilizadas para la recuperación, proporciona unos rendimientos de recuperación de ácidos nucleicos superiores al 95%, con factores de concentración de 100X.

Las columnas Microcon resultan útiles para una gran variedad de aplicaciones, incluyendo desalinización y limpieza de tampones y otros solventes, para eliminar contaminantes de bajo peso molecular y concentrar más de 20 veces el ADN extraído en cada lavado.

II.5.II. Pelo con raíz mediante Chelex®. Se utilizó la extracción con resina Chelex para obtener la fracción de ADN genómico de pelos con raíz (Higuchi et al., 1988; Walsh et

al., 1991). Partiendo de una muestra de 4 a 5 pelos con raíz, pueden obtenerse entre 150 y 200 ng/ μ L de ADN. Esta resina quelante es un compuesto con alta afinidad por iones metálicos polivalentes que suelen degradar al ADN a altas temperaturas, gracias a esta propiedad la muestra puede ser sometida a 100°C logrando romper la membrana celular y liberando el material genético.

El protocolo utilizado fue el siguiente:

- Lavar el pelo con agua miliQ.
- Cortar el extremo del pelo con la raíz en porciones pequeñas e introducir en un tubo tipo eppendorf.
- Añadir 200 μ L de Chelex al 5%.
- Incubar en termobloque a 56°C durante al menos 8 horas.
- Agitar los tubos en el vortex a durante 5-10 segundos.
- Incubar en termobloque a 100°C durante 8 minutos sellando los tubos con parafilm para que no se abran.
- Agitar los tubos en el vortex a durante 5-10 segundos.
- Centrifugar 3 minutos, 13200 rpm.
- Tomar el volumen de muestra del menisco superior asegurándose de no pipetear Chelex.
- El ADN se guarda en la nevera de 4°C cuando el uso va a ser inmediato o almacenado en el arcón de -80°C hasta el momento de su utilización (para lo que habrá que repetir los tres últimos pasos del protocolo).

II.5.III. Pelo con raíz mediante ChargeSwitch Kit (Invitrogen). Se utilizó un kit basado en partículas metálicas, novedad técnica en el momento de su uso. El kit ChargeSwitch Forensic DNA Purification está asociado a una gradilla magnética o *magnarack* para imantar las partículas metálicas, todos los búferes excepto el que se usa en primer lugar, *buffer mix*, están preparados y listos para su uso. Partiendo de una muestra de 4 a 5 pelos con raíz, pueden obtenerse entre 75 y 100 ng/ μ L de ADN.

El protocolo utilizado, basado en las recomendaciones técnicas del kit, fue el siguiente:

- Lavar el pelo con agua miliQ y trocearlo en porciones pequeñas, introducir en un tubo tipo eppendorf.
- Agregar 1mL de bufer mix de lisis con 1mL de búfer de lisis y 10 μ L de proteinasa K (Promega), la muestra debe quedar sumergida.
- Incubar 1 hora a 55°C o 30 minutos si se agitan las muestras durante la incubación.

Transferir el sobrenadante a un tubo estéril.

- Agregar 200 μ L de buffer de purificación y mezclar bien.
- Agregar 20 μ L de partículas magnéticas y mezclar bien. Usar una pipeta de 100-1000 μ L para evitar la formación de burbujas.
- Incubar a temperatura ambiente entre 1 y 5 minutos.
- Colocar los tubos en el magnarack y esperar a que se forme un *pellet* a la altura del imán de la gradilla. Sin quitar el tubo de la gradilla imantada pipetear y descartar el sobrenadante.
- Proceder de forma inmediata a lavar el ADN. Agregar 500 μ L de buffer de lavado y remover hasta resuspender totalmente las partículas magnéticas. Usar una pipeta 100-1000 μ L para evitar la formación de burbujas.
- Colocar los tubos en el magnarack y esperar a que se forme un *pellet* a la altura del imán de la gradilla. Sin quitar el tubo de la gradilla imantada pipetear y descartar el sobrenadante.
- Repetir los tres últimos pasos para hacer un segundo lavado al ADN.
- Proceder a eluir el ADN, añadiendo 150 μ L de búfer de elución y remover hasta resuspender totalmente las partículas magnéticas.
- Incubar a temperatura ambiente entre 1 y 5 minutos. Remover de nuevo hasta que las partículas estén en suspensión.
- Colocar los tubos en el magnarack y esperar a que se forme un *pellet* en a la altura del imán de la gradilla. Sin quitar el tubo de la gradilla imantada pipetear y transferir la totalidad del volumen a un tubo estéril.
- El ADN se guarda en la nevera de 4°C cuando el uso va a ser inmediato o almacenado en el arcón de -80°C hasta el momento de su utilización.

II.5.IV. Epitelio bucal. Se utilizó la extracción orgánica con fenol-cloroformo por su alto rendimiento y sencillez (Sambrook et al., 1989). Mediante este método se extrajo el ADN genómico lisando las membranas celulares. Partiendo de un hisopo con epitelio bucal pueden obtenerse entre 750 y 1.000 ng/ μ L de ADN. El protocolo utilizado fue el siguiente:

- Cortar el extremo del hisopo de algodón y colocarlo en un tubo tipo eppendorf.
- Añadir 500 μ L de DNA Lysis Buffer, 50 μ L de SDS (Dodecilsulfato sódico) al 10% y 10 μ L de Proteinasa K (20 mg/mL).
- Incubar los tubos, entre 45 minutos y 2 horas, a 65°C.
- Añadir 500 μ L de Fenol: Cloroformo: Alcohol isoamílico (25:24:1), y homogeneizar por inversión las dos fases.
- Centrifugar los tubos 5 minutos a 13.200 rpm y pasar la fase acuosa a columnas

Microcon-100.

- Tapar las columnas y centrifugar a 4.400 rpm durante 15 minutos.
- Vaciar el tubo con los productos de desecho.
- Repetir el lavado con 500 μ L de agua milliQ.
- Tapar las columnas y centrifugar a 4400 rpm durante 15 minutos.
- Tirar el tubo con los productos de desecho.
- En tubos nuevos, añadir 50 μ L de agua miliQ sin tocar la membrana e invertir las columnas.
- Centrifugar a 4.400 rpm durante 3 minutos para recoger la muestra purificada.
- El ADN se guarda en la nevera de 4°C cuando el uso va a ser inmediato o almacenado en el arcón de -80°C hasta el momento de su utilización.

La concentración de ADN se midió por espectrofotometría con un biofotómetro (BioPhotometer de Eppendorf), que efectúa tanto la medición de la concentración de ADN como otros datos importantes para estas mediciones.

Se hizo una dilución 1:50 del ADN que se quería cuantificar, en un volumen final de 150 μ l, utilizando de forma alternada con las muestras agua miliQ, la misma solución utilizada para diluir el ADN. Se midió la absorbancia en modo dsDNA para tener una estimación del grado de pureza del ADN cuantificado.

Los parámetros que se tuvieron en cuenta para valorar la calidad y cantidad de ADN presente en las muestras fueron los siguientes:

1. Cociente A260/A280: el valor debe estar entre 1.8 y 2.0 si la muestra es pura. Cuando existe contaminación con proteínas este cociente está fuera de rango, siendo significativamente menor.
2. Valor A260: entre 0,1 y 0,5 se considera que la medida de ADN es reproducible.
3. Cociente A260/A230: el valor debe ser inferior a 2,2 en muestras de ADN puro. Cuando existe contaminación por fenoles y carbohidratos el cociente es superior a ese valor.

A pesar de que las muestras fueron cuantificadas, finalmente se ajustaron las condiciones de amplificación para optimizar el proceso.

II.6. PCR: Amplificación de la región control o D-Loop

Con la introducción de la técnica de PCR (Polymerase Chain Reaction o Reacción en Cadena de la Polimerasa) por Mullis y Faloona (1987), fue posible el análisis de *loci* hipervariables a partir de vestigios biológicos que no podían ser analizados hasta el momento, provocando una revolución en la genética molecular y en especial, en la genética forense. La técnica permitía la obtención de millones de copias de un fragmento de ADN a partir de una cantidad de ADN molde inicial mínima.

La PCR es una técnica que permite la amplificación *in vitro* de una secuencia específica de ADN hasta generar millones de copias de manera rápida y sencilla (Figura II.3). Esta técnica es extremadamente sensible, ya que permite analizar muestras con cantidades mínimas de ADN (del orden de picogramos) e incluso muestras en las que el ADN se encuentra parcialmente degradado.

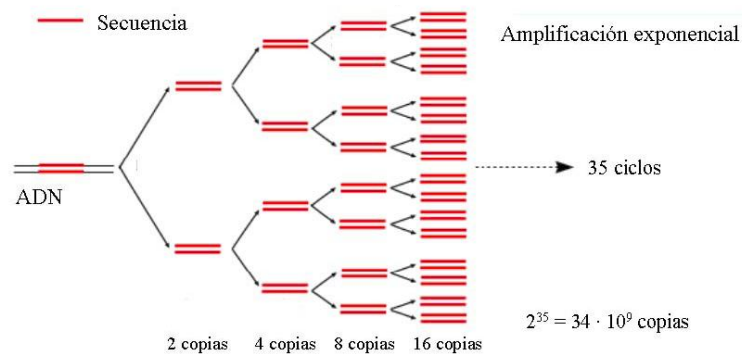


Figura II.3. Esquema explicativo de la Reacción en Cadena de la Polimerasa.

Uno de los aspectos más importantes es la optimización de la PCR, es decir, los parámetros que afectan al rendimiento de la PCR, con el objetivo de poder mejorar su rendimiento, especificidad y reproducibilidad.

Es importante profundizar también en las precauciones necesarias para trabajar en el laboratorio así como para evitar la contaminación, uno de los principales problemas derivados de la enorme sensibilidad de la PCR.

1. Fundamentos de la técnica. La PCR es una técnica diseñada para producir una amplificación selectiva, específica, de una determinada secuencia de ADN, es decir, es capaz de generar *in vitro* grandes cantidades de un determinado fragmento de ADN a partir de cantidades mínimas del mismo.

La técnica se basa en una de las propiedades de las ADN polimerasas, que es su capacidad de sintetizar una cadena complementaria de ADN en dirección 5' - 3', usando como molde un ADN de cadena sencilla, pero utilizando siempre el extremo 3' libre de una región de doble cadena. La PCR utiliza este mismo principio para amplificar, mediante replicaciones sucesivas, la secuencia de ADN deseada utilizando dos oligonucleótidos sintéticos o *primers* (cada uno complementario a las secuencias que flanquean el fragmento de ADN que se desea amplificar) que actúan como cebadores de la síntesis de ADN, catalizada normalmente por una enzima termoestable denominada Taq (*Thermus aquaticus*) polimerasa.

Además, la PCR también utiliza la capacidad del ADN de disociar *in vitro* sus dos cadenas (desnaturalización) al alcanzar temperaturas cercanas a los 100°C, y su capacidad de reasociarse (renaturalización) al disminuir la temperatura, siendo este proceso de hibridación altamente específico, y que en condiciones adecuadas tiene lugar solamente entre hebras de ADN cuyos nucleótidos son complementarios.

La reacción de amplificación se lleva a cabo en una serie de ciclos, cada uno de los cuales incluye normalmente tres fases o pasos: desnaturalización, hibridación o *annealing*, y extensión. Actualmente todo el proceso de amplificación se realiza de forma automática utilizando termocicladores que controlan tanto la temperatura como la duración de cada una de las fases.

- **Desnaturalización.** Para que comience la reacción de amplificación es necesario que la doble cadena de ADN molde, se encuentre en forma de cadena sencilla. Para ello es necesario aumentar la temperatura de la reacción hasta los 90-95°C, de este modo se rompen los puentes de hidrógeno intercatenarios que mantienen unidas las cadenas de ADN complementarias. Esta temperatura debe mantenerse durante algunos segundos para conseguir la completa separación de las hebras, ya que en caso contrario el ADN tenderá a renaturalizarse rápidamente, dificultando la eficiente hibridación de los *primers* en la siguiente fase.
- **Hibridación.** Esta fase también se denomina fase de *annealing* y en ella, una vez que todo el ADN se ha desnaturalizado, se disminuye la temperatura hasta un rango comprendido entre los 40 y 60°C (en función de la T_m de los *primers*, que se calcula según su composición nucleotídica) para permitir la hibridación de los *primers* con las zonas complementarias del ADN molde que flanquean el fragmento que queremos amplificar. Esta fase dura normalmente entre 15 y 60

segundos. Esta es la fase crítica del proceso de amplificación, ya que si la temperatura de *annealing* es demasiado alta no se producirá la unión completa de los *primers*, con lo que no se producirá amplificación alguna, mientras que si es demasiado baja, se producirá la amplificación de productos inespecíficos.

- **Extensión**. Una vez que los *primers* se han unido a las zonas complementarias del ADN molde, se generan fragmentos de ADN bicatenario con un extremo 3' libre que serán utilizados por la Taq polimerasa para incorporar nucleótidos, utilizando como molde la cadena de ADN previamente desnaturalizada. La temperatura a la que se lleva a cabo este paso suele ser 72°C, que es la temperatura a la que la Taq polimerasa alcanza su actividad máxima. Normalmente una extensión de 20-30 segundos es suficiente para fragmentos menores de 500pb.

Estas tres etapas (desnaturalización, annealing y extensión) constituyen un ciclo de amplificación, durante el transcurso del cual se duplica la cantidad de ADN diana producido en el ciclo previo. Así va aumentando de forma exponencial, el número de secuencias de ADN.

2. **Componentes y optimización de la PCR**. El objetivo final de la PCR es la obtención del mayor número de copias posibles de una secuencia específica de ADN. Para lograrlo es imprescindible conocer las condiciones óptimas de cada uno de los componentes de la reacción, así como la propia cinética de la reacción, para poder de este modo solucionar cualquier problema que surja (amplificaciones inespecíficas, falta de amplificado, etc.). A continuación se detallan todos los componentes necesarios para la PCR y su importancia en la optimización de la reacción:

- **ADN molde**. Es el ADN en el que se encuentra la secuencia que queremos amplificar y que la Taq polimerasa utilizará como molde para copiar repetidas veces la llamada secuencia diana. El éxito de la amplificación dependerá de la cantidad y de la calidad del ADN molde.

Con respecto a la calidad del ADN, se obtienen mejores resultados cuando se utiliza un ADN perfectamente purificado, exento de proteínas y de RNA, pero es importante tener en cuenta que la mayoría de los reactivos que se utilizan para purificar ácidos nucleicos (sales, proteasas, solventes orgánicos, detergentes, etc.) son potentes inhibidores de las ADN polimerasas, aunque la precipitación con etanol de las muestras o la purificación en columnas eliminan la mayoría de

los agentes inhibidores. En el caso de que el ADN se encuentre degradado, el éxito de la amplificación dependerá de que el fragmento que queremos estudiar se encuentre intacto o no.

En lo referente a la cantidad de ADN necesaria para la PCR, siempre que se analiza más de 1 ng de ADN los resultados suelen ser plenamente satisfactorios.

- **Primers**. El diseño de los *primers* que se van a utilizar en una PCR es un paso esencial para el éxito de la misma. El paso fundamental es conocer la secuencia exacta de la región que queremos amplificar, ya que de eso dependerá el diseño de los oligonucleótidos. Estos deben ser complementarios a cada una de las hebras del ADN en la región flanqueante de la secuencia de interés.

Es también muy importante el tamaño que se desea amplificar, ya que el rendimiento de la PCR es inversamente proporcional a la longitud del fragmento amplificado, es decir, un fragmento pequeño se amplificará con mayor eficiencia que uno más grande. En ningún caso se recomienda amplificar secuencias mayores de 1.500 pb.

Una vez seleccionada la secuencia diana y la longitud del fragmento que se pretende amplificar es conveniente seguir una serie de recomendaciones para el diseño de los *primers*:

- La longitud de los *primers* debe estar comprendida entre 18 y 25 nucleótidos.
- El porcentaje en GC debe estar entre un 40-60% y el contenido en GC debe ser similar al del ADN diana.
- Es importante evitar repeticiones largas de polipurinas o polipirimidinas (homopolímeros)
- También es importante evitar sucesiones de G o C en el extremo 3', ya que se producirían hibridaciones no específicas en regiones de ADN ricas en GC.
- Para evitar que los *primers* formen estructuras secundarias e hibriden consigo mismo, éstos no deben ser complementarios.
- Los extremos 3' de los *primers* no deben ser complementarios para evitar la formación de dímeros de *primer*.
- Ambos *primers* deben hibridar a temperaturas similares, por lo que se diseñarán de forma que tengan T_m similares.

Antes de trabajar con una pareja de *primers* es necesario comprobar que sean específicos y que únicamente puedan hibridar en la región que deseamos amplificar. Es fundamental que el extremo 3' esté desapareado en cualquier otra zona diferente de nuestra secuencia diana, ya que en ese caso la Taq polimerasa no podrá comenzar la síntesis de ADN.

Para comprobar la especificidad de los *primers* se suele utilizar la herramienta BLAST® (Basic Local Alignment Search Tool) para comparar la secuencia del primer que hemos seleccionado con la base de datos de secuencias de ADN del NCBI (National Center for Biotechnology Information) en la <http://www.ncbi.nlm.nih.gov/>. Para el diseño de *primers* se utilizan programas informáticos que trabajan con secuencias concretas de ADN, teniendo en consideración los factores descritos anteriormente.

También es necesario optimizar la concentración de primer que se va a utilizar en la PCR. Para amplificaciones simples, se suelen utilizar *primers* a una concentración final de entre 0.3 y 0.5 μM en un volumen final de 25 μl .

- **Polimerasa**. La enzima más utilizada es la Taq polimerasa, aislada de una bacteria termófila llamada *Thermus aquaticus*. Esta enzima es termoestable y su procesividad es relativamente alta, alcanzando su máxima actividad entre 65 y 72°C. Como casi todas las polimerasas necesita Mg^{+2} como cofactor para su actividad enzimática.

La cantidad óptima de Taq polimerasa que se recomienda para la síntesis de ADN se encuentra entre 1 y 2 unidades en 25 μl de volumen final de reacción. Con estas cantidades en la mayoría de los casos habrá un exceso de enzima, y la adición de más unidades no aumentará significativamente la cantidad de producto amplificado.

- **Cloruro de Magnesio**. Otro de los factores más críticos y con un efecto más relevante en el rendimiento y especificidad de la reacción de PCR es la concentración de ión magnesio (Mg^{+2}). Interviene en diferentes aspectos del proceso de amplificación, ya que por un lado es un cofactor imprescindible para la actividad enzimática de la Taq polimerasa y por otro, estabiliza la doble cadena de ADN, por lo que la concentración de magnesio afectará a la T_m de las distintas hibridaciones que se producen en la PCR: *primer-ADN* molde, ADN molde-ADN molde y *primer-primer*.

Es importante que la concentración de magnesio sea adecuada, si la concentración está por debajo de 0.5mM, se producirá poca o ninguna amplificación porque se reducirá enormemente la actividad enzimática de la polimerasa. Por otro lado, un exceso de Mg^{+2} aumentará la estabilidad de la doble cadena evitando la desnaturalización completa de la doble cadena en cada ciclo, por lo que se reducirá el rendimiento de la reacción. Por otro lado, una concentración excesiva de Mg^{+2} podría estabilizar también la hibridación incorrecta de los *primers*, reduciendo la especificidad y provocando una mayor cantidad de productos inespecíficos.

En la práctica, la concentración óptima de Mg^{+2} para una determinada reacción, se realiza de forma empírica, utilizando intervalos de 0.5 mM, en el rango desde 1 mM hasta 4 mM.

- **Buffer de amplificación**. Además de $MgCl_2$, los *buffers* de amplificación suelen estar compuestos por otras sales, principalmente KCl o NaCl a 50-100 mM. Como ya se ha comentado, la T_m de los ácidos nucleicos depende de la concentración de sales, por lo que un aumento de su concentración facilitaría la hibridación, estabilizando las dobles cadenas de ADN. Puede provocar amplificaciones inespecíficas a altas concentraciones.

Por otra parte, el buffer de amplificación debe proporcionar el pH óptimo para la Taq polimerasa, manteniéndolo constante a lo largo de la reacción. Generalmente se utilizan concentraciones entre 10 y 50 mM de Tris-HCl a un pH entre 8,3 y 8,8.

- **Desoxinucleótidos trifosfatos (dNTPs)**. La concentración de los diferentes dNTPs debe estar equilibrada aunque el ADN diana posea una secuencia de nucleótidos muy rica en una determinada base. En caso contrario pueden producirse elevadas tasas de error en la replicación. Suelen utilizarse concentraciones de dNTPs en torno a los 200 μM para cada uno de ellos. Es importante señalar que como los iones Mg^{+2} se unen a los dNTPs, una concentración demasiado alta de dNTPs capturaría Mg^{+2} , disminuyendo la actividad de la Taq polimerasa.
- **Adyuvantes de la PCR**. La presencia de inhibidores de la PCR en la solución de ADN, puede reducir el rendimiento de la reacción o incluso impedir la amplificación. Se conocen muchos inhibidores de la PCR como el exceso de

ADN, iones metálicos, etc. En estudios realizados en los últimos años se han purificado y caracterizado los componentes celulares con mayor capacidad de inhibición, como son la inmunoglobulina G en el plasma, la hemoglobina en los eritrocitos y la lactoferrina en los leucocitos (Al Soud *et al.*, 2000; Al Soud and Radstrom, 2001).

El adyuvante más extendido y generalizado, para mejorar el rendimiento y la especificidad de la PCR, es la BSA (Albúmina de Suero Bovino). A concentraciones por encima de 0,8 $\mu\text{g}/\mu\text{l}$, este compuesto incrementa considerablemente la eficiencia de la PCR, ya que actúa como una proteína captadora de iones inhibidores de la Taq polimerasa.

- **Condiciones de la amplificación.** Para optimizar la reacción son esenciales las temperaturas, los tiempos, las rampas de transición de temperaturas para cada una de las tres fases de la PCR por ciclo, así como el número de ciclos de la reacción.

A. Temperatura y duración de la fase de desnaturalización: Es importante que todo el ADN se haya desnaturalizado antes de que empiece a disminuir la temperatura para no entorpecer la hibridación de los *primers*. Suele ser suficiente mantener la reacción a 95°C durante 30 segundos.

La hibridación de los *primers* en sitios distintos de la secuencia diana que puede provocar la amplificación de productos inespecíficos, se producen normalmente durante la preparación de la mezcla de reacción, que suele realizarse a temperatura ambiente. A esta temperatura la Taq polimerasa tiene una actividad parcial. Para evitar este problema se puede realizar la preparación de la mezcla de reacción en recipientes con hielo, evitando la activación de la Taq polimerasa.

B. Temperatura y duración de la fase de hibridación: Para calcular la temperatura de anillamiento de la reacción, debemos conocer la T_m de cada uno de los *primers* escogidos para la amplificación. La T_m se define como la temperatura a la que la mitad del ADN se encuentra en forma de cadena sencilla y la otra mitad en forma de doble cadena, es decir, se podría definir como la temperatura a la que la mitad de cada uno de los *primers* ha hibridado totalmente con su secuencia diana.

La T_m de un oligonucleótido dependerá principalmente de tres factores: de su tamaño (en general, a mayor tamaño, mayor T_m); del porcentaje de GC que posea (a mayor porcentaje de GC, mayor T_m); y de la concentración de sales de la mezcla de reacción (a mayor cantidad de sales, mayor T_m).

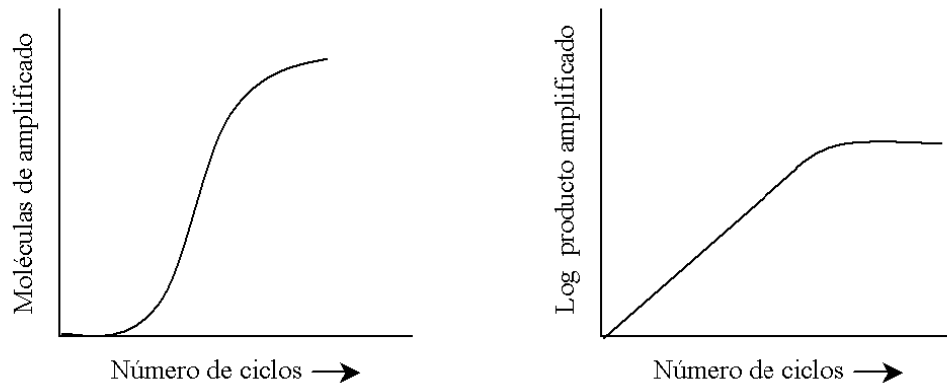
Existen numerosas fórmulas para determinar la T_m teórica de un ácido nucleico, lo que nos puede servir como punto de partida para alcanzar las condiciones de hibridación. De todas formas, la mejor manera de optimizar las condiciones de anillamiento es realizando varias reacciones a diferentes temperaturas, comenzando unos 5°C por debajo de la T_m calculada. A continuación se detalla una de esas fórmulas para calcular la T_m para un oligonucleótido:

$$T_m = 4*(C+G) + 2*(T+A) - 5^\circ\text{C}$$

Una vez que se conoce la T_m de todos los *primers* que van a intervenir en la reacción, se debe seleccionar siempre la más baja para asegurar que ningún primer va a dejar de hibridar con su secuencia diana. En el caso de la aparición de amplificaciones inespecíficas, se debe aumentar la temperatura a intervalos de 2-5°C hasta alcanzar la T_m óptima.

La duración de la fase de hibridación no es un parámetro crítico y suele ser suficiente con 20-40 segundos.

- C. Temperatura y duración de la fase de extensión: Fase menos crítica de la reacción de amplificación, bastan entre 30 y 60 segundos a 72°C para amplificar fragmentos menores a 1.000 pb. Esto es debido a la elevada procesividad de la Taq polimerasa (35-70 bases por segundo).
- D. Número de ciclos: El número de ciclos de una reacción de amplificación suele estar entre 25 y 40. Un incremento en los mismos no aumenta la cantidad de producto amplificado, ya que la propia cinética de la reacción de la PCR lo impide. Durante los primeros ciclos no se observa un aumento considerable en la cantidad de producto aunque posteriormente, la amplificación sigue un proceso exponencial, para alcanzar por último un punto de saturación (entre 10^8 y 10^9 moléculas). Este fenómeno se conoce como *efecto Plateau* (Figura II.4):

Figura II.4. Efecto *plateau*.

Los factores que pueden contribuir al *efecto plateau* pueden ser la inactivación progresiva de la polimerasa; que la concentración relativa de la polimerasa comienza a ser limitante; la caída progresiva de la eficacia de desnaturalización del ADN molde; la caída progresiva de la eficacia de hibridación de los *primers* (el ADN molde se rehibrida), etc.

E. Rampas de transición de temperaturas: Un factor importante son las rampas de transición de temperaturas, podrían definirse como la velocidad a la que se pasa de una temperatura a otra durante la reacción. Muchas veces es necesario que esta transición sea lenta, especialmente de desnaturalización (95° C) a anillamiento, de manera que la hibridación de los *primers* sea lo más eficiente posible.

3. Precauciones para evitar la contaminación. La gran sensibilidad de la reacción de PCR puede llegar a ser un problema si las muestras o los reactivos utilizados están contaminados. La mayor parte de los procesos susceptibles de sufrir contaminaciones ocurren en el laboratorio, siendo la mayor fuente de contaminación el ADN amplificado en reacciones previas. Hay que tener en cuenta que un tubo en el que se ha realizado la PCR contiene cantidades enormes de ADN por lo que abrir el tubo puede provocar que parte del vapor de agua que contiene (y en el que el ADN amplificado está disuelto), salga al exterior, quedando en la atmósfera del laboratorio. Para evitar este problema es recomendable centrifugar los tubos antes de abrirlos y destaparlos con precaución.

Una medida fundamental para minimizar el problema de la contaminación consiste en separar físicamente las zonas de trabajo pre-PCR y post-PCR. Es conveniente también utilizar material de laboratorio exclusivo para cada área (pipetas, reactivos,

equipos, etc.) y preparar las reacciones de amplificación en una cámara de flujo laminar exclusiva para este proceso. También se recomienda utilizar micropipetas de desplazamiento positivo y puntas con filtro desechables, así como trabajar con material esterilizado en ambas áreas, poniendo especial cuidado en el área de pre-PCR.

Para evitar la contaminación provocada por personal del laboratorio es imprescindible la utilización de batas y guantes, siendo conveniente cambiar de guantes frecuentemente. Igualmente es necesario que todo el personal del equipo esté genotipado para todos los marcadores genéticos que se utilizan de rutina.

Para evitar que los reactivos se contaminen conviene alicuotarlos en volúmenes pequeños y añadir siempre como último elemento a la mezcla de reacción, el ADN. Es importante hacer una buena planificación para manipular y pipetear los reactivos lo menos posible.

Por último, es absolutamente imprescindible para asegurar que no ha existido ningún tipo de contaminación, la utilización de controles para garantizar que el ADN no se ha contaminado durante el proceso de extracción o que los reactivos están contaminados. Como control negativo se agregará agua a un tubo de PCR con todos los reactivos excepto el ADN. Como control positivo se incluirá una muestra de genotipo conocido y cuya amplificación sea efectiva.

La amplificación de la región control o D-Loop se realizó en dos fragmentos para todos los individuos utilizando las parejas de *primers* 15997L- 017H, y 16555L- 611H, a excepción de la población Huichol. En estos individuos se trabajó con tres fragmentos utilizando las siguientes parejas de *primers* 15997L- 16401H, 16380L- 285H y 172L- 599H que igualmente cubrirían la totalidad de la región control como se observa en la figura II.5.

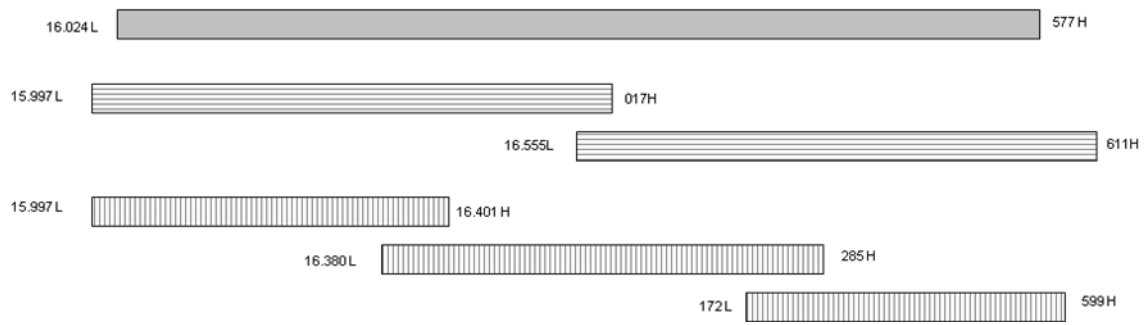


Figura II.5. Localización y solapamiento de los *primers* utilizados.

Primer	Secuencia	Origen
15997L	5' CAC CAT TAG CAC CCA AAG CT 3'	Ward et al., 1991
16380L	5' TCA GAT AGG GGT CCC TTG AC 3'	Álvarez-Iglesias et al., 2007
16401H	5' TGA TTT CAC GGA GGA TGG TG 3'	Vigilant et al., 1989
16555L	5' CCC ACA CGT TCC CCT TAA AT 3'	Álvarez-Iglesias et al., 2007
017H	5' CCC GTG AGT GGT TAA TAG GGT 3'	Álvarez-Iglesias et al., 2007
172L	5' ATT ATT TAT CGC ACC TAC GT 3'	Álvarez-Iglesias et al., 2007
285H	5' GGG GTT TGG TGG AAA TTT TTT G 3'	Álvarez-Iglesias et al., 2007
599H	5' TTG AGG AGG TAA GCT ACA TA 3'	Brandstätter et al., 2004
611H	5' AGT GTA TTG CTT TGA GGA GG 3'	Álvarez-Iglesias et al., 2007

Tabla II.10. Secuencias de los *primers* utilizados y origen de los mismos.

De las muestras de Mayos y Huicholes se cargó 1 μ L de ADN (A); de Mayas, Nahuas, Otomíes y Tepehuas, se cargaron 3 μ L de ADN (B); de las muestras de Pimas se cargaron 9 μ L de ADN (C). Las cantidades de cada reactivo se especifican en la tabla II.11.

Se trabajó en termocicladores Mastercycler de Eppendorf y GeneAmp 6700 de Applied Biosystems, con el programa de amplificación: 95°C 1 minuto; 36 ciclos de 95°C 30 segundos, 55°C 59 segundos, 72°C 30 segundos; una extensión final a 72°C durante 10 minutos y 12°C hasta la finalización del programa.

	A	B	C
H ₂ O miliQ	13,50	11,50	5,00
Buffer (10X)	2,50	2,50	2,50
BSA (1µg/µL)	3,00	3,00	3,00
MgCl ₂ (25mM)	1,50	1,50	1,50
dNTPs (2mM)	2,50	2,50	2,50
Primers (2,5mM)	0,50	0,50	1,00
Taq Polimerasa (5U/µL)	0,50	0,50	0,50
MIX	24,00	22,00	16,00
DNA	1,00	3,00	9,00
Total	25,00	25,00	25,00

Tabla II.11. Condiciones de amplificación, todas las cantidades están expresadas en µL.

II.7. Electroforesis en gel de agarosa

Los sistemas electroforéticos permiten que los productos de amplificación sean separados de acuerdo con su tamaño en geles de electroforesis habitualmente de agarosa (o de poliacrilamida), que son posteriormente visualizados mediante tinción con bromuro de etidio (o nitrato de plata). La electroforesis en geles de agarosa se utiliza para separar moléculas de ADN con un tamaño comprendido entre los 200 pb y las 50 Kb, sometiéndolos a electroforesis horizontales en un campo eléctrico de dirección constante. Aunque el poder de resolución es menor que el de los geles de poliacrilamida, existe agarosa de alta resolución capaz de separar fragmentos con una diferencia de sólo 4 pb.

La movilidad del ADN en geles de agarosa se ve afectada por una serie de factores como son:

- **Concentración de agarosa y tamaño del ADN:** Las moléculas lineales de ADN de doble hebra migran a través del gel a una velocidad que es inversamente proporcional al logaritmo de su masa molecular. Las moléculas de mayor tamaño migran más despacio debido a la resistencia que encuentran a su paso por los poros del gel. Para separar fragmentos grandes, entre 0.5 y 30 Kb, se utilizan

concentraciones de agarosa bajas, entre el 0.5 y 1%; mientras que con concentraciones más elevadas se resuelven los fragmentos más pequeños, debido a que se reduce el tamaño del poro al aumentar el porcentaje de agarosa. Por lo tanto, utilizando geles de diferentes concentraciones puede resolverse un amplio intervalo de tamaños de moléculas de ADN.

- Voltaje: En general, se puede afirmar que los fragmentos de ADN migran a una velocidad proporcional al voltaje aplicado. De todos modos, hay que tener en cuenta que la capacidad de separación en geles de agarosa disminuye conforme se aumenta el voltaje, por lo que para una máxima resolución de fragmentos de ADN no se recomienda utilizar un voltaje muy elevado. Es habitual utilizar un voltaje un poco inferior a 100V.
- Composición de la solución amortiguadora (tampón) de electroforesis: La movilidad electroforética del ADN varía según la composición y la fuerza iónica del tampón empleado. En ausencia de iones, la conductancia eléctrica es mínima y el ADN migra muy despacio. Si por el contrario la fuerza iónica es muy alta, aumenta mucho la resistencia al movimiento de cargas, y el ADN también migra muy despacio.

La tinción de geles de agarosa se suele realizar con bromuro de etidio, añadiéndolo al gel antes de la electroforesis, cuando está en fase líquida y no ha polimerizado todavía, o bien añadiéndolo al tampón de electroforesis, generalmente TAE o TBE. El bromuro de etidio es una molécula que se intercala entre las bases del ADN (por lo que es potencialmente cancerígeno) y emite una fluorescencia anaranjada al ser iluminado con luz ultravioleta. Esta fluorescencia puede ser plasmada en una película fotográfica.

Actualmente la utilización de los geles de agarosa está prácticamente limitada a la cuantificación relativa de la cantidad de ADN presente en una muestra extraída, la comprobación de la integridad del ADN, y como comprobación del correcto funcionamiento de la reacción de PCR, previa secuenciación del ADN mitocondrial, como es este caso.

El protocolo para realizar un gel de agarosa se detalla a continuación:

1. En un matraz de 250mL de capacidad añadir 50mL de TAE (0,04M Tris-Acetato, 0,001M EDTA) y 1g de agarosa.
2. Calentar utilizando un microondas hasta que se funda la agarosa.

3. Dejar enfriar hasta aproximadamente 50°C y añadir 2,5µL de bromuro de etidio.
4. Mientras se enfría la agarosa hasta 50°C sellar con cinta adhesiva el molde en el que se preparará el gel. Colocar el molde en una superficie horizontal y situar el peine para crear los pocillos a unos centímetros del borde.
5. Añadir la solución al molde con cuidado de retirar las burbujas que se formen. Dejar gelificar la agarosa hasta que adquiera una apariencia traslúcida.
6. Retirar el peine y la cinta adhesiva con cuidado. Colocar el gel en la cubeta de electroforesis y cubrirlo con tampón de recorrido (TAE).
7. Con cuidado de no romper los pocillos, cargar en el gel las diferentes muestras de ADN amplificado. Para ello se mezclan 5µL de tampón de carga con 5µL de ADN amplificado en tubos de 1.5mL. El tampón de carga contiene un compuesto para aumentar la densidad (por ejemplo glicerol), de manera que las muestras se depositan en los pocillos. Además se añaden uno o dos colorantes que migran en el gel marcando el frente de migración de las muestras durante la electroforesis. Una vez mezclado ADN y tampón de carga ayudados de una micropipeta cargar dicha mezcla en un pocillo del gel.
8. Cargar del mismo modo que el amplificado un pocillo con un marcador de peso molecular o *ladder* que permita localizar las bandas de aproximadamente 400pb.
9. Conectar la fuente de alimentación y correr el gel durante 30 minutos aproximadamente a 100V.
10. Analizar los resultados mediante observación del gel en un transiluminador de luz ultravioleta.

II.8. Purificación de amplificadores y reacción de secuenciación

La purificación de los productos de PCR se hizo en columnas Microcon-100, cuyas características se han explicado en el apartado I.6.I. Este paso es necesario para eliminar *primers* no consumidos durante el proceso de amplificación, dNTPs, y cualquier otro elemento que pueda interferir en la reacción de secuenciación.

El proceso seguido fue el siguiente:

- Añadir 200µL de agua milliQ a las columnas.
- Tapar las columnas y centrifugar a 4.400 rpm durante 15 minutos.
- Vaciar el tubo con los productos de desecho.

- Tirar el tubo con los productos de desecho.
- En tubos nuevos, añadir 30 ó 20 μ L sin tocar la membrana e invertir las columnas de agua miliQ. La diferencia en el volumen añadido para la resuspensión de los productos de PCR se hizo en función de la intensidad de las bandas observadas en el gel de agarosa; si eran intensas se agregaron 30 μ L, en bandas más tenues se agregaron 20 μ L.
- Centrifugar a 4.400 rpm durante 3 minutos para recoger la muestra purificada.
- El producto purificado se guarda en la nevera de 4°C cuando el uso va a ser inmediato o almacenado en el arcón de -80°C hasta el momento de su utilización.

Esta reacción permite conocer la secuencia nucleotídica de cualquier producto amplificado por PCR, utilizando una química de secuenciación cíclica con terminadores marcados con fluorocromos. Este tipo de secuenciación combina la técnica de la PCR con el método clásico de secuenciación de Sanger. Es un método muy simple en el que como consecuencia de sucesivos ciclos, se consigue una amplificación lineal de los productos de extensión (Sanger et al., 1977).

La mezcla de reacción de secuenciación cíclica incluye una polimerasa, dNTPs, ddNTPs (en menor proporción que los dNTPs), buffer, el ADN molde (normalmente un producto de PCR purificado) y un único primer complementario a una de las cadenas del ADN molde. La química de secuenciación aprovecha la capacidad de los secuenciadores automáticos para detectar fluorescencia a cuatro longitudes de onda simultáneamente, para asignar a cada uno de los nucleótidos del ADN una de esas longitudes o colores uniéndole directamente el fluorocromo al didexosinucleótido. Es la llamada química *dye-labeled terminators*. Cada uno los ddNTPs están marcados con un fluorocromo diferente, de modo que las cadenas crecientes terminan simultáneamente y quedan marcadas con el fluorocromo de la base que provocó el final de la síntesis (figura II.6).

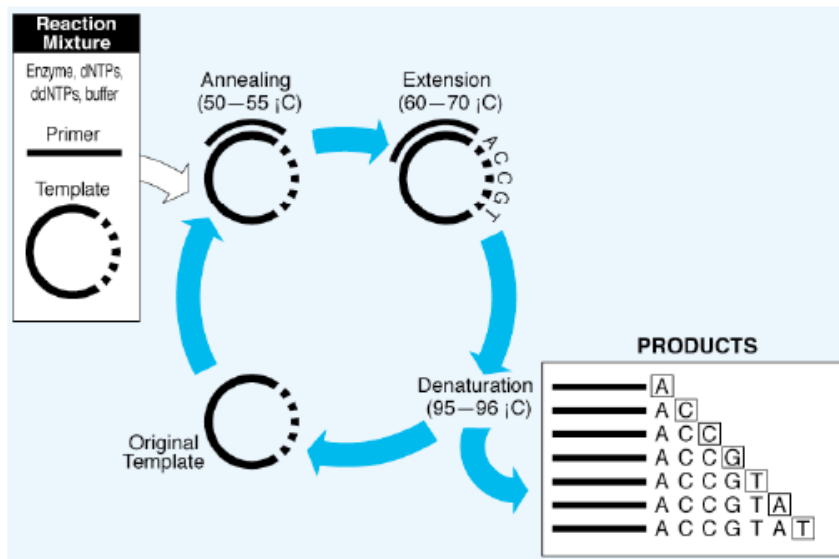


Figura II.6. Reacción de secuenciación por *dye-labeled terminators*.

Funcionamiento del secuenciador automático:

Esta tecnología basada en el empleo de fluorocromos permite la detección con sistemas automatizados de electroforesis capilar, que incorporan un láser y un lector para la detección de la fluorescencia. El análisis de polimorfismos de ADN utilizando esta tecnología automatizada ha supuesto una mayor facilidad y rapidez en el análisis, gran sensibilidad y precisión en la lectura de las muestras. La química de secuenciación está diseñada en base a la capacidad de los secuenciadores automáticos para detectar fluorescencia a cuatro longitudes de onda simultáneamente, de modo que se asigna a cada uno de los nucleótidos del ADN una de esas longitudes.

Para la reacción de secuenciación se utilizó el kit BigDye Terminator v1.1 de Applied Biosystems basado en el método de secuenciación de Sanger (Sanger et al., 1977), que utiliza 2'-3'-dideoxynucleótido trifosfato (ddNTPs), estos difieren de los deoxynucleótidos precursores de la síntesis de ADN en que el grupo 3'-hidroxilo está sustituido por un hidrógeno. Los ddNTPs son reconocidos por la ADN polimerasa como los dNTPs y son añadidos al extremo de la cadena en síntesis. Al carecer del grupo 3'-hidroxilo no pueden unirse al nucleósido 5'-trifosfato de modo que su

incorporación supone la interrupción de la síntesis de la cadena de ADN. Así la inclusión de ddNTPs hace que de ADN termine al azar en cualquier sitio.

Las reacciones de secuenciación se hicieron con las parejas de *primers* de cada uno de los fragmentos, 15997L- 017H, y 16555L- 611H, y en ambos sentidos para cada uno de ellos. Las cantidades de cada uno de los reactivos se especifican en la tabla II.12.

BigDye Reaction Premix	2,00
BigDye sequencing buffer 5X	1,00
Primer (3,2pmol)	0,40
MIX	3,40
DNA	6,60
Total	10,00

Tabla II.12. Condiciones de la reacción de secuenciación (las cantidades están expresadas en μL).

Se trabajó en termocicladores Mastercycler de Eppendorf y GeneAmp 6700 de Applied Biosystems. El programa fue el siguiente: 96°C 1 minuto y 25 ciclos de 96°C 10 segundos, 50°C 5 segundos, 60°C 4 minutos; 12°C hasta la finalización del programa.

II.9. Purificación de secuencias y análisis

Las secuencias deben ser purificadas antes de ser analizadas, este paso es necesario para eliminar *primers* no consumidos durante el proceso de secuenciación, ddNTPs, y cualquier otro elemento que pueda interferir en su análisis.

Para ello se llevó a cabo un método de purificación con etanol, siguiendo el protocolo siguiente:

- Preparar tantos tubos tipo eppendorf como muestras.
- Añadir 80 μL de etanol al 76% a 4°C y mezclar con cada muestra.
- Dejar reposar 15 minutos a temperatura ambiente.
- Centrifugar a 13.200 rpm durante 20 minutos.
- Vaciar los tubos por decantación.
- Añadir 250 μL de etanol al 70% a 4°C y mezclar con cada muestra.

- Centrifugar a 13.200 rpm durante 10 minutos.
- Vaciar los tubos por decantación.
- Secar en un evaporador centrífugo tipo Speed-Vac durante 10 minutos.
- Añadir 25 μ L de formamida desnaturalizante y transferir a una placa de secuenciador.
- El producto purificado se puede guardar en la nevera de 4°C si el uso va a ser inmediato o almacenado en el arcón de -80°C hasta el momento de su utilización.

Las muestras se analizaron en secuenciadores de Applied Biosystems, ABI PRISM 3100 Genetic Analyzer y ABI PRISM 3130 Genetic Analyzer.

II.10. Alineamiento de secuencias y determinación del haplogrupo

El análisis del ADNmt se hace en referencia a una secuencia conocida como secuencia de Anderson o Secuencia de Referencia de Cambridge (CRS) (Anderson et al., 1981; Andrews et al., 1999). Cualquier muestra analizada para ADNmt se reporta en forma de diferencias con respecto a la cadena L de CRS. Por ejemplo la posición 280 de la secuencia de Anderson se corresponde con una C, si la secuencia de la muestra analizada posee una T en esa misma posición, la nomenclatura adecuada para reportar esa mutación sería 280T. Aquellas posiciones en las que no existan diferencias entre la secuencia Anderson y la secuencia problema no serán reportadas.

- Inserciones y deleciones: Las deleciones de nucleótidos con respecto a CRS se reportan como el número de la posición delecionada seguido de una "d". Por ejemplo, si la base entre las posiciones 245 y 247 ha sido delecionada, se nombra como 246d. Si se observa la inserción de un nucleótido, ésta se nombraría como la posición adyacente (*upstream*) al lugar de la inserción, seguido de ".1" y la base insertada. Por ejemplo, si se observa una adenina adicional entre las posiciones 245 y 246, se designará como 245.1A.

Si estas inserciones ocurren dentro de un tracto homopolimérico, no se puede conocer la posición exacta de la inserción. El estándar adoptado para estos casos es asumir que la inserción ha ocurrido en la posición final de la región. Por ejemplo, si ha ocurrido 1 inserción en las poliC entre las posiciones 302 y 310, se debe designar como 309.1C; si ha habido dos inserciones se debe nombrar como 309.1C y 309.2C, y así sucesivamente.

- Heteroplasmas puntuales: La heteroplasma se define como el fenómeno por el que un individuo presenta más de un genotipo de ADNmt. De este modo, cuando una mutación surge en una de las moléculas se crea una mezcla intracelular de moléculas mutantes y normales. El primer ejemplo documentado de puntos de mutación heteroplásmicos dentro de la región control fue el reportado en la identificación de los restos esqueléticos del Zar Nicolás II (Gill et al., 1994; Coble et al., 2009; Rogaev et al., 2009).

Si se observa de forma clara una heteroplasma en la que la proporción de los nucleótidos para la misma posición sea similar (~50%), se debe utilizar la designación IUB (International Code for Unresolved Bases) correspondiente (Tabla II.13.). Por ejemplo, si en la posición 152 existe una mezcla de C y T, se designará como 152Y. De modo alternativo, también se podría utilizar la designación T ~ C.

De todos modos muchas veces es difícil diferenciar las heteroplasmas del ruido de fondo, por lo que se recomienda comprobar la existencia de la heteroplasma secuenciando ambas cadenas. En los casos en los que exista un nucleótido predominante se podría usar, por ejemplo, la fórmula A>G para indicar que A está en una mayor proporción que G. Si la presencia de dos bases en una misma posición no puede ser confirmada en una segunda secuenciación, la posición debe designarse como ambigua, utilizando la letra N en la posición conflictiva.

Una vez cargadas las muestras en el secuenciador automático y obtenido los electroferogramas correspondientes, se compara la secuencia obtenida con la secuencia de Anderson de referencia, utilizando un software de alineamiento de secuencias que existen en el mercado. Uno de los problemas más frecuentes en la secuenciación de la región control son las zonas de poliC existentes tanto en HVI (entre las posiciones 16.184 y 16.193, numeración acorde a CRS) como en HVII (entre las posiciones 310 y 315, numeración acorde a CRS). En las ocasiones en que se produce una mutación en las timinas presentes en estas zonas de la secuencia, es complejo determinar si el análisis es correcto por lo que la revisión de los datos en ambos sentidos es necesaria.

El análisis y revisión de las secuencias resultantes se hizo con tres programas en función de su calidad. Se trabajó con Sequencing Analysis versión 5.2, y Chromas Lite versión 2.01 y BioEdit versión 7.0.5.3 para la revisión y alineamiento de las secuencias mediante comparación directa con la secuencia de referencia o secuencia de Cambridge revisada (rCRS) (Anderson et al., 1981; Andrews et al., 1999) con

SeqScape versión 2.5. Algunas de ellas se alinearon a mano y todas fueron revisadas al menos por dos personas y al menos en tres ocasiones. Para evitar errores se secuenció la muestra procedente de la autora de este trabajo como precaución y con el fin de detectar posibles contaminaciones.

Designación IUB	Bases ambiguas	Designación IUB	Bases ambiguas
M	AC	K	GT
R	AG	V	ACG
W	AT	H	ACT
S	CG	D	AGT
Y	CT	B	CGT

Tabla II.13. Código IUB para designar posiciones con más de una base.

La asignación de haplogrupos individuales se hizo describiendo los lugares polimórficos de cada secuencia y asignándole un haplogrupo concreto. Para ello se utilizaron diferentes fuentes de información. Por una parte se utilizaron las publicaciones de Achilli, (Achilli et al., 2008), Behar (Behar et al., 2007), Tamm (Tamm et al., 2007), O'Rourke (O'Rourke and Raff, 2009) y Perego (Perego et al., 2009, 2010). Por otra parte se explotaron diferentes recursos *online*, como fueron las siguientes páginas

- MITOSEARCH: www.mitosearch.org. Es una página de servicio público que permite entre otras cosas, cargar uno a uno los puntos polimorfismos de una muestra y hacer una búsqueda en un base de datos individual. Los resultados indican los individuos a nivel mundial que tienen este punto variable, también da información sobre el haplogrupo asignado así como el origen de las muestras.
- MITOWEB: www.mitomap.org/Main/WebHome. MitoWeb es una plataforma para distribuir, analizar y desarrollar herramientas bioinformáticas basadas en la Web para el análisis de ADNmt. Presenta múltiples funciones aunque en este caso se ha utilizado sobre todo la opción *Mitosearch* que permite introducir los puntos polimórficos y efectuar una búsqueda. Los resultados indican si ese polimorfismo ya ha sido descrito y la referencia bibliográfica de su publicación.
- GENOGRAPHIC PROJECT: nfhgtool.nationalgeographic.com/classify/. Es una herramienta diseñada y desarrollada por el consorcio Genographic que contiene la información producida en los primeros 18 meses de trabajo de este proyecto. Ha resultado ser una útil herramienta de clasificación de haplogrupos. Los datos se

pueden cargar como sitios polimórficos individuales o como una base de datos generada en un editor de texto. El resultado es la asignación, en función de la base de datos, de un haplogrupo.

- PHYLOTREE: www.phylotree.org/. Es una página que ofrece un árbol filogenético basado en la variación global del ADN mitocondrial humano. Para su construcción se han tenido en cuenta ADNmt completos y permite la ubicación de un polimorfismo en este árbol, asignando la nomenclatura correcta y el haplogrupo. Es una herramienta muy versátil y funcional y tiene la ventaja de que se actualiza cada seis meses, siempre y cuando haya nuevos datos (Van Oven and Kayser, 2009)

II.11. Software y determinaciones estadísticas

A partir de las secuencias completas y alineadas se creó una base de datos en Excel que sirvió de referencia para modificar los formatos, y compatibilizarlos con *software* específicos. Esta base de datos contenía más de 620 filas correspondientes a las secuencias estudiadas y 1.115 columnas correspondientes a cada uno de los nucleótidos que conforman la región D-Loop, así que la base original completa estaba formada por casi 700.000 celdas. Posteriormente se dividió en bases de datos complementarias por poblaciones, por regiones HVI y HVII, así como otras clasificaciones que generaron bases parciales necesarias para la organización y manipulación específica de la información.

Estas bases de datos se cargaron en diferentes programas para su tratamiento. Uno de estos programas es DnaSP v5.10 (Librado and Rozas, 2009). A partir de este programa se calcularon diferentes estimadores estadísticos e índices de diversidad como el número de haplotipos diferentes (k), número de sitios polimórficos (S), diversidad nucleotídica (π) (Tajima 1983, 1993), diversidad haplotípica (H) (Nei 1987), número medio de nucleótidos diferentes (M). Para π y H se calcularon los errores estándares. Aplicando este mismo programa se calculó el índice de D de Tajima (Tajima 1989, 1996) así como su probabilidad asociada, los cambios en el tamaño de la población expresados mediante *pairwise difference*. En algunos casos se utilizó para comprobar los resultados obtenidos en análisis realizados en otros programas y para la representación gráfica de las distribuciones *mismatch* (Rogers and Harpending, 1992; Ray et al., 2003; Excoffier 2004). Por último también se calcularon los haplotipos únicos en cada población y los compartidos entre poblaciones.

Los datos de DnaSP fueron importados a diferentes programas para realizar análisis específicos. Se utilizó el programa Arlequin v. 3.5.1.2 (Excoffier and Lischer, 2010) para el cálculo del test exacto de diferenciación poblacional (Raymond and Rousset, 1995; Goudet et al. 1996) utilizando 10.000 repeticiones en la cadena de Markov para cada comparación, y las distancias genéticas basadas en el estimador F_{st} de Wright (Wright 1965) y sus probabilidades asociadas. En todos los casos el límite de significación fue de 0,05.

En este mismo programa se calcularon las AMOVAS (Análisis Molecular de la Varianza) basadas en el estimador F_{st} que permiten estimar la diferenciación genética dentro de las poblaciones, entre poblaciones de un mismo grupo y entre grupos, así como determinar la hipótesis de diferenciación. Este análisis se realizó sobre todo en agrupaciones jerarquizadas. La asignación de categorías permite construir agrupaciones en las poblaciones y determinar el grado de estructuración de la variación genética observada. Así, puede dividirse la variación en un componente intragrupal (F_{sc}) y en un componente intergrupalo (F_{ct}) (Nei, 1977; Weir and Cockerham, 1984) y en F_{st} que estima las diferencias entre poblaciones (Excoffier et al., 1992; Weir 1996). También se calcularon los haplotipos únicos en cada población y los compartidos entre poblaciones y se contrastaron con los mismos resultados obtenidos con el software DnaSP.

A partir de las distancias genéticas F_{st} obtenidas con el programa Arlequin, se creó una matriz de distancias genéticas que se representaron en un Multidimensional Scaling (MDS) aplicando la aplicación estadística para Excel XL-Stat. El MDS no es un método estadístico exacto, sino más bien una manera de reorganizar las poblaciones en el espacio para obtener la configuración que mejor se aproxima a la matriz de distancias genéticas. Mediante un proceso de permutaciones, el programa cambia la configuración espacial de las poblaciones definiendo el espacio dimensional más adecuado, y reproduciendo la configuración que mejor se aproxima a la matriz de distancias genéticas. Este método viene acompañado de un índice de *stress* que indica cuánto se ha distorsionado la representación respecto a los datos originales.

Los *networks* se han realizado mediante el programa Network v. 4.5.1.6. y aplicando la opción *reduced-median algorithm* (Bandelt et al., 1995) a partir de la variación de las secuencias de la región HVI. Este método se ha utilizado para representar la variación genética de un haplogrupo presente en diferentes poblaciones. Es una diagrama filogenético representado por un gráfico formado por círculos conectados por líneas, que se relacionan con las posiciones nucleotídicas donde se han producido cambios,

siendo la longitud de las líneas proporcional al número de cambios producidos. Cada círculo está asociado a un haplogrupo y por interpolación al individuo portador de esta variante del ADNmt. El tamaño del círculo es proporcional al número de individuos. A partir del *network* se puede estimar la edad del haplogrupo representado (Morrall et al., 1994; Saillard et al., 2000). El método de datación se basa en considerar el número de mutaciones acumuladas desde la secuencia ancestral como una función lineal de la tasa de mutación y del tiempo.

Aplicando el programa XL-Stat se han realizado análisis de componentes principales (ACP), métodos estadísticos que consisten en simplificar la complejidad de los datos reduciendo el número de dimensiones del espacio en el que se trabaja, pero con una mínima pérdida de información. En este caso se aplicó utilizando las frecuencias relativas de los haplogrupos del ADNmt para cada una de las poblaciones. Los componentes principales obtenidos mediante este método no están correlacionados, por lo que pueden ser analizados independientemente. La cantidad de variabilidad explicada por cada componente principal disminuye conforme se avanza en los diferentes componentes: el primer componente acumula la máxima variación, el segundo menos y así sucesivamente. Los valores de los componentes principales se pueden representar en mapas sintéticos que resumen los patrones espaciales de forma resumida. La aplicación de este método permite determinar qué variables determinan las diferencias entre poblaciones.

Por último, se ha utilizado un método que permite expresar la topografía genética de las poblaciones de acuerdo a la geografía. Este método se basa en el modelo de triangulación de Delaunay (Brasel and Reif, 1979) y en los trabajos de Monmonier. Este autor basó sus estudios en el desarrollo del algoritmo de máxima diferenciación, que permitían construir e identificar gráficamente las fronteras existentes entre pares de poblaciones de un área determinada (Monmonier 1973). Es un método que se ha utilizado en múltiples estudios con datos de diferente naturaleza como apellidos (Manni and Barraï, 2001) o marcadores genéticos (Bosch et al., 2000; Comas et al., 2000; Iriondo et al., 2003; Via 2006).

Para tener una visión general sobre las posibles barreras que existen en el área estudiada, se ha recopilado de la bibliografía científica las frecuencias de los haplogrupos amerindios de diferentes grupos indígenas estimando su posición geográfica utilizando la página web de Global Gazetteer Version 2.1. (<http://www.fallingrain.com/world/index.html>). Sobre esta información se ha aplicado el programa de libre acceso en la red llamado Barrier (Manni et al., 2004), que permite la

reconstrucción gráfica de estas fronteras genéticas además de asignar un valor numérico que indica la intensidad o robustez de estas fronteras.

III. RESULTADOS

III.1. Consideraciones metodológicas

Inicialmente se recopilaron 916 muestras, aunque el número final una vez eliminados los individuos emparentados y habiendo hecho los análisis de laboratorio, fue de 628 (Tabla III.1) que representa un 68,56% de efectividad.

Grupo indígena	Muestras recogidas	Muestras finales	Porcentaje efectividad
Pimas	127	49	38%
Mayos	55	55	100%
Huicholes	36	36	100%
Nahuas	252	192	76%
Otomíes	200	177	88%
Tepehuas	66	61	92%
Mayas	180	58	32%

Tabla III.1. Número final de muestras analizadas por grupo indígena

La disminución en el número final de muestras estuvo causada por la eliminación de individuos emparentados así como por ADN que no amplificó. Estos datos están reflejados en la tabla III.2. La población Maya fue la más afectada por la eliminación de parientes con 105 individuos seguida del grupo Pima con 75. Si atendemos al porcentaje respecto al total de la muestra disponible la población Pimas se redujo en un 59,05% y la Maya en un 58.33%. En la muestra de Nahuas de la Huasteca el número de familiares detectados fue de 53 (21,02%), 14 (7%) en Otomíes y 3 en Tepehuas (4,54%). No se detectaron parientes en las poblaciones Mayo y Huichol.

Grupo indígena	Muestras recogidas	Emparentadas	No amplificación	Muestras finales
Pimas	127	75	3	49
Mayos	55	-	-	55
Huicholes	36	-	-	36
Nahuas	252	53	7	192
Otomíes	200	14	9	177
Tepehuas	66	3	3	61
Mayas	180	105	17	58

Tabla III.2. Causas de reducción en el número efectivo de muestras

La efectividad de los ADN procedentes de muestras de sangre fue del 100%, y variable para el resto de muestras, presentando la menor efectividad por fallos en la amplificación en el caso de la población Maya en la que fallaron 17 muestras.

De la totalidad de estas muestras se hicieron amplificaciones de dos o tres fragmentos para obtener la secuencia completa de la región control, obteniéndose imágenes similares a las siguientes:

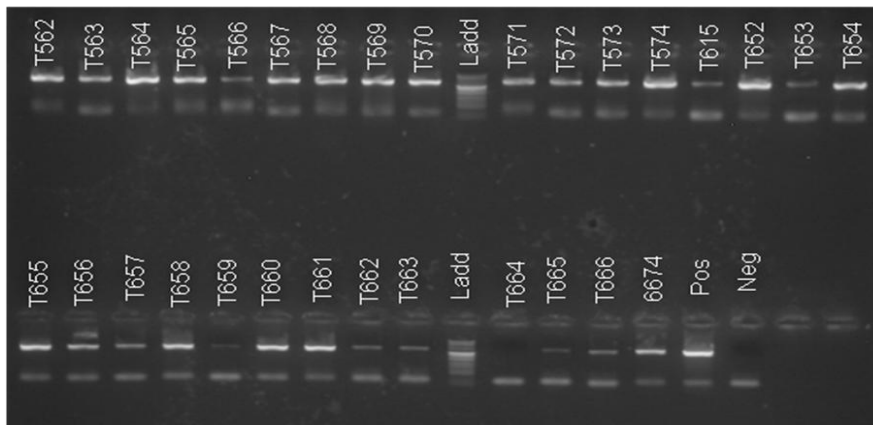


Figura III.1. Gel correspondiente al fragmento II de individuos Tepehuas

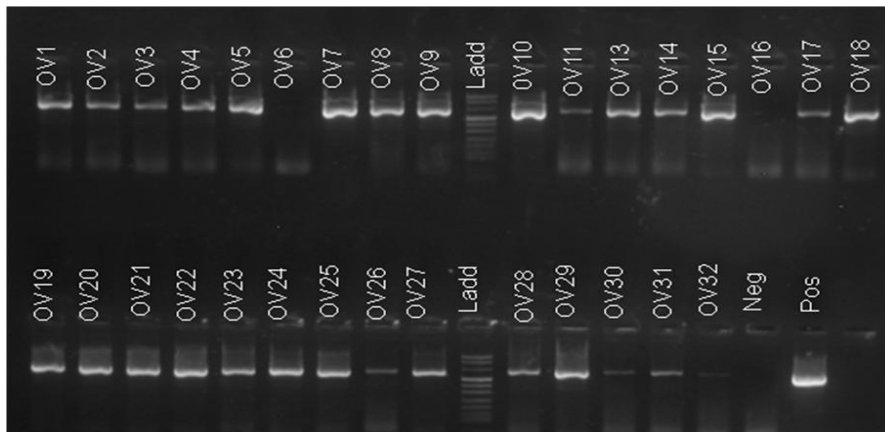


Figura III.2. Gel correspondiente al fragmento I de individuos Otomíes

Como se observa en las figuras III.1 y III.2, no aparecen bandas de amplificación para algunos individuos (por ejemplo OV06), o son tenues para otros (OV30). En ambos casos son correctas para los controles positivo y negativo. En ningún caso se detectaron contaminaciones o coincidencias con el haplotipo de la autora de esta tesis, determinado como paso previo al inicio del trabajo de laboratorio.

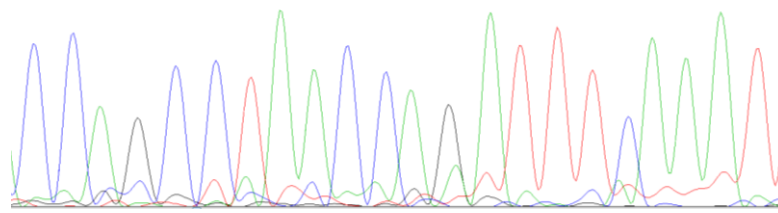


Figura III.3. Secuencia de un fragmento de HVII correspondiente al individuo Tepehua 664

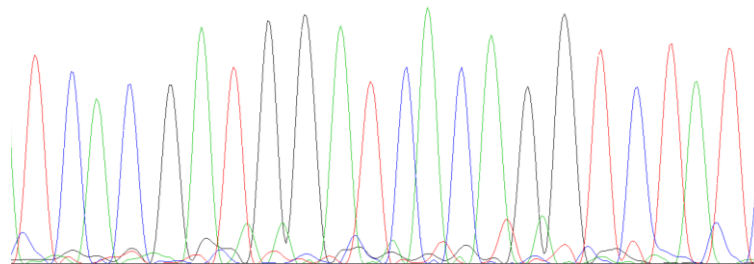


Figura III.4. Secuencia de un fragmento de HVI correspondiente al individuo Nahua 357

Las secuencias obtenidas para la mayoría de los casos fueron nítidas y sin ruido de fondo como puede verse en las figuras III.3 y III.4, correspondientes a un individuo Tepehua (T664) y uno Nahua (N357).

Grupo	HVI	HVII
HgA2	16111T, 16223T, 16290T, 16319A, 16362C	64T, 73G, 146G, 153C, 235G, 263G
HgB2	16189C, 16217C	73G, 263G
HgC1	16223T, 16298, 16327	73G, 249d, 263G, 290-291d
HgD1	16223T, 16325C, 16362C	73G, 263G

Tabla III.3. Criterios de asignación de los haplogrupos y mutaciones diagnóstico para cada uno de ellos

Algunos individuos que amplificaron sin problema para el fragmento I de D-Loop con la pareja de *primers* 15.997L-017H, no amplificaron sin embargo para el fragmento II con la pareja de *primers* 16.555L-611H, aún obteniéndose en el primer caso bandas intensas de PCR en el gel. Estos individuos fueron finalmente amplificados partiendo el fragmento II a su vez en otro par de fragmentos con las parejas de *primers* utilizadas para el grupo Huichol, 16.380L- 285H y 172L-599H (Tabla II.10. del capítulo II. Material

y Métodos). En todos los casos se obtuvieron electroferogramas de buena calidad, con picos claros y sin ruido de fondo.

Grupo	A2	B2	C1	D1	T	H	V	NC	Total
Pimas	1 (2,04)	7 (14,28)	40 (81,6)	1 (2,04)	-	-	-	-	49 (7,80)
Mayos	13 (23,6)	17 (30,9)	7 (12,7)	18 (32,7)	-	-	-	-	55 (8,75)
Huicholes	20 (55,6)	9 (25,0)	6 (16,7)	1 (2,8)	-	-	-	-	36 (5,73)
Nahuas	109 (57,4)	54 (28,4)	18 (9,5)	9 (4,7)	-	2 (1,1)	-	-	192 (30,57)
Otomíes	89 (50,28)	23 (12,99)	45 (25,42)	17 (9,60)	1 (0,56)	1 (0,56)	1 (0,56)	-	177 (28,18)
Tepehuas	38 (62,3)	15 (24,6)	3 (4,9)	2 (3,3)	-	2 (3,3)	-	1 (1,6)	61 (9,71)
Mayas	43 (74,1)	9 (15,5)	4 (6,9)	2 (3,4)	-	-	-	-	58 (9,23)
Total	313 (49,84)	134 (21,33)	123 (19,58)	50 (7,96)	1 (0,16)	5 (0,79)	1 (0,16)	1 (0,16)	628

Tabla III.4. Proporción y porcentajes de haplogrupos para cada población

La tabla III.3 refleja los criterios de asignación de los haplogrupos y las mutaciones diagnóstico para cada uno de ellos, aunque en algunos casos la asignación fue compleja. En la tabla VII.1 incluida en el apartado VII. Anexos se han volcado algunas consideraciones que se tuvieron en cuenta para definir los haplogrupos de algunos individuos. La proporción de haplogrupos asignados finalmente para cada población así como los porcentajes de los mismos, están reflejados en la tabla III.4. En esta tabla se puede observar como la población más representada en el estudio es la Nahua (30,57%) seguida de Otomíes (28,18%), la que presenta menor tamaño muestral es la Huichol (5,73%).

III.2. Resultados D-Loop para la población total

El número total de secuencias analizadas ha sido de 628 aunque finalmente se dispone únicamente 605 han quedado para el análisis poblacional conjunto, ya que en 23 de ellas no se ha podido secuenciar la región D-Loop completa. Concretamente en 2 individuos (Tabla III.5) no se pudo secuenciar la región HVI, y en 21 casos el

problema afectó a HVII. El total de secuencias utilizadas en los análisis para la HVI es por lo tanto de 626 y de 607 para HVII.

En una revisión por poblaciones se puede observar que la totalidad de las muestras de los grupos Pimas, Mayos, Huicholes y Nahuas, pudieron ser secuenciadas completamente. En las poblaciones Otomíes fallaron tres muestras, una que afectó a HVI, concretamente OV052, y dos a HVII, OV125 y OS688. Para la población Tepehua no se pudo secuenciar la región HVI del individuo TE552 ni HVII de otros cinco individuos (Tabla III.5). La población Maya ha sido la más afectada ya que aunque se pudo secuenciar la totalidad de las muestras para la región HVI, fallaron 14 de la región HVII.

Población	HVI	HVII
Pimas	-	-
Mayos	-	-
Huicholes	-	-
Otomíes	OV052	OV125, OS688,
Nahuas	-	-
Tepehuas	TE552	TE567, TE571, TE665, TE668, TE477
Mayas	-	MA093, MA103, MA116, MA096, MA100, MA082, MA117, MA102, MA115, MA119, MA110, MA114, MA087, MA109,

Tabla III.5. Muestras clasificadas por poblaciones que no se han podido secuenciar completamente

La longitud total de la D-Loop es de 1.121pb aunque el número de posiciones oscila en función de la presencia de inserciones/delecciones, conocidas también como *indels*, que pueden afectar al tamaño total de la secuencia. Es interesante hacer una descripción de estos eventos ya que permite caracterizar a las poblaciones aunque su interés filogenético sea menor (Van Oven and Kayser, 2009).

Se han detectado tres eventos *indels* en HVI, el resto corresponde a HVII. Respecto a la primera, uno de estos eventos es una inserción y los otros dos corresponden a delecciones. A efectos descriptivos, la muestra MO077 es la única de todas las estudiadas que presenta una inserción en la región HVI, concretamente en la posición 16.193 (16.193.1C) lo que implica que el tamaño real, en términos de contaje de posiciones, será de 1.122pb.

Las dos deleciones situadas en HVI corresponden a las posiciones 16.183 y 16.566. La primera de ellas afecta a 10 individuos, un 1,59% del total de regiones HVI secuenciadas en el estudio, y todos ellas proceden del grupo Mayo. La descripción de la deleción en la posición 16.566 es en cierto sentido semejante a la 16.183, ya que afecta a nueve individuos todos ellos procedentes del mismo grupo étnico, en este caso Nahuas de la Huasteca.

Un caso excepcional en cuanto a la presencia de *indels* en HVI y que requiere una descripción detallada, lo constituye el individuo NA409 procedente de la Huasteca. En él se ha detectado una deleción que afecta a siete pares de bases, entre las posiciones 16.130 y 16.216 tomando como referencia la secuencia de Anderson, esta deleción implica la pérdida del segmento GTACCAT. Es el único individuo que presenta esta característica, siendo además un suceso no descrito hasta el momento en ninguna otra población. De cualquier forma se trata de un evento puntual con poca repercusión en los estudios poblacionales tratados en esta tesis.

El panorama de inserción/deleción en HVII es muy diferente al descrito para la región HVI, se han detectado 12 deleciones y 8 inserciones. Las primeras tienen diferentes rangos de importancia atendiendo a su frecuencia poblacional. Algunas de ellas, como la 60del ó 202del afectan a un único individuo, y en este caso corresponden a los individuos NA322 y HU013 respectivamente. Sin embargo otras deleciones son mucho más frecuentes y pueden afectar a individuos de diferentes grupos. Así, las deleciones en las posiciones 522, 523 y 524 afectan a un total de 101, 384 y 257 individuos respectivamente. Estas pérdidas puntuales de un nucleótido representan un 16,63%, un 63,26% y un 42,33% respecto a los 607 fragmentos HVII secuenciados. Por otra parte, la mutación puntual 63del afecta a 22 individuos, un 3,62% del total. Esta deleción tiene cierto interés ya que podría caracterizar al grupo Pima, puesto que todos los individuos afectados proceden de estas poblaciones ubicadas en el noroeste de la región estudiada. Esta variante ha sido previamente descrita en poblaciones de la provincia de Córdoba (Argentina) (Álvarez-Iglesias et al., 2007).

Las deleciones en las posiciones 308 y 309 aparecen siempre conjuntamente y están presentes en 5 individuos de la región Huasteca. Las mutaciones que afectan a las posiciones 249del, 290del y 291del han de valorarse más profundamente, ya que son puntos diagnósticos para definir el haplogrupo C1.

Se ha detectado una deleción que afecta a cuatro individuos procedentes de la población Nahua de la Huasteca. Se trata de una deleción de los nucleótidos GGAGCA entre las posiciones 106 y 111. Es un evento infrecuente y en principio podría pensarse que los cuatro individuos están emparentados. La revisión de las bases de datos y las encuestas confirman que no comparten ningún apellido, y que los antecesores de todos ellos proceden de asentamientos indígenas distintos. Se han considerado como no emparentados y se han incluido en los análisis. Representan el 0,66% de la población total.

Por último, hay que hacer mención a la mutación 267del que implica la desaparición de una T y que afecta a 14 individuos, 12 de los cuales proceden de la población Tepehua y dos del grupo Otomí originario del valle del Mezquital.

Respecto a las inserciones se han detectado nueve, ocho de las cuales se presentan dos a dos. La primera de ellas es 296.1C y es un evento puntual, no se presenta en la tabla ya que únicamente afecta a un individuo procedente del grupo indígena Huichol. Es el mismo caso que se ha detectado para la doble inserción 313.1C- 313.2T, solo se presenta en un individuo Pima.

La inserción de una C en la posición 309 es un evento relativamente frecuente y afecta a 445 muestras, el 73,31% de los individuos estudiados. La inserción adyacente a esta última, 309.2C, también implica la inserción de una C aunque su frecuencia es mucho menor, solo 70 individuos la presentan (11,53%). La doble inserción 315.1C/.2C la presentan tres individuos, todos ellos procedentes de la Huasteca mientras que la presencia solo de la 315.1C es un evento muy frecuente, detectándose en el 98,86% de las muestras (588).

La doble inserción AC detectada en las posiciones 524.1A, 524.2C se ha detectado en 20 ocasiones estando ausente en las poblaciones Nahua, Huichol y Tepehua. En tres ocasiones se ha dado la mutación 524.1A en solitario y afecta exclusivamente a individuos Otomíes procedentes de la Sierra Otomí-Tepehua.

Los índices de diversidad son una estimación de la cantidad de variación genética que presenta una población y permiten realizar análisis comparativos entre poblaciones. Los más utilizados en los estudios de la variabilidad del ADNmt son el número de posiciones variables o sitios polimórficos (S), el número de haplotipos diferentes (K), la diversidad genética (o haplotípica) de Nei (H), la diversidad nucleotídica (π) y el

número medio de diferencias por parejas (M). En la tabla III.6 se expresan estos índices, en el caso de π y H van acompañados de la desviación estándar. También se da el valor D del test de Tajima que expresa si los datos se encuentran bajo la hipótesis neutral, es decir, si están o no sometidos a presión selectiva. El valor D se acompaña del nivel de significación expresado mediante un valor de p .

	N	k	S	π	H	M	D
Total	605	278	183	0,00993 \pm 0,00013	0,9915 \pm 0,0010	10,789	-1,7734; $p < 0.05$

Tabla III.6. Índices de diversidad poblacional para la totalidad de las muestra al analizar D-Loop completa. **N**: número de muestras; **k**: número de haplotipos diferentes; **S**: número de sitios polimórficos; **π** : diversidad nucleotídica \pm desviación estándar; **H**: diversidad haplotípica \pm desviación estándar; **M**: número medio de nucleótidos diferentes; **D**: índice de Tajima

Una primera aproximación a la diversidad poblacional muestra valores relativamente altos, el número de sitios polimórficos se eleva a 183, lo que indica que por cada 6 nucleótidos existe un punto polimórfico. Este elevado valor de S hay que tratarlo con cautela a la hora de valorarlo como estimador de la diversidad, sobre todo si se considera que 52 de estos polimorfismos son puntuales y afectan a un único individuo, mientras que 131 aparecen en más de una ocasión entre las secuencias analizadas. Este alto nivel de sitios polimórficos favorece la variación de los haplotipos que presentan también un valor elevado de 278.

Los valores de diversidad haplotípica y de diversidad nucleotídica presentan unos valores de 0,9915 \pm 0,0010 y 0,00993 \pm 0,00013 respectivamente. Al comparar dos a dos las diferencias que hay entre secuencias, el valor medio es de 10,789. Por otra parte, el estimador de neutralidad es de -1,7734 con un valor de $p < 0,005$ y por lo tanto, significativo. En el contexto de los resultados no es extraño, ya que un valor negativo de la D indica que existe un exceso de polimorfismos de baja frecuencia como resultado de una importante expansión poblacional y un deficiente efecto de la selección purificadora.

La figura III.5 representa las *pairwise differences*, y aporta la misma información general que los valores de diversidad, el modelo gráfico de distribución bi-modal es indicativo de poblaciones en expansión.

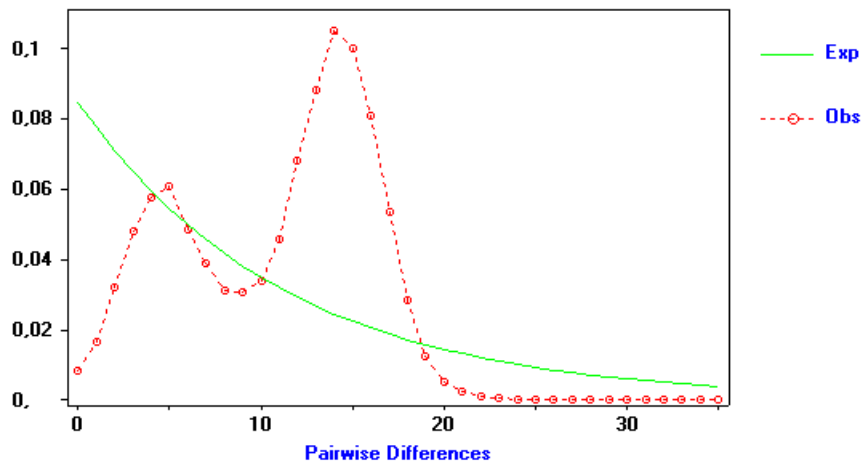


Figura III.5. Representación gráfica de las *pairwise differences* para el total de la población utilizando la región D-Loop completa.

Los resultados expuestos hasta el momento requieren ser analizados con mucha cautela ya que, al margen de estudiar una amplia región de más de 1.000pb, abarcan poblaciones de diferentes regiones con historias demográficas y evolutivas muy distintas. Un detalle importante que puede influir sobre estos primeros resultados es el hecho de que se ha incluido en los análisis haplogrupos procedentes de otras regiones geográficas, claros indicadores de mestizaje.

Los haplogrupos amerindios se han clasificado, como ya se ha explicado en el apartado II.10. Alineamiento de secuencias y determinación del haplogrupo del capítulo II. Material y Métodos, utilizando preferentemente las publicaciones más recientes (Achilli et al., 2008; O'Rourke and Raff, 2009; Perego et al 2009, 2010; Tamm et al., 2007). Para el resto de haplogrupos no amerindios así como para confirmar la clasificación de haplogrupos amerindios dudosos, se han utilizado otros recursos *online* ya explicados.

Los haplogrupos más frecuentes, aunque oscilan entre poblaciones, son A2 (49,84%), seguido de B2 (21,33%), C1 (19,58%) y por último D1 (7,96%). Los haplogrupos no amerindios representan un 1,11%. Estos están representados por tres haplogrupos distintos. El haplogrupo H se detectó en 5 ocasiones (0,79%); NA151, NA249, TE457, TE664 y OS604. El haplogrupo T2 lo presenta la muestra OV055 y el V el individuo OS688. Por último indicar que un único individuo procedente de la población Tepehua

(TE552), no ha podido ser clasificado dado que no se pudo secuenciar la región HVI y los datos de HVII no fueron lo suficientemente informativos.

Hay que advertir que el ADNmt no es el mejor estimador para determinar mestizaje, ya que es indicativo del origen femenino de los linajes. De esta manera sólo se detectarán aquellos linajes que se hayan heredado por vía materna, por lo que se detecta la combinación de cruzamientos entre varones indígenas y mujeres europeas. Además, es posible que algunos linajes de origen no amerindio hayan desaparecido debido al efecto de la deriva génica o a que sus portadoras no hayan tenido descendientes femeninos o, sencillamente a que no hayan tenido hijos. En todo caso de las siete poblaciones estudiadas únicamente se han detectado haplogrupos no amerindios en tres de ellas, Otomíes, Tepehuas y Nahuas. Es interesante detallar que estas poblaciones habitan la región montañosa conocida como sierra Otomí-Tepehua zona *a priori* de difícil acceso y muy aislada, aunque como consecuencia de estos resultados es preciso revisar y plantear nuevas hipótesis.

Una vez conocidos los haplogrupos no americanos se recalcularon los índices de diversidad poblacional (Tabla III.7). De hecho, a partir de ahora, y si no se indica lo contrario, los análisis se realizarán en base a los haplogrupos amerindios. En este caso los cálculos se realizaron excluyendo las siete secuencias europeas y la que no se ha podido clasificar. La población total para realizar estos análisis fue, por lo tanto, de 620 aunque a efectos de análisis de la región D-Loop completa se incluyeron 599.

	N	k	S	π	H	M	D
Total	599	272	180	0,00992 \pm 0,0014	0,9913 \pm 0,0010	10,773	-1,7586; p < 0,05

Tabla III.7. Índices de diversidad poblacional utilizando exclusivamente la muestras con linajes mitocondriales amerindios y para D-Loop completa. **N**: número de muestras; **k**: número de haplotipos diferentes; **S**: número de sitios polimórficos; **π** : diversidad nucleotídica \pm desviación estándar; **H**: diversidad haplotípica \pm desviación estándar; **M**: número medio de nucleótidos diferentes; **D**: índice de Tajima

Una primera aproximación a la diversidad poblacional muestra valores inferiores a los primeros cálculos realizados, incluyendo los haplogrupos europeos. Aunque sustancialmente se mantiene la misma estructura de valores, se detecta un menor número de haplotipos diferentes, disminuyendo de 278 a 272, así como un decremento del número de polimorfismos, pasando de 183 a 180. El descenso en el número de polimorfismos y de haplotipos diferentes tiene como consecuencia una

disminución en el valor de π que es de $0,00992 \pm 0,0014$, y de H que pasa a ser $0,9913 \pm 0,0010$. El número medio de diferencias, como es lógico, también disminuye ligeramente pasando de 10,789 a 10,773. Respecto a la D de Tajima también se produce una disminución, aunque continua con valor negativo y significativo ($D = -1,7586$; $p < 0.05$).

El estudio de la región D-Loop da una amplia visión sobre la diversidad nucleotídica del ADNmt que puede ser engañosa, ya que la mayor parte de los puntos polimórficos se acumulan en las regiones las dos regiones, HVI y HVII. En la figura III.6 se puede observar la distribución de los sitios polimórficos a lo largo de la región control.

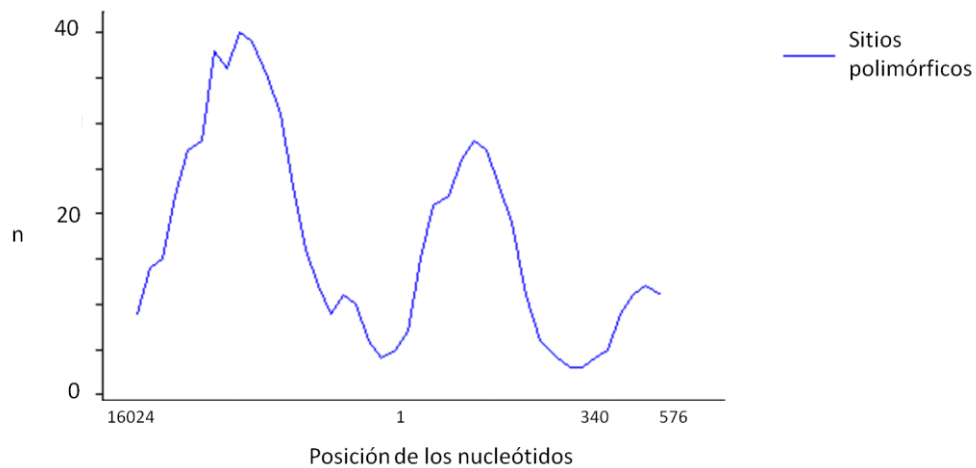


Figura III.6. Distribución de los sitios polimórficos a lo largo de la región control.

La mayor parte de los sitios polimórficos se distribuyen en HVI y en HVII detectándose una clara disminución en la región intermedia entre las posiciones 16.333 y 73, y a partir de la posición 340, nucleótido en el que termina HVII. Aprovechando esta circunstancia, y que la mayor parte de los estudios poblacionales se han centrado tradicionalmente en una u otra región hipervariable, se realizará un estudio independiente de las dos regiones.

III.3. Resultados HVI y HVII para población total

La región HVI se distribuye entre los nucleótidos 16.024 y 16.365, mientras que HVII está delimitada por los nucleótidos 73 y 340. A partir de la base de datos original se construyeron otras dos bases de datos, una que contenía exclusivamente las secuencias de HVI y otra las de HVII. Para la primera, el número total de muestras es 619, mientras que para HVII el total es 600.

Para HVI el número total de nucleótidos es de 343, la región HVII es algo más pequeña y ocupa una extensión de 268 nucleótidos. La primera de ellas presenta 89 posiciones polimórficas, 23 son eventos únicos y 66 son compartidos por varios individuos, mientras que HVII tiene tan sólo 49 polimorfismos, 14 de los cuales están representados en un único individuo y 35 son compartidos por más de un individuo. Los índices de diversidad nucleotídica son proporcionales a estos valores, obteniéndose un valor de $0,01793 \pm 0,00031$ para HVI y de $0,00936 \pm 0,00021$ para HVII (Tabla III.8).

Los valores de K y H son también superiores en la región HVI, con $K = 148$ y $H = 0,9703 \pm 0,0027$, frente a $K = 78$ y $H = 0,8570 \pm 0,0100$ para HVII. Donde los resultados presentan tendencias diferentes es en el test de Tajima. Aunque en ambos casos el resultado es negativo, el valor no es significativo para la región HVI ($D = -1,56582$; $p > 0,05$) y sí lo es para HVII ($D = -1,85140$; $p < 0,05$) (tabla III.9).

	N	K	S	π	H	M	D
HVI	619	148	89	$0,01793 \pm 0,00031$	$0,9703 \pm 0,0027$	5,99	$-1,56582$; $p > 0,05$
HVII	600	78	49	$0,00936 \pm 0,00021$	$0,8570 \pm 0,0100$	2,36	$-1,85140$; $p < 0,05$

Tabla III.8. Índices de diversidad poblacional utilizando exclusivamente la muestras con linajes mitocondriales amerindios para las regiones HVI y HVII. **N**: número de muestras; **k**: número de haplotipos diferentes; **S**: número de sitios polimórficos; **π** : diversidad nucleotídica \pm desviación estándar; **H**: diversidad haplotípica \pm desviación estándar; **M**: número medio de nucleótidos diferentes; **D**: índice de Tajima

En ambas regiones el valor medio de diferencias es distinto, siendo casi seis (5,99) para HVI y algo superior a 3 para HVII (2,36). Un reflejo de estas diferencias medias lo constituye la *pairwise differences* representada en la figura III.7, como distribuciones para las dos regiones estudiadas de forma independiente. En ellas puede observarse que el modelo de distribución para HVI y para HVII es diferente. Las diferencias

radican fundamentalmente en un modelo bimodal para la primera de ellas y unimodal para la otra.

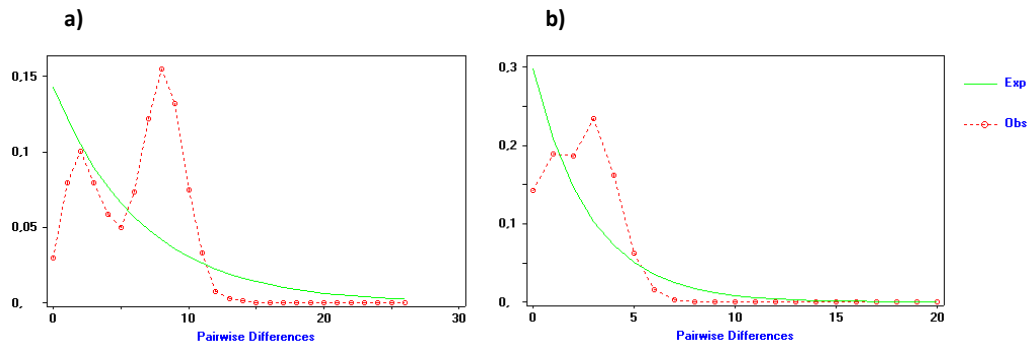


Figura III.7. Representación de *pairwise differences* para la región HVI (a) y HVII (b) independientemente.

III.4. Estudios poblacionales mediante D-Loop

En este apartado se ha abordado la diversidad de los grupos indígenas de forma independiente y trabajando con la región control completa. En la tabla III.9 se dan los valores de estos índices.

	N	k	S	π	H	M	D
Pimas	49	17	34	0,00532 ± 0,00064	0,894 ± 0,028	5,935	-0,74737; p > 0,10
Mayos	55	33	52	0,00962 ± 0,00036	0,956 ± 0,017	10,737	-0,18887; p > 0,10
Huicholes	36	25	41	0,00965 ± 0,00070	0,963 ± 0,019	10,765	0,31779; p > 0,10
Nahuas	190	94	110	0,00988 ± 0,00030	0,983 ± 0,003	10,787	-1,36540; p > 0,10
Otomíes	172	86	97	0,01013 ± 0,00025	0,982 ± 0,003	11,298	-1,03900; p > 0,10
Tepehuas	53	27	40	0,00821 ± 0,00056	0,945 ± 0,016	9,157	0,15930; p > 0,10
Mayas	44	26	50	0,00748 ± 0,00089	0,951 ± 0,020	8,336	-1,77375; p < 0,05

Tabla III.9. Índices de diversidad poblacional por grupos indígenas y para la región D-Loop completa; **k**: número de haplotipos diferentes; **S**: número de sitios polimórficos; π : diversidad nucleotídica ± desviación estándar; **H**: diversidad haplotípica ± desviación estándar; **M**: número medio de nucleótidos diferentes; **D**: índice de Tajima

En la tabla III.9 se dan los índices de diversidad para las siete poblaciones estudiadas para la región D-Loop completa. El número disponible de muestras para cada población es variable oscilando entre las 36 muestras procedentes del grupo Huichol, y los 190 que representan a los Nahuas de la Huasteca. El número muestral se reduce, ya que al considerar la región control completa hay que eliminar aquellos individuos de los que solo se dispone de una u otra de las regiones hipervariables.

Los grupos mejor representados, que más adelante permitirán realizar estudios intra-grupales o intra-comunidades, son Nahuas de la Huasteca y Otomíes, estos últimos con 172 individuos. Mayos están representados por 55 muestras y Tepehuas por 53, Pimas y Mayas están representados por 49 y 44 muestras respectivamente.

Una revisión detallada de la tabla pone de manifiesto que, atendiendo a los diferentes índices de diversidad, las poblaciones pueden agruparse en dos modelos. Uno que integraría poblaciones con elevada diversidad, y otro caracterizado por valores moderados. En el primer grupo se incluyen las poblaciones de Mayos, Huicholes, Nahuas y Otomíes. Mayas, Tepehuas y Pimas podrían incluirse en el segundo grupo.

Entre las poblaciones con índices de diversidad elevados destacan los Otomíes, que presentan los valores máximos para todos los indicadores. Presentan 97 puntos polimórficos y 86 haplotipos diferentes, además de un número medio de nucleótidos diferentes superior a 11 (11,298). Estos datos están en consonancia con los índices de diversidad que presentan valores de $\pi = 0,01013 \pm 0,00025$ y $H = 0,982 \pm 0,003$.

En orden decreciente de diversidad la siguiente población a describir es la Nahua, con 110 sitios polimórficos y 94 haplotipos distintos. La media de nucleótidos diferentes es cercana a 11 (10,787) y los índices de diversidad nucleotídica y haplotípica, son $0,00988 \pm 0,00030$ y $0,983 \pm 0,003$ respectivamente.

Las dos últimas poblaciones integradas en este modelo de alta diversidad son Mayos y Huicholes, ambos relativamente cercanos geográficamente. Estas dos poblaciones presentan un número medio de nucleótidos diferentes semejante, la población Mayo tienen un valor de 10,737 y la Huichol de 10,765. A pesar de que los primeros presentan un número mayor de sitios polimórficos (52) frente a Huicholes (41), sus índices de diversidad nucleotídica son algo menores, con $\pi = 0,00962 \pm 0,00036$ y $\pi = 0,00965 \pm 0,00070$, ligeramente superior en el segundo caso. El número de haplotipos en Huicholes es ligeramente superior a los detectados en la población

Mayo, 25 frente a 33, mientras que en los valores de H esta tendencia se invierte, siendo más elevados en Mayos ($0,956 \pm 0,017$) que en Huicholes ($0,963 \pm 0,019$).

De los tres grupos con menores índices de diversidad el que presenta valores más altos es el Tepehua, con 40 sitios polimórficos y 27 haplotipos distintos y una media de nucleótidos diferentes inferior a 10, concretamente de 9,157. Los índices de diversidad para esta población son $0,00821 \pm 0,00056$ y $0,945 \pm 0,016$, para π y H respectivamente.

Las dos poblaciones con menor diversidad son Mayas y Pimas. Los primeros presentan 50 sitios polimórficos y 26 haplotipos diferentes, con una media de 8,336 puntos polimórficos. Sus índices son $\pi = 0,00748 \pm 0,00089$ y $H = 0,951 \pm 0,020$. El grupo Pima representan la única población de la región externa a Mesoamérica, y es el que presenta los menores índices de diversidad con valores de $\pi = 0,00532 \pm 0,00064$ y de $H = 0,894 \pm 0,028$. El número de sitios polimórficos es 34 y presentan 17 haplotipos diferentes, el valor de 5,935, correspondiente a la media de sitios polimórficos diferentes, completa el análisis descriptivo de la diversidad genética de esta población.

Un aspecto que resulta interesante es el resultado de los test D de Tajima. Dos poblaciones presentan valores positivos, Tepehuas y Huicholes, aunque los resultados son no significativos ($p > 0,1$). El resto de grupos indígenas tienen un valor negativo y solamente para una de ellas, concretamente para Mayas, el valor es significativo ($D = -1,77375$, $p < 0,05$).

Estos resultados resultan interesantes sobre todos si los comparamos con los valores de D para la totalidad de las poblaciones. El valor de $D = -1,7586$ (Tabla III.7), daba a entender que nos encontrábamos ante una región con un comportamiento demográfico homogéneo, sin embargo los resultados obtenidos por poblaciones nos plantean otro panorama y permiten postular la idea de una alta diversidad en cuanto al comportamiento demográfico.

En las figuras III.8 y III.9 se representan los *pairwise differences* para las siete poblaciones estudiadas. En la figura III.8 la distribución de las cuatro poblaciones es muy parecida presentando todos los grupos gráficas bimodales, aunque los valores máximos no son coincidentes. Llama la atención el comportamiento del grupo Nahua, que presenta un desplazamiento en su primer máximo modal y el segundo atenuado

respecto a las otras tres poblaciones. También la población Huichol tiene particularidades relevantes como un atenuado valor en la primera moda y un elevado valor, el máximo para todas las poblaciones, en la segunda.

La figura III.9 representa la distribución de las tres poblaciones restantes: Pimas, Otomíes y Mayos. Estas poblaciones tiene un comportamiento ligeramente distinto a los descritos para la III.8, y no cumplen claramente el modelo bimodal. Se observa como detalle más relevante, que Otomíes y Mayos presentan muy atenuada la primera moda que prácticamente ni existe en la última. En Otomíes, al margen del ya comentado bajo valor en la primera moda, se observa la segunda un tanto desplazada hacia la derecha coincidiendo en este aspecto con la población Nahuas.

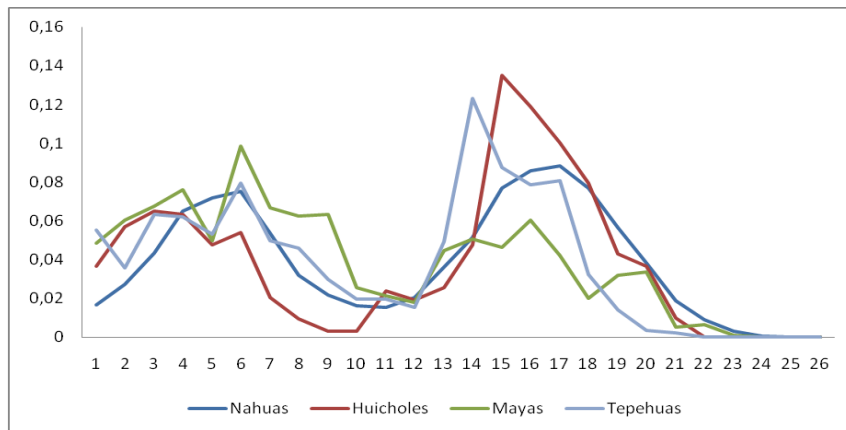


Figura III.8. *Pairwise differences* de la región D-Loop para cuatro poblaciones indígenas. El eje X representa el número de diferencias y el Y la frecuencia

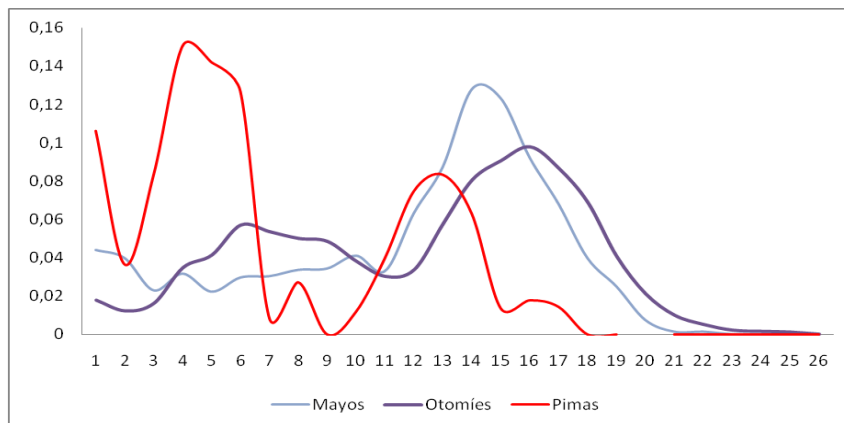


Figura III.9. *Pairwise differences* de la región D-Loop para tres poblaciones indígenas. El eje de la X representa el número de diferencias y el Y la frecuencia

Pero la población que tiene un comportamiento más atípico es la Pima. Aunque se puede apreciar una distribución bimodal, llaman la atención dos detalles. El primero de ellos es que es la única población que presenta el primer valor modal mayor que el segundo, y además es destacable el desplazamiento hacia valores menores que presenta el segundo máximo modal.

Ya se ha descrito con anterioridad la distribución de la frecuencia de haplogrupos por grupos indígenas (Tabla III.4). Ahora bien, revisando en detalle su distribución (Figura III.10) se puede describir un modelo general. Dentro del contexto la población Pima presenta un porcentaje de haplogrupos muy distinto al resto de las poblaciones. Este modelo se caracteriza por una ausencia casi total de haplogrupos A2 y D1, representados por un único individuo cada uno, y una elevada proporción de haplogrupo C1, llegando a alcanzar un 81,63%. En esta población los valores de haplogrupo B2 son moderados (n = 7; 14,29%).

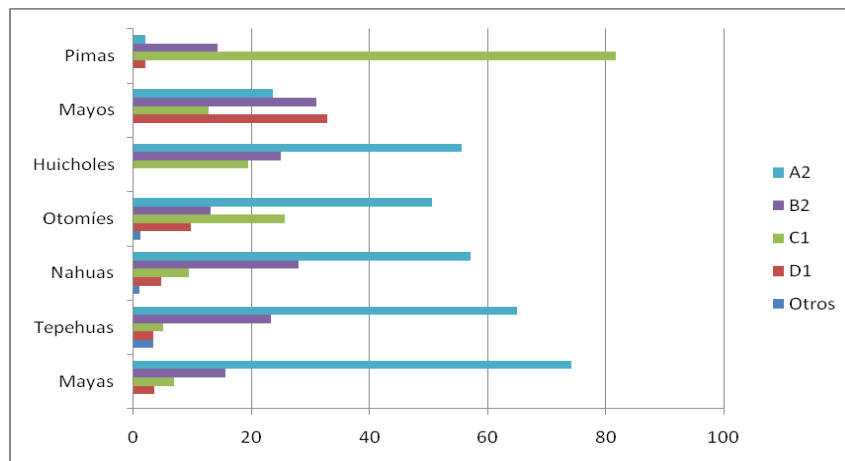


Figura III.10. Representación gráfica de los porcentajes de los diferentes haplogrupos en las siete poblaciones analizadas. En el apartado Otros se incluyen los haplogrupos no amerindios; H, V y T2.

Otra población con una frecuencia de haplogrupos relativamente particular es la Mayo. Este grupo presenta elevados valores de los haplogrupos D1 (n = 18; 32, 73%) y B2 (n = 17; 30,91%), y menor representación del A2 (n = 13; 23,4%) y C1 (n = 7; 12,73%).

Las otras cinco poblaciones presentan un modelo relativamente semejante, caracterizado por una elevada presencia del haplogrupo A2 y B2, C1 y D1, en orden decreciente. En este modelo general solo hay una excepción en la población Otomí,

donde la frecuencia del haplogrupo C1 (n = 45; 25,86%) es superior a la del B2 (n = 23; 13,22%).

Una revisión de la distribución de la frecuencia de los haplogrupos sugiere la existencia de una posible relación geográfica, por lo que se han calculado los índices de correlación de Pearson entre las frecuencias de los haplogrupos y las coordenadas de cada uno de los grupos indígenas, expresadas en longitud y latitud. La tabla III.10 muestra los resultados, en la hemimatriz inferior se representan los coeficientes de correlación y en la superior la probabilidad asociada.

	HgA2	HgB2	HgC1	HgD1	Otros	Latitud	Longitud
HgA2	1	0,938	0,025	0,430	0,370	0,000	0,008
HgB2	0,037	1	0,249	0,279	0,887	0,999	0,500
HgC1	-0,816	-0,504	1	0,664	0,433	0,036	0,186
HgD1	-0,358	0,477	-0,202	1	0,662	0,445	0,301
Otros	0,403	0,067	-0,356	-0,203	1	0,314	0,611
Latitud	-0,972	-0,001	0,787	0,347	- 0,447	1	0,021
Longitud	0,884	-0,309	-0,565	-0,458	0,263	-0,830	1

Tabla III.10. Coeficientes de correlación entre la frecuencia de los haplogrupos amerindios y las coordenadas geográficas expresadas como latitud y longitud. La hemimatriz inferior expresa los índices de correlación de Pearson y la superior la probabilidad asociada. En negrita se expresan los coeficientes y las probabilidades significativas

Los resultados muestran cinco correlaciones significativas aunque una de ellas tiene poca importancia para los objetivos de este trabajo, ya que expresa con toda lógica, la correlación entre latitud y longitud.

Existe una correlación negativa entre la distribución del haplogrupo A2 y el C1, el aumento de la frecuencia de un haplogrupo implica la disminución de la frecuencia del otro. En la figura III.11 puede observarse este resultado.

Las poblaciones más meridionales tenderán a presentar, al menos en el área estudiada, mayor frecuencia del haplogrupo A2 y menor del C1, por lo que existe un claro gradiente en cuanto a la frecuencia de ambos haplogrupos. Por la misma lógica

las poblaciones distribuidas a elevadas latitudes presentan menor frecuencia del haplogrupo A2 y mayor del C1.

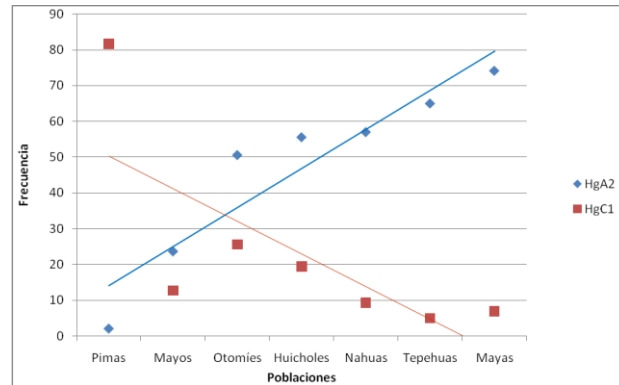


Figura III.11. Representación gráfica de las tendencias entre los haplogrupos A2 y C1. Las líneas marcan las tendencias de ambos haplogrupos

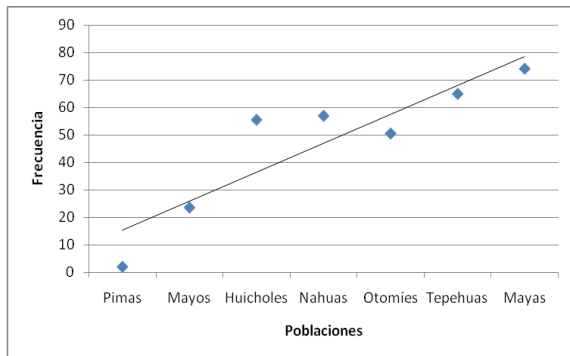


Figura III.12. Representación gráfica de la correlación entre la frecuencia del haplogrupo A2 y la latitud. Las poblaciones en el eje de las X están ordenadas según su latitud, de mayor a menor.

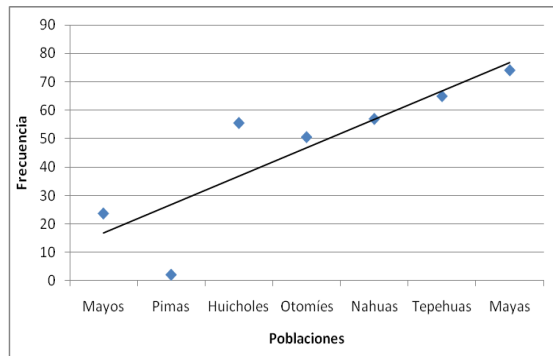


Figura III.13. Representación gráfica de la correlación entre la frecuencia del haplogrupo A2 y la longitud. Las poblaciones en el eje de las X están ordenadas según su longitud, de menor a mayor.

Otros resultados destacables son la correlación detectada entre el haplogrupo A2 y la latitud ($r = -0,972$; $p = 0,000$) y la longitud ($r = 0,884$; $p = 0,008$). El modelo de distribución resulta interesante, ya que la presencia de A2 disminuye hacia latitudes meridionales y aumenta hacia las regiones orientales. La figura III.12 representa la tendencia de las frecuencias del haplogrupo A2 respecto a la latitud, mientras que la figura III.13 refleja la relación con la longitud.

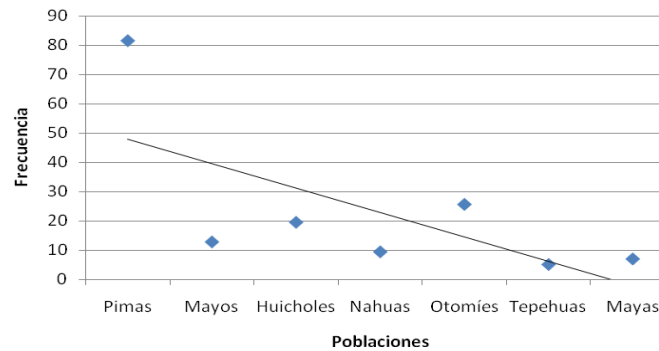


Figura III.14. Representación gráfica de la correlación entre la frecuencia del haplogrupo C1 y la latitud. Las poblaciones en el eje de las X están ordenadas según su latitud, de Mayor a menor.

También resulta significativa la correlación entre la distribución del haplogrupo C1 y la latitud; la presencia de C1 aumenta con la latitud (Figura III.14). Se espera, por lo tanto, que las poblaciones más septentrionales presenten mayores valores de C1.

III.5. Estudios poblacionales mediante HVI y HVII

Para un análisis más detallado de las poblaciones se han realizado los cálculos de los índices de diversidad, trabajando independientemente con las regiones HVI y HVII. En la tabla III.11 se dan los valores de estos índices.

Al considerar HVI las siete poblaciones pueden clasificarse en dos grupos. El primero estaría formado por Pimas, Tepehuas y Mayas, y se caracterizaría por presentar valores relativamente bajos de π y H con desviaciones elevadas, como puede apreciarse en las figuras III.15 y III.16. En el otro grupo podrían incluirse el resto de poblaciones, con valores de diversidad elevados y desviaciones relativamente bajas.

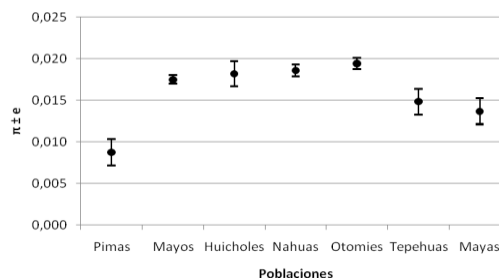


Figura III.15. Representación gráfica de los valores de π y sus errores estándares para las siete poblaciones estudiadas (análisis de HVI)

Grupo		N	K	S	π	H	M	D
Pimas	HVI	49	17	7	0,00873 \pm 0,00160	0,707 \pm 0,0019	2,986	-0,67896; p > 0,10
	HVII	49	11	11	0,00567 \pm 0,00051	0,832 \pm 0,0290	1,502	-1,14346; p > 0,10
Mayos	HVI	55	23	21	0,01750 \pm 0,00057	0,895 \pm 0,0005	5,966	0,59715; p > 0,10
	HVII	55	17	17	0,00957 \pm 0,00087	0,866 \pm 0,0300	2,537	-0,97524; p > 0,10
Huicholes	HVI	36	24	18	0,01818 \pm 0,00150	0,916 \pm 0,0310	6,218	0,25402; p > 0,10
	HVII	36	10	11	0,00880 \pm 0,00075	0,767 \pm 0,0610	2,322	-0,11279; p > 0,10
Nahuas	HVI	190	59	59	0,01859 \pm 0,00068	0,962 \pm 0,0060	6,228	-1,16272; p > 0,10
	HVII	190	22	26	0,00852 \pm 0,00035	0,802 \pm 0,0230	2,190	-1,22272; p > 0,10
Otomíes	HVI	173	57	58	0,01941 \pm 0,00066	0,971 \pm 0,0040	6,638	-1,01213; p > 0,10
	HVII	173	19	21	0,00783 \pm 0,00027	0,793 \pm 0,0200	2,067	-1,01572; p > 0,10
Tepehuas	HVI	53	25	15	0,01485 \pm 0,00155	0,831 \pm 0,0330	5,079	-0,19046; p > 0,10
	HVII	53	9	12	0,00785 \pm 0,00061	0,855 \pm 0,0280	2,073	0,15882; p > 0,10
Mayas	HVI	44	32	25	0,01367 \pm 0,00158	0,913 \pm 0,0260	4,675	-1,06081; p > 0,10
	HVII	44	16	15	0,01118 \pm 0,00124	0,874 \pm 0,0380	2,941	-0,63412; p > 0,10

Tabla III.11. Índices de diversidad poblacional por grupos indígenas, para las regiones HVI y HVII independientemente; **k**: número de haplotipos diferentes; **S**: número de sitios polimórficos; **π** : diversidad nucleotídica \pm desviación estándar; **H**: diversidad haplotípica \pm desviación estándar; **M**: número medio de nucleótidos diferentes; **D**: índice de Tajima

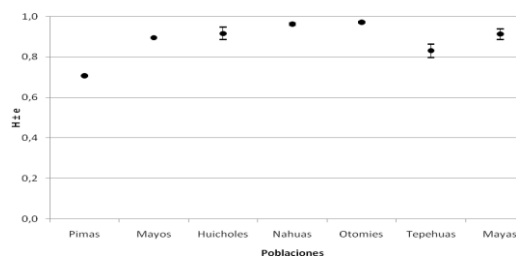


Figura III.16. Representación gráfica de los valores de H y sus errores estándares para las siete poblaciones estudiadas (análisis de HVI)

La información obtenida al estudiar HVII es ligeramente diferente. Los valores de π son algo más bajos que los detectados en HVI, aunque mantienen prácticamente el mismo modelo entre las poblaciones. Una de las diferencias más remarcable es la que presenta la población Maya, con índices ligeramente superiores a las poblaciones Otomíes y Tepehuas (Figura III.17).

Las estimas de H en la región HVII presentan unas diferencias bastante marcadas, la más importante es la obtenida para la población Pima (Figura III.18). Este grupo indígena presenta de forma general bajos valores de diversidad para todos los estimadores utilizados, pero para H el valor mínimo lo presenta la población Huichol ($H = 0,767 \pm 0,0610$), seguida de Otomíes ($H = 0,793 \pm 0,0200$ y Nahuas de la Huasteca ($H = 0,802 \pm 0,0230$).

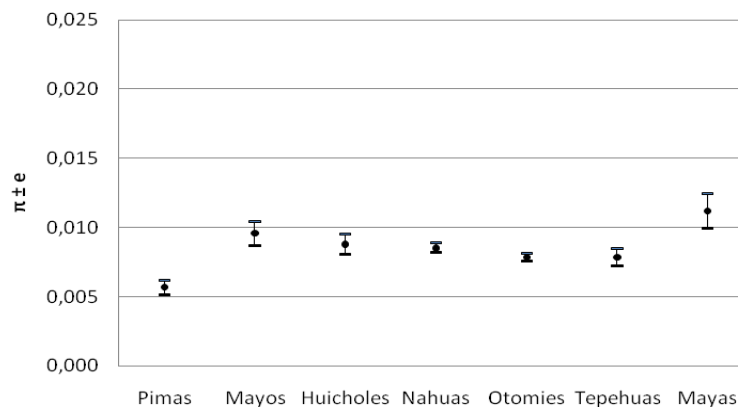


Figura III.17. Representación gráfica de los valores de π y sus errores estándares para las siete poblaciones estudiadas (análisis de HVII).

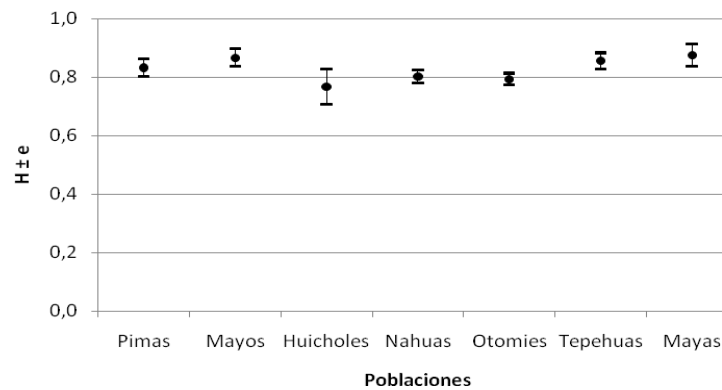


Figura III.18. Representación gráfica de los valores de H y sus errores estándares para las siete poblaciones estudiadas (análisis de HVII).

El número medio de diferencias (Tabla III.9, Figura III.19) es mayor en la región HVI que en la HVII. Los valores máximos para la HVI los presentan las poblaciones Otomíes (6,638), Nahuas (6,228) y Huichol (6,218) y el mínimo la población la Pima (2,986). Respecto a HVII, son Mayas (2,941) y Mayos (2,537) los que presentan mayores valores medios, y nuevamente Pimas, los que presentan los valores menores (1,502).

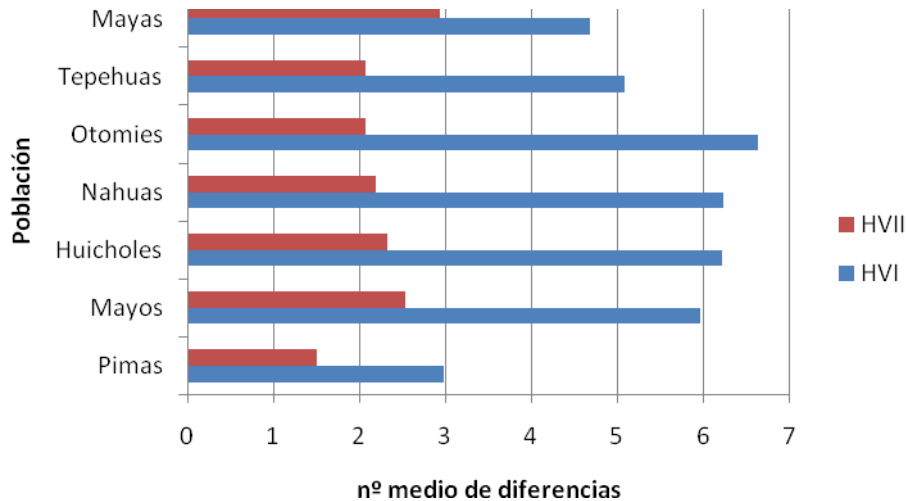


Figura III.19. Número medio de diferencias para las siete poblaciones estudiadas (análisis HVI y HVII)

Estos mismos resultados se pueden observar al representar gráficamente el número medio de polimorfismos y el número de haplotipos para cada una de las dos regiones (Figura III.20 y III.21).

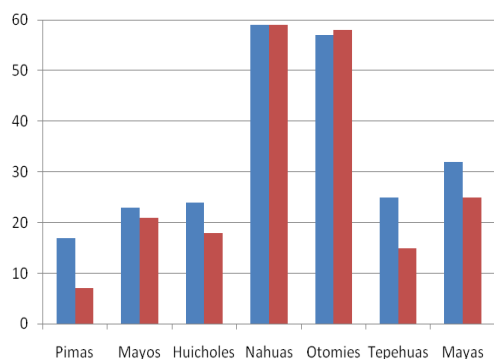


Figura III.20. Valor absoluto de los sitios polimórficos (S) y el número de diferentes haplotipos (k) para HVI

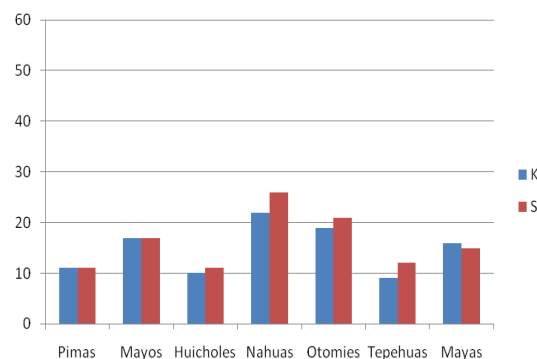


Figura III.21. Valor absoluto de sitios polimórficos (S) y el número de diferentes haplotipos (k) para HVII

En las figuras puede apreciarse como el mayor número de polimorfismos así como el de haplotipos distintos aparecen, tanto para HVI como para HVII, en las poblaciones Nahuas y Otomíes. Nuevamente, la población con valores más bajos es la Pima. El resto de poblaciones mantienen valores intermedios.

III.6. Estudios intrapoblacionales: subestructura interna

Una vez descritas las diversidades poblacionales mediante distintos parámetros, el siguiente paso es determinar la relación que existe entre ellas. Pero antes de exponer los resultados de las pruebas estadísticas para determinar las relaciones intergrupales, es preciso hacer una reflexión sobre los dos grupos que presentan los índices de diversidad más elevados. Tanto la población Nahua de la Huasteca como la Otomí, presentan unas particularidades históricas y geográficas que podrían haber afectado a su estructura genética interna, cuestión que se analizará en el capítulo IV. Discusión. Estudiar la sub-estructura de estas dos poblaciones, determinará el efecto que la cultura y la geografía ha tenido sobre la conformación genética poblacional. Estas características ya se tuvieron en cuenta al diseñar el muestreo, integrando grupos procedentes de varios municipios y en cada uno de ellos varias comunidades, como se ha detallado en los apartados II.3. Estrategias de muestreo y II.4. Toma de muestras en las comunidades indígenas del capítulo II. Material y Métodos.

III.6.I. Otomíes del Valle vs Otomíes de la Sierra

Los primeros resultados obtenidos al analizar la región D-Loop completa, muestran una mayor diversidad para todos los indicadores en las poblaciones de la sierra. Así, M tiene un valor de 11,55 para Otomíes de la Sierra frente a 10,88 de los grupos del Valle. Esta misma tendencia se detecta para π , $0,01034 \pm 0,00036$ frente a $0,00975 \pm 0,00035$, y para H , $0,972 \pm 0,008$ y $0,968 \pm 0,008$. Los cálculos de los índices a partir de la HVI y la HVII de forma independiente, reflejan la misma tendencia (Tabla III.12).

En el anexo se adjunta la distribución *pairwise differences* para las dos poblaciones (Figuras VII.5 y VII.6) al considerar la región D-Loop completa. Ambos modelos son muy parecidos y representan poblaciones en expansión, con una clara distribución bimodal.

En la tabla III.13 se da, en valores absolutos y en porcentaje, la distribución de los haplogrupos para ambas poblaciones. En los dos casos se mantiene un patrón en el que el haplogrupo más frecuente es A2, seguido de C1, de B2 y por último D1. Como se ha comentado anteriormente, se detectaron dos haplogrupos europeos en las poblaciones de la sierra que no se han tenido en cuenta para realizar estos cálculos, al considerarse que son de procedencia europea e indicativos de flujo génico y mestizaje.

Otomíes del Valle							
	N	k	S	π	H	M	D
HVI	81	34	37	0,01785 \pm 0,00088	0,964 \pm 0,008	6,10	-0,5751; $p > 0,10$
HVII	81	13	9	0,00729 \pm 0,00029	0,802 \pm 0,029	1,92	0,1577; $p > 0,10$
D-Loop	82	62	39	0,00975 \pm 0,00035	0,968 \pm 0,008	10,88	-0,3762; $p > 0,10$.
Otomíes de la Sierra							
	N	k	S	π	H	M	D
HVI	91	37	47	0,02035 \pm 0,00093	0,956 \pm 0,009	6,59	-0,7908; $p > 0,10$
HVII	91	15	16	0,00828 \pm 0,00045	0,791 \pm 0,026	2,19	-1,1768; $p > 0,10$
D-Loop	91	83	52	0,01034 \pm 0,00036	0,972 \pm 0,008	11,55	-0,9652; $p > 0,10$

Tabla III.12. Índices de diversidad poblacional para los dos grupos Otomíes considerando de forma independientemente D-Loop, HVI y HVII; **k**: número de haplotipos diferentes; **S**: número de sitios polimórficos; **π** : diversidad nucleotídica \pm desviación estándar; **H**: diversidad haplotípica \pm desviación estándar; **M**: número medio de nucleótidos diferentes; **D**: índice de Tajima

Grupo		A2	B2	C1	D1	Total
Otomíes Valle	N	40	13	24	6	83
	%	48,19	15,66	28,91	7,22	100
Otomíes Sierra	N	49	10	22	10	91
	%	53,84	10,98	24,17	10,98	100

Tabla III.13. Distribución, en valores absolutos y porcentajes, de los distintos haplogrupos en los dos grupos Otomíes. En estos cálculos no se han incluido los dos haplogrupos europeos detectados en Otomíes de la Sierra

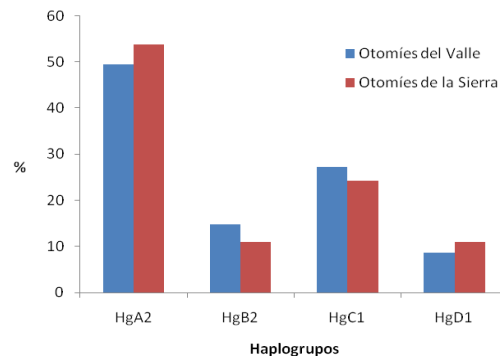


Figura III.22. Distribución en porcentajes de los distintos haplogrupos en los dos grupos Otomíes. En la representación no se han incluido los dos haplogrupos europeos

La lectura de la figura III.22 permite hacer una comparación de los porcentajes de los haplogrupos por población. Las diferencias no son muy marcadas, observándose que las frecuencias de los haplogrupos B2 y C1 son mayores en Otomíes del Valle, mientras que los haplogrupos A2 y haplogrupos D1 son más frecuentes en Otomíes de la Sierra.

Para determinar el grado de diferenciación que presentan estas dos poblaciones, se ha calculado el índice exacto de diferenciación poblacional, obteniéndose un valor significativo ($p = 0,0000$). Estos resultados sugieren que las diferencias entre las dos poblaciones Otomíes son lo suficientemente importantes como para considerar que se tratan de grupos genéticamente diferenciados. Como consecuencia de estos resultados, en los análisis comparativos se considerarán Otomíes de la Sierra y Otomíes del Valle como poblaciones distintas. Este resultado avala las propuestas lingüísticas, que proponen unas diferencias acusadas (INALI, 2007, 2008) entre las lenguas habladas por los grupos del valle del Mezquital y los de la sierra Otomí-tepehua.

Para profundizar en la estructura poblacional de los Otomíes, se estudiaron las muestras primero a nivel de municipios y posteriormente de comunidades. En estas nuevas agrupaciones, y para aprovechar el máximo de información molecular y compensar la posible pérdida de potencial estadístico asociado a la reducción del tamaño muestral, se utilizó la región D-Loop completa.

En el estudio a nivel municipal hay que advertir que El Cardonal coincide con los Otomíes del Valle, ya que todas las comunidades de esta región pertenecen a este municipio.

	N	K	S	π	H	M	D
Tenango de Doria	61	32	59	0,01000 ± 0,00039	0,952 ± 0,00022	11,168	-0,38749; p > 0,10
San Bartolo Tutotepec	30	26	66	0,01084 ± 0,00075	0,989 ± 0,00016	12,103	-1,03156; p > 0,10
El Cardonal	81	39	61	0,00975 ± 0,00035	0,968 ± 0,00800	10,880	-0,37623; p > 0,10

Tabla III.14. Índices de diversidad poblacional por municipios Otomíes y para la región D-Loop completa; **k**: número de diferentes haplotipos; **S**: número de sitios polimorfismos; π : diversidad nucleotídica ± desviación estándar; **H**: diversidad haplotípica ± desviación estándar; **M**: número medio de nucleótidos diferentes; **D**: índice de Tajima.

En la tabla III.14 se muestran los resultados de los índices de diversidad a nivel municipal para la población Otomí. El Cardonal es el único municipio del valle del Mezquital, mientras que Tenango de Doria y San Bartolo Tutotepec pertenecen a la sierra Otomí-tepehua. En general se puede apreciar que los valores de la diversidad son menores en el municipio del Cardonal; $\pi = 0,00975 \pm 0,00035$, $H = 0,968 \pm 0,00800$ y $M = 10,880$. Respecto a los municipios de la sierra, observamos que San Bartolo Tutotepec presenta resultados ligeramente superiores a Tenango de Doria. Los índices de D de Tajima son todos ellos no significativos.

Las representación de las *pairwise differences* (Figuras III.23, III.24 y III.25) para los tres municipios muestran un modelo bimodal para Tenango y El Cardonal, mientras que para San Bartolo las modas parecen estar muy atenuadas.

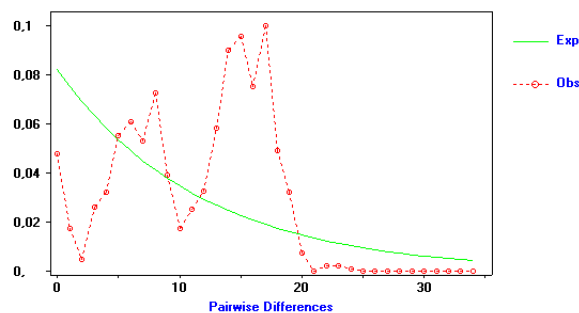


Figura III.23. *Pairwise differences* en el municipio de Tenango de Doria al estudiar la región D-Loop completa

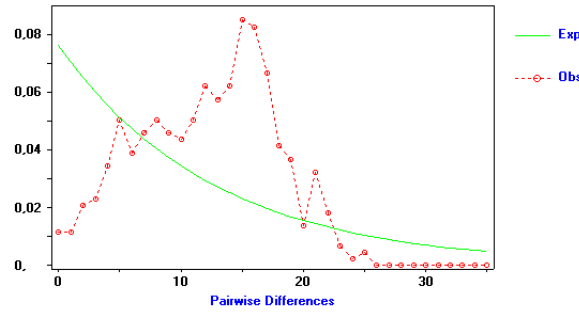


Figura III.24. *Pairwise differences* en el municipio de San Bartolo Tutotepec al estudiar la región D-Loop completa

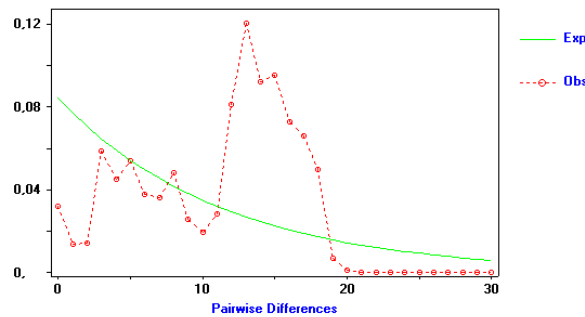


Figura III.25. *Pairwise differences* en el municipio de el Cardonal al estudiar la región D-Loop completa

El cálculo de las distancias genéticas entre los tres municipios (Tabla III.15) expresado por F_{st} , muestra diferencias significativas para todas las comparaciones, tanto entre los municipios de la sierra ($F_{st} = 0,01842$; $p = 0,03180$), como entre la sierra y el Cardonal.

	El Cardonal	San Bartolo Tutotepec	Tenango de Doria
El Cardonal		0,00901*	0,03180*
San Bartolo Tutotepec	0,01928		0,01842*
Tenango de Doria	0,03180	0,01842	

Tabla III.15. Distancias genética F_{st} entre parejas de municipios Otomíes utilizando la región D-Loop completa. En la hemimatriz inferior se muestran los valores de F_{st} y en la superior el valor de p . *, significación

Estos mismos resultados se obtienen al aplicar el test exacto de diferenciación poblacional, en el que todos los resultados arrojan diferencias significativas (Tabla III.16).

	El Cardonal	San Bartolo Tutotepec	Tenango de Doria
El Cardonal		+	+
San Bartolo Tutotepec	0,00000		+
Tenango de Doria	0,00000	0,00000	

Tabla III.16. Diferenciación entre parejas de municipios utilizando la región D-Loop completa para los municipios Otomíes. En la hemimatriz inferior se muestran los valores de significación (p) del test exacto de diferenciación poblacional, en la superior el resultado del test; (+) diferencias significativas; (-) diferencias no significativas

El análisis a nivel de comunidad implica trabajar con seis poblaciones. El Buena, el Deca y San Andrés Dabothá, pertenecen al municipio de El Cardonal, mientras que Santa Mónica y San Nicolás, son de Tenango de Doria. Santiago es el único municipio que pertenece al municipio de San Bartolo Tutotepec. En la tabla III.17 se muestran los valores de los índices de diversidad, considerando la región D-Loop completa, para estas comunidades.

	N	k	S	π	H	M	D
El Buena	24	19	46	0,01055 \pm 0,00062	0,975 \pm 0,021	11,790	-0,16578; p > 0,10
El Deca	30	21	40	0,00846 \pm 0,00068	0,966 \pm 0,019	9,448	-0,23659; p > 0,10
San Andrés Dabothá	27	16	44	0,00959 \pm 0,00073	0,940 \pm 0,025	10,707	-0,23428; p > 0,10
San Nicolás	32	16	45	0,00999 \pm 0,00055	0,910 \pm 0,033	11,166	-0,03225; p > 0,10
Santa Mónica	29	19	48	0,00956 \pm 0,00076	0,966 \pm 0,017	10,680	-0,47228; p > 0,10
Santiago	30	26	66	0,01084 \pm 0,00075	0,989 \pm 0,013	12,103	-1,03156; p > 0,10

Tabla III.17. Índices de diversidad poblacional por comunidades Otomíes y para la región D-Loop completa; **k**: número de diferentes haplotipos; **S**: número de sitios polimórficos; **π** : diversidad nucleotídica \pm desviación estándar; **H**: diversidad haplotípica \pm desviación estándar; **M**: número medio de nucleótidos diferentes; **D**: índice de Tajima

Inicialmente no parecen existir modelos administrativos ni geográficos en cuanto a la distribución de los valores de los índices. Las comunidades más diversas atendiendo a los valores de π , son el Buena (0,01055 \pm 0,00062) que pertenece a un municipio del valle y la de Santiago (0,01084 \pm 0,00075), adscrita a San Bartolo Tutotepec en la sierra. El resto de comunidades, siempre haciendo referencia a la diversidad nucleotídica, presenta unos valores semejantes que oscilan entre 0,00999 de San Nicolás, y 0,00959 de San Andrés Dabothá. Una excepción lo constituye la comunidad de El Deca, con un índice de diversidad nucleotídica algo menor (0,00846). Este modelo en la distribución de los valores de π , se mantiene al calcular el número

medio de nucleótidos diferentes, donde el valor máximo lo representa Santiago (12,103) y el mínimo El Deca (9,448). Respecto a los índices de diversidad haplotípica, los resultados son ligeramente diferentes, en este caso la población con valores más elevados sigue siendo Santiago ($H = 0,989 \pm 0,013$), seguido de El Buena ($H = 0,975 \pm 0,021$), y el valor mínimo lo presentan las comunidades de San Andrés Daboxthá, procedente del municipio de El Cardonal en el valle del Mezquital ($H = 0,940 \pm 0,025$) y San Nicolás ($H = 0,910 \pm 0,033$), perteneciente al municipio de Tenango de Doria, en la sierra. En todos los casos los valores de la D de Tajima son negativos y no significativos.

Las distancias genéticas entre comunidades medidas con Fst (Tabla III.18), aportan unos interesantes resultados ya que las comunidades del valle, todas ellas pertenecientes al municipio de El Cardonal, presentan valores no significativos. Por otro lado, las comunidades de la sierra tienen distancias genéticas significativas tanto al calcularlas entre ellas como al compararlas con las comunidades del valle.

	El Buena	El Deca	San Andrés Daboxthá	San Nicolás	Santa Mónica	Santiago
El Buena		0,22523	0,12613	0,00901*	0,00000*	0,01802*
El Deca	0,00792		0,05405	0,00000*	0,00000*	0,00000*
San Andrés Daboxthá	0,01377	0,01550		0,00000*	0,00000*	0,00000*
San Nicolás	0,03761	0,05535	0,06176		0,00000*	0,00000*
Santa Mónica	0,02436	0,03337	0,04466	0,02789		0,01802*
Santiago	0,01425	0,02190	0,03195	0,03647	0,01390	

Tabla III.18. Distancias genética Fst entre parejas de las comunidades Otomíes utilizando la región D-Loop completa. En la hemimatriz inferior se muestran los valores de Fst y en la superior el valor de p. *, significación

En este mismo sentido se observan los resultados el cálculo de los índices exactos de diferenciación poblacional (Tabla III.19). Las comunidades del valle no son distintas entre ellas pero sí que lo son al compararlas con las de la sierra. Estas últimas a su vez, se diferencian entre ellas y con las del valle.

	El Buena	El Deca	San Andrés Daboxthá	San Nicolás	Santa Mónica	Santiago
El Buena		-	-	+	+	+
El Deca	0.18000		-	+	+	+
San Andrés Daboxthá	0.14355	0.06635		+	+	+
San Nicolás	0.00000	0.00000	0.00000		+	+
Santa Mónica	0.00065	0.00000	0.00000	0.00425		+
Santiago	0.00965	0.00000	0.00215	0.00000	0.01150	

Tabla III.19. Diferenciación entre parejas de municipios utilizando la región HVI para los municipios Otomíes. En la hemimatriz inferior se muestran los valores de significación (p) del test exacto de diferenciación poblacional, en la superior el resultado del test; (+) diferencias significativas; (-) diferencias no significativas

La representación de los *pairwise differences* de las comunidades Otomíes pueden consultarse en el capítulo VII. Anexo (Figuras de VII.7 a VII.12). Todas ellas presentan un mismo modelo, aunque hay que analizarlas con cierta cautela ya que las fluctuaciones observadas pueden estar sometidas al tamaño muestral utilizado.

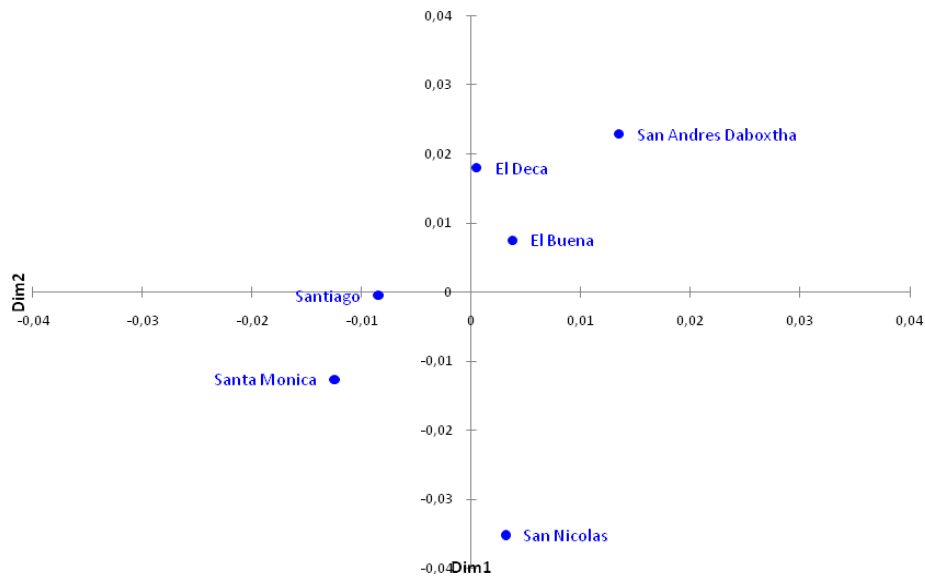


Figura III.26. Multidimensional scaling que reconstruye, a partir de las distancias Fst y utilizando la información extraída de la región D-Loop completa, la relación entre las seis comunidades Otomíes estudiadas (stress = 0,068).

Criterio	Agrupación	Grupo	Fuente de variación	Porcentaje varianza	Índice fijación	P
Administrativo	El Cardonal	El Deca, El Buena, San Andrés Daboxthá	Dentro de poblaciones	96.31	Fst = 0,03693	0,0000*
	Tenango de Doria	Santa Mónica, San Nicolás	Entre poblaciones dentro de grupos	1.94	Fsc = 0,01979	0,0000*
	San Bartolo Tutotepec	Santiago	Entre grupos	1.75	Fct = 0,01748	0,1026
Geográfica	Otomíes Valle	El Deca, El Buena, San Andrés Daboxthá	Dentro de poblaciones	92.41	Fst = 0,07589	0.00000*
	Otomíes Sierra	Santa Mónica, San Nicolás, Santiago	Entre poblaciones dentro de grupos	4.29	Fsc = 0,04440	0.00196*
			Entre grupos	3.30	Fct = 0,03296	0.19159

Tabla III.20. Resultados del cálculo de AMOVAS para diferentes agrupaciones de las comunidades Otomíes.

En la figura III.26 se representa un MDS de las comunidades Otomías utilizando la distancia Fst. La distribución espacial de las comunidades muestra un claro *cluster* que agrupa a las comunidades del valle, es decir, El Deca, El Buena y San Andrés Daboxthá. Las otras tres comunidades, todas ellas de la sierra, no parecen agruparse con una lógica administrativa ya que Santiago y Santa Mónica, pertenecientes a municipios diferentes, tienden a agruparse conjuntamente mientras de San Nicolás aparece más aislado.

Para comprobar si los parámetros administrativos y/o geográficos influyen en la conformación de la estructura genética de las comunidades Otomías, se han realizado dos AMOVAS atendiendo a estos dos criterios: geográfico y administrativo. En ambos casos (Tabla III.20) los resultados para las agrupaciones entre grupos, $F_{ct} = 0,01748$ con una $p = 0,1026$ para el criterio administrativo y $F_{ct} = 0,03296$ con una $p = 0,19159$ para el criterio geográfico, han resultado no significativas. Al menos para las comunidades Otomías no existe estructura genética relacionada con estos dos criterios.

Como último análisis descriptivo para los grupos Otomías se han analizado las comunidades atendiendo a las frecuencias de sus haplogrupos (Tabla III.21; Figura III.27).

Grupo		A2	B2	C1	D1	Total
El Buena	N	7	4	12	1	24
	%	29,16	16,66	50,00	4,16	100
El Deca	N	19	5	5	1	30
	%	63,33	16,66	16,66	3,33	100
San Andrés Daboxthá	N	14	3	5	5	27
	%	51,85	11,11	18,51	18,51	100
San Nicolás	N	15	1	15	1	32
	%	46,87	3,12	46,87	3,12	100
Santa Mónica	N	18	3	3	5	29
	%	62,06	10,34	10,34	17,24	100
Santiago	N	16	6	4	4	30
	%	53,33	20,00	13,33	13,33	100

Tabla III.21. Valores absolutos y porcentajes de los haplogrupos por comunidades Otomías

El Deca, San Andrés, Santa Mónica y Santiago presentan un modelo en el que el haplogrupo A2 es el más frecuente, seguido por el haplogrupo B2 o C1. La comunidad

de El Buena rompe este modelo, ya que presenta como haplogrupos más frecuente C1, seguido de A2, B2 y, por último, D1. La comunidad de San Nicolás también presenta un modelo diferente, en este caso los haplogrupos más frecuentes son el A2 y el C1 y los menos representados B2 y D1.

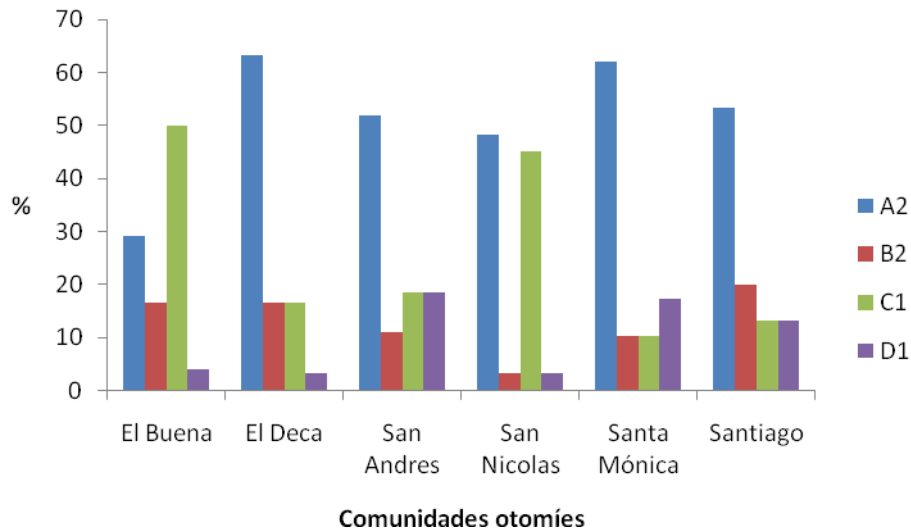


Figura III.27. Porcentaje de los haplogrupos amerindios en las comunidades Otomíes

III.6.II. Nahuas de la Huasteca

Debido a la compleja historia de la región de la Huasteca, el muestreo se diseñó para hacerse de forma exhaustiva, tratando de representar la máxima diversidad de la región. Por este motivo se trabajó en siete comunidades indígenas diferentes pertenecientes a cuatro municipios de la Huasteca hidalguense.

Siguiendo la misma metodología utilizada para los Otomíes, se realizarán análisis a nivel de municipios para luego profundizar a nivel de comunidades. En este caso, y dado que en algunos municipios el número de muestras es bastante elevado, los análisis se han realizado utilizando la región control completa, y HVI y HVII de forma independiente.

En la tabla III.22 se reflejan los índices de diversidad para los cuatro municipios muestreados divididos en D-Loop, HVI y HVII. La totalidad de los resultados va en mismo sentido, el municipio más diverso es Xochiatipan ($\pi = 0,01003 \pm 0,0004$; $H = 0,976 \pm 0,007$) seguido de Atlapexco ($\pi = 0,00980 \pm 0,0005$; $H = 0,955 \pm 0,017$), Yahualica ($\pi = 0,00972 \pm 0,0005$; $H = 0,973 \pm 0,009$) y Huejutla de Reyes ($\pi =$

0,00913 \pm 0,0011; H = 0,967 \pm 0,019). La misma información se obtiene al utilizar los valores de S, K o M.

Al representar gráficamente los índices de diversidad por municipios y región del ADNmt, se observa que la información aportada por cada una de las regiones genómicas es distinta (Figura III.28).

Municipio	Región	N	k	S	π	H	M	D
Atlápexco	D-Loop	42	25	52	0,00980 \pm 0,0005	0,955 \pm 0,017	10,85	-0,361; p > 0,10
	HVI	42	21	30	0,01845 \pm 0,0010	0,933 \pm 0,001	6,31	-0,322; p > 0,10
	HVII	42	10	9	0,00934 \pm 0,0006	0,841 \pm 0,037	2,40	0,427; p > 0,10
Huejutla de Reyes	D-Loop	25	17	49	0,00913 \pm 0,0011	0,967 \pm 0,019	10,04	-0,928; p > 0,10
	HVI	25	14	26	0,01554 \pm 0,0023	0,927 \pm 0,035	5,20	-0,902; p > 0,10
	HVII	25	8	10	0,00884 \pm 0,0013	0,747 \pm 0,080	2,33	-0,667; p > 0,10
Xochiatipan	D-Loop	68	41	59	0,01003 \pm 0,0004	0,976 \pm 0,007	11,18	-0,309; p > 0,10
	HVI	68	25	33	0,01936 \pm 0,0011	0,950 \pm 0,010	6,62	-0,123; p > 0,10
	HVII	68	12	10	0,00724 \pm 0,0005	0,777 \pm 0,037	1,90	-0,238; p > 0,10
Yahualica	D-Loop	56	32	60	0,00972 \pm 0,0005	0,973 \pm 0,009	10,76	-0,603; p > 0,10
	HVI	56	29	35	0,01886 \pm 0,0011	0,969 \pm 0,009	6,45	-1,162; p > 0,10
	HVII	56	13	12	0,00891 \pm 0,0006	0,809 \pm 0,039	2,29	-0,358; p > 0,10

Tabla III.22. Índices de diversidad para los cuatro municipios de la Huasteca muestreados. Los datos se dan para D-Loop, HVI y HVII; **k**: número de haplotipos diferentes; **S**: número de sitios polimórficos; **π** : diversidad nucleotídica \pm desviación estándar; **H**: diversidad haplotípica \pm desviación estándar; **M**: número medio de nucleótidos diferentes; **D**: índice de Tajima

Si se analiza la información extraída de la secuencia completa de D-Loop en el patrón municipal de los Nahuas de la Huasteca, el municipio más diverso es Xochiatipan y el menos diverso Huejutla de Reyes. Este mismo modelo se mantiene con la información aportada por HVI, pero para HVII se detectan algunas variaciones, ya que en este caso la población más diversa sería Atlápexco, seguida de Huejutla de Reyes, Yahualica y Xochiatipan. Respecto al test D de Tajima en todos los casos el resultado fue negativo, excepto para el municipio de Atlápexco cuando se utilizaba la región HVII, y no significativo.

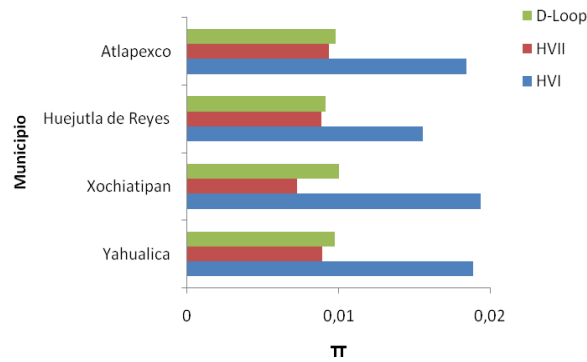


Figura III.28. Representación de los valores de π para los cuatro municipios de la Huasteca y para las regiones D-Loop, HVI y HVII.

Para profundizar más en la estructura genética de la población Nahua se ha calculado el test exacto de diferenciación poblacional, un estimador semejante al test exacto de Fisher (Raymond and Rousset, 1995) pero extendido a tablas de contingencia de cualquier tamaño. Los valores de probabilidad asociada a cada comparación se han estimado con 10.000 repeticiones de la cadena de Markov. Los resultados son concluyentes y todas las poblaciones presentan diferencias significativas excepto Atlapexco con Yahualica, de manera que las variantes haplotípicas acumuladas en la región D-Loop son lo suficientemente grandes como para determinar que son poblaciones diferentes (Tabla III.23).

	Atlapexco	Huejutla	Xochiatipan	Yahualica
Atlapexco		+	+	-
Huejutla	0,0000		+	+
Xochiatipan	0,0000	0,0000		+
Yahualica	0,0978	0,0000	0,0000	

Tabla III.23. Diferenciación entre parejas de municipios utilizando la región D-Loop completa. En la hemimatriz inferior se muestran los valores de significación (p) del test exacto de diferenciación poblacional, y en la superior el resultado del test; (+) diferencias significativas; (-) diferencias no significativas

Al utilizar la región HVI (Tabla III.24) y HVII (Tabla III.25) de forma independiente para determinar la diferenciación entre los municipios, los resultados son los mismos, aunque los valores de p oscilan ligeramente, y tienden a ser menores al utilizar una u otra región mitocondrial.

	Atlapexco	Huejutla	Xochiatipan	Yahualica
Atlapexco		+	+	-
Huejutla	0,00340		+	+
Xochiatipan	0,00105	0,00360		+
Yahualica	0,21450	0,01885	0,01135	

Tabla III.24. Diferenciación entre parejas de municipios utilizando la HVI. En la hemimatriz inferior se muestran los valores de significación (p) del test exacto de diferenciación poblacional, en la superior el resultado del test; (+) diferencias significativas; (-) diferencias no significativas

	Atlapexco	Huejutla	Xochiatipan	Yahualica
Atlapexco		+	+	-
Huejutla	0,00040		+	+
Xochiatipan	0,00070	0,00280		+
Yahualica	0,25160	0,02330	0,00750	

Tabla III.25. Diferenciación entre parejas de municipios utilizando la HVI. En la hemimatriz inferior se muestran los valores de significación (p) del test exacto de diferenciación poblacional, en la superior el resultado del test; (+) diferencias significativas; (-) diferencias no significativas

	Atlapexco	Huejutla	Xochiatipan	Yahualica
Atlapexco		0,01802*	0,15315	0,18018
Huejutla	0,05534		0,01802*	0,10811
Xochiatipan	0,00929	0,03418		
Yahualica	0,00698	0,01620	0,00461	

Tabla III.26. Distancias genéticas entre los cuatro municipios de la Huasteca estimados mediante Fst y utilizando la región D-Loop. En la hemimatriz inferior se muestran los valores de Fst y en la superior la significación (p). * valor significación al 0,05.

También se han calculado las distancias genéticas estimadas por Fst entre las cuatro poblaciones (Tabla III.26), y se han representado mediante un *multidimensional scaling* (MDS) (Figura III.29). Los valores de Fst son significativos entre los municipios de Atlapexco y Huejutla (Fst = 0,05534; p = 0,01802), y Huejutla y Xochiatipan (Fst = 0,03418; p = 0,01802). Estos datos inducen a pensar que el municipio de Huejutla presenta una composición genética diferente al resto de municipios Nahuas de la Huasteca. Este planteamiento se confirma al construir un MDS (Figura III.29), dónde se observa claramente como este municipio presenta una marcada diferencia respecto a los otros tres. De todas formas los valores de *stress* son muy elevados para los cuatro objetos representados (Sturrock and Rocha, 2000), por los que hay que analizar con sumo cuidado este resultado.

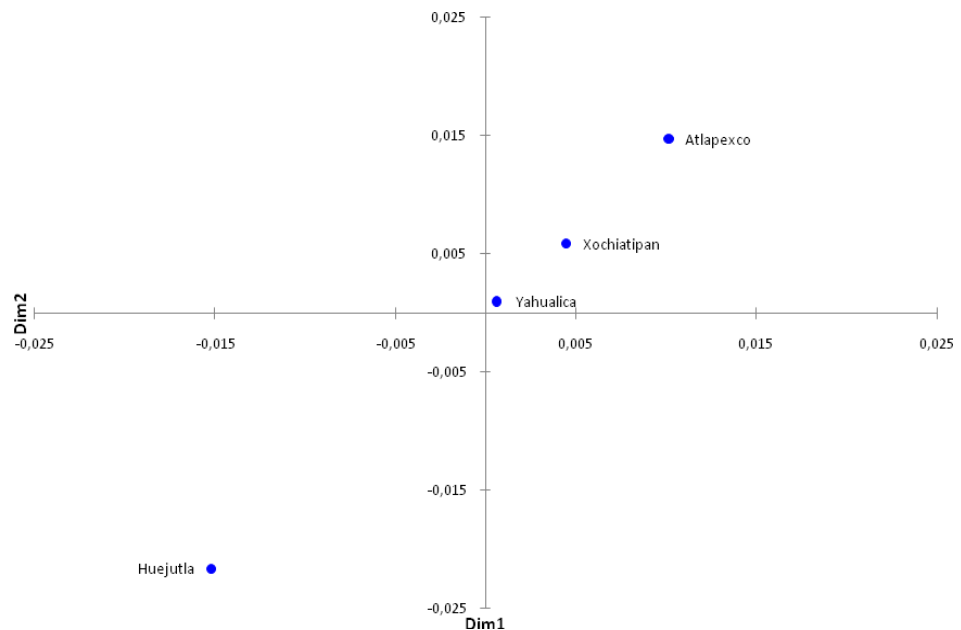


Figura III.29. MDS construido a partir de las F_{st} calculadas con información de las secuencias de D-Loop para los cuatro municipios del grupo indígena Nahuatl, ($stress = 0,286$).

En la tabla III.27 se reflejan los valores de los índices de diversidad para las siete comunidades muestreadas calculadas únicamente para D-Loop. Como ya se ha comentado en el apartado anterior, en este nivel de análisis se utilizará toda la información posible para tratar de comprender las relaciones entre las distintas comunidades. Los municipios de Atlapexco, Xochiatipan y Yahualica están representados por dos comunidades, mientras que Huejutla de Reyes lo está únicamente por Santa Catarina. Se puede apreciar que los valores de π son superiores en las comunidades indígenas de Texoloc e Ixtaczoquico, ambas pertenecientes al municipio de Xochiatipan. Con valores intermedios siguen en cuanto a diversidad las dos comunidades pertenecientes al municipio de Yahualica, es decir Tlalchihualica y Atlalco, y a continuación Tecacahuaco, perteneciente al municipio de Atlapexco. Las dos comunidades con menores índices de diversidad son, en orden decreciente, Santa Catarina, única comunidad representante del municipio de Huejutla de Reyes, y Pahactla del de Atlapexco.

La información que se obtiene del análisis de la diversidad, no aporta ningún conocimiento específico sobre la estructura genética del grupo Nahuatl, ni de las relaciones genéticas que existen entre municipios y comunidades. Por este motivo se han realizado dos niveles de análisis. Por una parte se han determinado las relaciones

genéticas siguiendo diferentes estrategias estadísticas, y por otro se ha realizado una AMOVA para determinar si la distribución de los haplotipos por comunidades y municipios en la Huasteca, es lo suficientemente importante como para determinar la existencia de poblaciones con subestructura genética.

Municipio	Comunidad	N	A2 (%)	B2 (%)	C1 (%)	D1 (%)
	Tecacahuaco	22	8 (36,36)	13 (59,91)	1 (4,54)	-
	Pahactla	11	8 (72,72)	1 (9,01)	1 (9,01)	1 (9,01)
Atlapexco	Total	33	16 (48,48)	14 (42,42)	2 (6,06)	1 (3,03)
	Santa Catarina	25	16 (64,00)	2 (8,00)	4 (16,00)	3 (12,00)
Huajuquila de Reyes	Total	25	16 (64,00)	2 (8,00)	4 (16,00)	3 (12,00)
	Ixtaczoquico	38	20 (52,63)	12 (31,58)	4 (10,52)	2 (5,26)
	Texoloc	30	19 (63,33)	8 (26,66)	3 (10,00)	-
Xochiatipan	Total	68	39 (57,35)	20 (51,28)	7 (10,29)	2 (2,94)
	Tlalchihualica	31	17 (54,83)	8 (25,80)	4 (12,90)	2 (6,45)
	Atlalco	25	17 (68,00)	7 (22,58)	1 (4,00)	-
Yahualica	Total	56	34 (60,71)	15 (26,78)	5 (8,93)	2 (3,57)

Tabla III.28. Valor absoluto y porcentajes de los haplogrupos del ADNmt para los municipios y las comunidades de la Huasteca.

Existen distintas metodologías para determinar las relaciones genéticas entre poblaciones. Para iniciar este proceso, el primer paso ha consistido en la descripción de los haplogrupos según municipio y comunidad (Tabla III.28 y Figuras III.30 y III.31). Destaca el municipio de Huejutla de Reyes, que presenta un valor bajo de haplogrupo B2 y valores de C1 y D1 elevados y mayores que el resto de municipios. De cualquier manera, los resultados son bastante heterogéneos, lo que podría ser el resultado del

Municipio	Comunidad	N	k	S	π	H	M	D
Atlapexco	Tecacahuaco	22	12	36	0,00969 \pm 0,0007	0,905 \pm 0,039	10,76	0,3501; p > 0,10
	Pahactla	20	15	40	0,00850 \pm 0,0013	0,958 \pm 0,001	9,47	-0,6364; p > 0,10
Huejutla de Reyes	Santa Catarina	25	18	50	0,00918 \pm 0,0010	0,969 \pm 0,018	10,11	-0,9314; p > 0,10
Xochiatipan	Ixtaczoquico	37	24	44	0,01003 \pm 0,0004	0,967 \pm 0,014	11,20	0,2507; p > 0,10
	Texoloc	30	23	51	0,01004 \pm 0,0009	0,979 \pm 0,000	11,20	-0,4850; p > 0,10
Yahualica	Tlalchihualica	31	24	53	0,00978 \pm 0,0006	0,981 \pm 0,014	10,86	-0,6718; p > 0,10
	Atlalco	25	13	38	0,00937 \pm 0,0011	0,927 \pm 0,030	10,43	0,1394; p > 0,10

Tabla III.27 Índices de diversidad para las siete comunidades indígenas correspondientes a los cuatro municipios estudiados en la Huasteca. Los datos se dan para D-Loop, **k**: número de haplotipos diferentes; **S**: número de sitios polimórficos; **π** : diversidad nucleotídica \pm desviación estándar; **H**: diversidad haplotípica \pm desviación estándar; **M**: número medio de nucleótidos diferentes; **D**: índice de Tajima

tamaño muestral relacionado con la fragmentación del grupo indígena Nahua a nivel de comunidad.

De todos modos, y asumiendo que algunos resultados podrían estar reflejando este aspecto, hay algunos detalles que caracterizan claramente a algunas comunidades. Por ejemplo, aunque la falta de haplogrupo D1 en las comunidades de Tecacahuaco, Texoloc y Atlalco podría ser un reflejo del tamaño muestral, el alto porcentaje del B2 en la población de Tecacahuaco parece ser indicativo de una composición diferencial respecto a otras comunidades, al menos en lo que se refiere a la distribución de los haplogrupos.

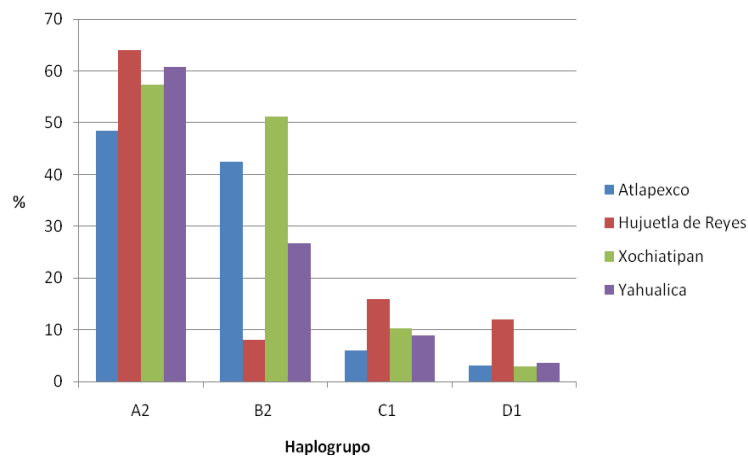


Figura III.30. Distribución en porcentaje de los haplogrupos por municipios en la Huasteca

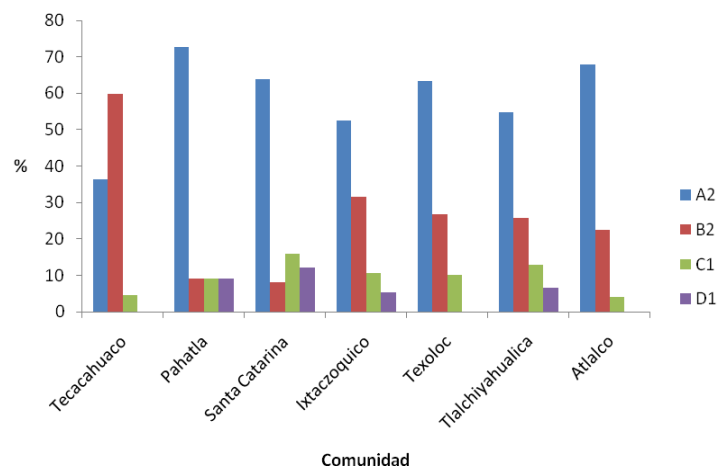


Figura III.31. Distribución en porcentaje de los haplogrupos amerindios en las siete comunidades indígenas de la Huasteca hidalguense

Para facilitar la interpretación de los datos se ha realizado un análisis de componentes principales (ACP) (Figura III.32). Los dos ejes principales explican el 66,45% de la

variabilidad confirmando que la topografía de las comunidades es muy sólida. En el ACP también se han representado los haplogrupos, lo que permite determinar la importancia que cada uno de ellos tiene sobre la distribución de las comunidades.

Texoloc e Ixtaczoquico están claramente influenciadas por el haplogrupo A2, mientras que en Santa Catarina tiene un mayor peso relativo el haplogrupo D1. Por otra parte, Tlalchihualica está muy relacionado con el haplogrupo C1 y el resto de comunidades con el haplogrupo B2. Esta distribución permite definir cuatro *clusters* o agrupaciones representados en color verde: Tecoloc, Ixtaczoquico y Tlalchihualica por una parte, Atlalco y Tecacahuaco por otra, y Santa Catarina y Pahactla representarían las otras dos agrupaciones independientes. No parece detectarse una relación entre comunidades asociada a la pertenencia al mismo municipio, englobadas en color azul. La observación de la figura muestra claramente la falta de concordancia entre las agrupaciones resultantes en el ACP y las asociaciones administrativas entre las comunidades.

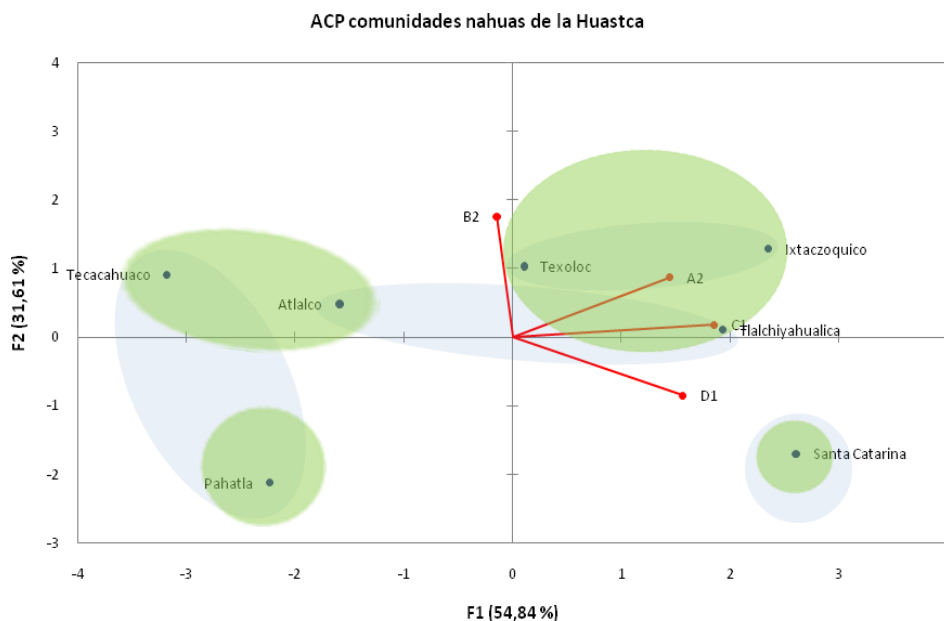


Figura III.32. Análisis de componentes principales de las siete comunidades indígenas Nahuas de la Huasteca. Se relacionan en color verde los *clusters* del ACP y en color azul las agrupaciones de las comunidades por municipios.

Para confirmar estos resultados se ha calculado el test exacto de diferenciación poblacional para las siete comunidades, cuyos resultados se exponen en la tabla III.29. Los resultados son coherentes con los observados en el ACP. Las dos únicas

comunidades pertenecientes al mismo municipio que presentan una coherencia con los *clusters* en el ACP, también tienen índices de diferenciación poblacional no significativos, como es el caso de Texoloc e Ixtaczoquico del municipio de Xochiatipan, que con un valor de 0.06090 no presenta significación en sus diferencias. El resto de comunidades pertenecientes a un mismo municipio sí que presentan diferencias significativas entre ellas. Es el caso de Tecacahuaco y Pahactla, del municipio de Atlapexco, que presentan un valor significativo del índice de diferenciación poblacional ($p = 0.00035$), o el de Atlalco y Tlalchiyalualica, las dos comunidades del municipio de Yahualica, que tienen también valores significativos ($p = 0,00310$). Estos resultados implican dos deducciones. Por una parte, es el reflejo de que en algunos casos entre comunidades se han acumulado diferencias lo suficientemente importantes como para que sean poblaciones diferentes (aunque habría que tener en cuenta el posible efecto del tamaño poblacional). Y por otra, parece no existir una relación de diferenciación entre comunidades del mismo municipio, por lo que es fácilmente deducible que la coherencia administrativa no parece relacionarse con la genética.

Las distancias F_{st} y sus probabilidades se muestran en la tabla III.30. Los resultados de esta tabla muestran que las distancias genéticas son significativas en nueve casos, en seis de los cuales está implicada la población de Tecacahuaco. Esta población presenta distancias genéticas distintas al compararla con todo el resto de comunidades. La comunidad de Atlalco presenta diferencias con Tlalchiyajualica e Ixtaczoquico, y esta última también con Santa Catarina.

De cualquier manera, la mejor forma de aproximarse a la posible estructura genética de las comunidades atendiendo a su pertenencia a un municipio es mediante una AMOVA (Tabla III.31). Los resultados de la AMOVA demuestran que no existe una relación entre la clasificación administrativa y composición genética en el caso concreto de estos municipios. La diversidad explicada entre grupos es baja (0,45), siendo $F_{ct} = 0.00450$ con una probabilidad asociada de 0,301, es decir, no significativa. Además se han realizado otras dos AMOVA, la primera de ellas con el criterio "Levante-Poniente" de las vertientes de la Sierra Otomí-tepehua que explica poca diversidad (0,78) aunque más elevada que en el caso anterior, $F_{ct} = 0,00778$ y tampoco es significativa ($p = 0,102$). La última AMOVA realizada es la llamada "Geográfica", en que se han agrupado las comunidades alrededor de un eje imaginario norte-sur y aunque la diversidad explicada es la mayor de las tres agrupaciones (1,60), sigue siendo insuficiente ($F_{ct} = 0,01604$) y no significativa ($p = 0,336$).

Por último, y para concluir los estudios intrapoblacionales que se han realizado en la Huasteca, se han calculado los haplotipos únicos y los compartidos para las siete comunidades de la región atendiendo a los haplotipos de la región D-Loop completa (Tabla III.32 y Figura III.33). Se han detectado 93 haplotipos diferentes de un total de 191 secuencias, lo que representan un 48,69%. Obviamente los valores de los haplotipos compartidos son complementarios a los haplotipos únicos. En la tabla III.33 se da el número de haplotipos compartidos entre comunidades.

Comunidad	n	Únicos	%	Compartidos	%
Tecacahuaco	12	8	66,66	4	33,33
Pahactla	15	6	40,00	9	60,00
Santa Catarina	18	15	83,33	3	16,66
Ixtaczoquico	24	14	58,33	10	41,66
Texoloc	23	11	47,82	12	52,18
Atlalco	13	7	53,84	6	46,16
Tlalchihualica	24	12	50,00	12	50,00

Tabla III.32. Número de haplotipos únicos y compartidos por comunidades de la Huasteca para la región D-Loop. El tanto por ciento se da respecto al total de haplotipos distintos que tiene cada una de las comunidades (n)

	Tecacahuaco	Pahactla	Santa Catarina	Ixtaczoquico	Texoloc	Tlalchihualica	Atlalco
Tecacahuaco	-						
Pahactla	2	-					
Santa Catarina	0	1	-				
Ixtaczoquico	1	2	1	-			
Texoloc	0	4	3	6	-		
Tlalchihualica	4	7	1	4		-	
Atlalco	2	4	1	0	3	5	-

Tabla III.33. Número de haplotipos compartidos por comunidades de la Huasteca para la región D-Loop

		ATLAPEXCO		HUEJUTLA	XOCHIATIPAN		YAHUALICA	
		Tecacahuaco	Pahactla	Santa Catarina	Ixtaczoquico	Texoloc	Atlalco	Tlalchihualica
ATLAPEXCO	Tecacahuaco		+	+	+	+	+	+
	Pahactla	0,00035		+	+	-	-	-
HUEJUTLA	Santa Catarina	0,00000	0,01105		+	+	+	+
XOCHIATIPAN	Ixtaczoquico	0,00000	0,00000	0,00000		-	+	+
	Texoloc	0,00000	0,37335	0,00920	0,06090		+	+
YAHUALICA	Atlalco	0,00000	0,07930	0,00000	0,00000	0,00000		+
	Tlalchihualica	0,01125	0,44555	0,00000	0,00065	0,01085	0,00310	

Tabla III.29. Diferenciación entre parejas de comunidades utilizando D-Loop. En la hemimatriz inferior se muestran los valores de significación (p) del test exacto de diferenciación poblacional, en la superior el resultado del test; (+) diferencias significativas; (-) diferencias no significativas

		ATLAPEXCO		HUEJUTLA	XOCHIATIPAN		YAHUALICA	
		Tecacahuaco	Pahactla	Santa Catarina	Ixtaczoquico	Texoloc	Atlalco	Tlalchihualica
ATLAPEXCO	Tecacahuaco		0,04505	0,00901	0,03604	0,01802	0,03604	0,02703
	Pahactla	0,10052		0,45946	0,34234	0,48649	0,25225	0,64865
HUEJUTLA	Santa Catarina	0,13178	-0,00475		0,01802	0,10811	0,06306	0,15315
XOCHIATIPAN	Ixtaczoquico	0,05551	0,00307	0,03641		0,42342	0,04505	0,74775
	Texoloc	0,07592	-0,01010	0,01877	-0,00137		0,18919	0,45946
YAHUALICA	Atlalco	0,11403	-0,00108	0,03630	0,05050	0,01687		0,01802
	Tlalchihualica	0,04707	-0,01242	0,01088	-0,01253	-0,00594	0,03055	

Tabla III.30. Distancias genéticas Fst entre parejas de comunidades de la Huasteca utilizando la región D-Loop. En la hemimatriz inferior se muestran los valores de Fst y en la superior los valores de significación (p)

Criterio	Agrupación	Grupo	Fuente de variación	Porcentaje varianza	Índice fijación	P
Administrativo	Atlapexco	Tecacahuaco, Pahactla	Dentro de poblaciones	96,90	Fst = 0,03104	0.000*
	Huejutla	Santa Catarina	Entre poblaciones dentro de grupos	2,65	Fsc = 0.02667	0.000*
	Xochiatipan	Ixtaczoquico, Texoloc	Entre grupos	0,45	Fct = 0.00450	0.301
	Yahualica	Tlalchiyalualica, Atlalco				
Levante-Poniente	Levante	Ixtaczoquico, Texoloc	Dentro de poblaciones	96,61	Fst = 0,03388	0.000*
	Poniente	Tecacahuaco, Pahactla, Santa Catarina, Tlalchiyalualica, Atlalco	Entre poblaciones dentro de grupos	2,61	Fsc = 0.02630	0.000*
			Entre grupos	0,78	Fct = 0.00778	0.102
Geografía	Norte	Santa Catarina	Dentro de poblaciones	97,66	Fst = 0,02339	0.000*
	Centro	Tecacahuaco, Pahactla, Tlalchiyalualica, Atlalco	Entre poblaciones dentro de grupos	0,73	Fsc = 0.00747	0,134
	Sur	Ixtaczoquico, Texoloc	Entre grupos	1,60	Fct = 0.01604	0,336

Tabla III.31. Resultados de las AMOVA realizadas para distintas agrupaciones. En la primera de ellas las agrupaciones son de tipo administrativo por municipio, en la segunda por una ubicación levante poniente en las vertientes de la Sierra Otomí-Tepehua, y la última por un orden geográfico en un eje imaginario norte-sur. Valor de significación (p)

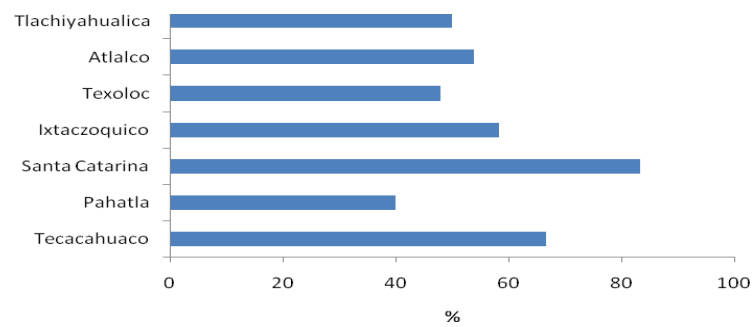


Figura III.33. Representación gráfica de los porcentajes de haplotipos únicos para cada una de las comunidades de la Huasteca

Tanto en la tabla III.32 como en la figura III.33 se puede comprobar que la comunidad con mayor número de haplotipos únicos es Santa Catarina (83,33%), seguida de Tecacahuaco (66,66%), Ixtaczoquico (58,33%) y Atlalco (53,84%). La comunidad con menor número de haplotipos únicos es Pahactla (40%) y Texoloc (47,87%). Respecto a los haplotipos compartidos (Tabla III.33) Tlachiyahualica y Pahactla comparten siete y Texoloc e Ixtaczoquico seis. Para el resto de poblaciones el número es variable y en tres ocasiones el valor es cero (Tecacahuaco con Santa Catarina y Texoloc, y Atlalco con Ixtaczoquico).

III.6.III. El Estado de Hidalgo

Aprovechando que gran parte de la información procede del estado de Hidalgo y que es la región donde se han podido realizar más detalladamente los muestreos, se ha creado este apartado en el que se analizarán las relaciones a nivel de comunidades entre Otomíes, Nahuas de la Huasteca y Tepehuas utilizando la región D-Loop. Hay que advertir que el estado de Hidalgo no tiene por qué tener una identidad poblacional ya que se trata de una definición político-administrativa. De hecho la referencia utilizada como Estado de Hidalgo es para situarla en el contexto de la República Mexicana. Lo realmente interesante es que esta clasificación administrativa contiene, por una parte, al menos tres grupos indígenas distintos con lenguas bien diferenciadas y con historias demográficas diferentes. Por otro lado esta región contiene una excepcional orografía, diferenciándose claramente una región central semidesértica, o valle de Ixmiquilpan donde se localizan los Otomíes del Valle, y un eje neovolcánico situado hacia el levante del estado. Esta región montañosa representa una frontera natural entre el altiplano mexicano y su conexión con las costas del Atlántico. Estos

dos aspectos, cultural y geográfico, serán los puntos de enfoque del apartado, utilizando la referencia al estado de Hidalgo para facilitar su localización espacial.

	N	k	S	π	H	M	D
D-Loop	416	181	148	0,00995 \pm 0,00020	0,9889 \pm 0013	10,857	-1,57673; p > 0,05

Tabla III.34. Índices de diversidad para los individuos del estado de Hidalgo. Los datos se dan para D-Loop, **k**: número de haplotipos diferentes; **S**: número de sitios polimórficos; **π** : diversidad nucleotídica \pm desviación estándar; **H**: diversidad haplotípica \pm desviación estándar; **M**: número medio de nucleótidos diferentes; **D**: índice de Tajima

En la tabla III.35 se muestran los índices de diversidad para los grupos indígenas del estado de Hidalgo, es decir, para los grupos indígenas Otomíes del Valle, Otomíes de la Sierra, Nahuas de la Huasteca y Tepehuas. Los índices de diversidad son relativamente elevados con $\pi = 0,00995 \pm 0,00020$ y $H = 0,9889 \pm 0013$, y una media de diferencias para la región D-Loop de 10,857. El test de neutralidad de Tajima es negativo y no significativo.

El modelo de *pairwise differences* es propio de poblaciones en expansión, con una distribución bi-modal, alcanzando la segunda moda valores más elevados (Figura III.34).

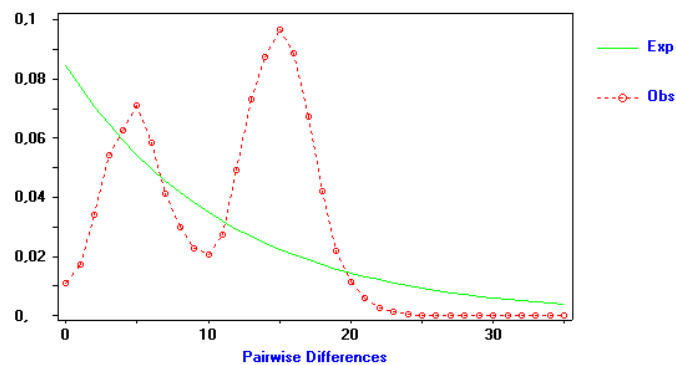


Figura III.34. *Pairwise differences* para las comunidades del estado de Hidalgo utilizando la región D-Loop completa.

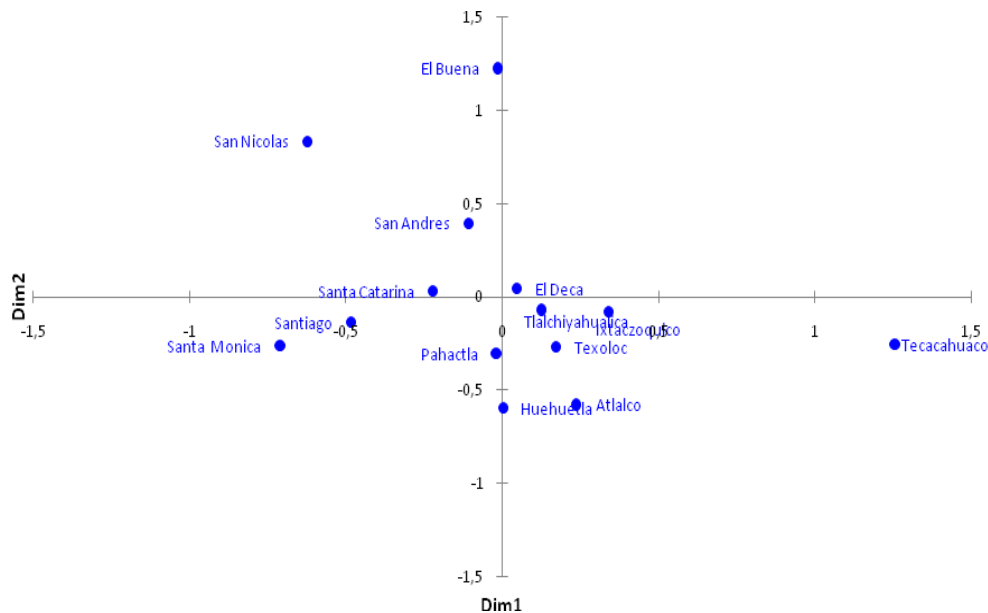


Figura III.35. MDS que representa las catorce comunidades del estado de Hidalgo. EL MDS se ha construido utilizando la región D-Loop y la distancia genética Fst (*stress*: 0,144)

En la tabla III.35 se dan los valores de Fst así como sus significaciones. En este caso no parece existir un modelo claro. Las diferencias, significativas o no significativas, no se corresponden necesariamente a las poblaciones del mismo municipio o del mismo grupo indígena indicando una falta de lógica administrativa. De todas formas las mayores distancias se detectan entre grupos indígenas diferentes o entre regiones geográficas distintas. Para analizar este resultado se pueden observar los valores de Fst para Atlalco. Esta comunidad, perteneciente al municipio de Yahualica situado en la región de la sierra y englobada dentro del grupo indígena Nahuatl, presenta un valor de $F_{st} = 0,16743$ ($p = 0,0000$) al compararlo con El Buena, comunidad del municipio de El Cardonal de cultura Otomí y situado en la región semidesértica de Ixmiquilpan.

Los valores obtenidos al calcular los índices exactos de fijación poblacional entre todos los pares de comparaciones posibles entre todas las comunidades del estado, pueden aportar una información complementaria a las distancias genéticas ya comentadas (Tabla III.36). En este caso, la gran mayoría de poblaciones presenta diferencias

	El Buena	El Deca	San Andrés	San Nicolás	Santa Mónica	Santiago	Atlalco	Ixtaczoquico	Pahactla	Santa Catarina	Tecacahuaco	Texoloc	Tlalchihualica	Huehuetla
El Buena	0,00000	0.00000	0.03604	0.09910	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
El Deca	0,09060	0,00000	0.25225	0.00000	0.02703	0.03604	0.10811	0.24324	0.47748	0.32432	0.00000	0.52252	0.51351	0.15315
San Andrés	0,04314	0,00565	0,00000	0.03604	0.01802	0.03604	0.01802	0.03604	0.10811	0.27928	0.00000	0.04505	0.11712	0.00901
San Nicolás	0,03497	0,08691	0,04664	0,00000	0.00901	0.00901	0.00901	0.00000	0.00000	0.01802	0.00000	0.00000	0.00901	0.00000
Santa Mónica	0,15535	0,06834	0,07231	0,06755	0,00000	0.45946	0.00901	0.00000	0.01802	0.05405	0.00000	0.00901	0.04505	0.00000
Santiago	0,12659	0,04511	0,04746	0,06039	0,00000	0,00000	0.01802	0.04505	0.06306	0.04505	0.00000	0.06306	0.13514	0.00000
Atlalco	0,16743	0,02697	0,06807	0,13875	0,07263	0,07100	0,00000	0.08108	0.42342	0.08108	0.00901	0.21622	0.13514	0.06306
Ixtaczoquico	0,10679	0,01041	0,03988	0,11055	0,06570	0,03701	0,03947	0,00000	0.32432	0.03604	0.00000	0.42342	0.62162	0.06306
Pahactla	0,14385	0,00000	0,02976	0,11410	0,04503	0,03176	0,00000	0,00484	0,00000	0.51351	0.01802	0.60360	0.54054	0.27027
Santa Catarina	0,10617	0,00360	0,00678	0,06918	0,04358	0,02800	0,03196	0,04167	0,00000	0,00000	0.00000	0.13514	0.24324	0.02703
Tecacahuaco	0,17933	0,12879	0,15894	0,18928	0,14055	0,10156	0,11469	0,06512	0,11347	0,15028	0,00000	0.00901	0.04505	0.00000
Texoloc	0,12604	0,00000	0,03965	0,11462	0,05416	0,03892	0,00844	0,00000	0,00000	0,02191	0,08688	0,00000	0.59459	0.32432
Tlalchihualica	0,08940	0,00000	0,02133	0,07979	0,04369	0,01960	0,01823	0,00000	0,00000	0,01154	0,05628	0,00000	0,00000	0.22523
Huehuetla	0,16476	0,01372	0,05940	0,15272	0,08201	0,05457	0,03696	0,01947	0,00657	0,03827	0,11866	0,00549	0,01177	0,00000

Tabla III.35. Resultados Fst para la comparación par a par entre las catorce comunidades estudiadas en el Estado de Hidalgo. En la hemimatriz inferior se representa en valor de Fst, en la superior el valor de p. En negrita los resultados significativos

III. Resultados

	El Buena	El Deca	San Andrés	San Nicolás	Santa Mónica	Santiago	Atlalco	Ixtaczoquico	Pahactla	Santa Catarina	Tecacahuaco	Texoloc	Tlalchihualica	Huehuetla
El Buena		-	-	+	+	+	+	+	+	+	+	+	+	+
El Deca	0,26210		-	+	+	+	+	+	+	+	+	+	+	+
San Andrés	0,16600	0,06470		+	+	+	+	+	+	+	+	+	+	+
San Nicolás	0,00375	0,00000	0,00010		+	+	+	+	+	+	+	+	+	+
Santa Mónica	0,00115	0,00000	0,00000	0,00020		+	+	+	+	+	+	+	+	+
Santiago	0,02645	0,00000	0,00000	0,00350	0,01625		+	+	+	+	+	+	+	+
Atlalco	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000		+	-	+	+	+	+	+
Ixtaczoquico	0,00075	0,00055	0,00000	0,00000	0,00000	0,00000	0,00000		+	+	+	+	+	+
Pahactla	0,00410	0,00305	0,00000	0,00000	0,00000	0,00210	0,07125	0,00000		+	+	-	-	+
Santa Catarina	0,00675	0,00095	0,00000	0,00000	0,00000	0,01810	0,00000	0,00000	0,00680		+	+	+	+
Tecacahuaco	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00235	0,00000	0,00000	0,00000		+	+	+
Texoloc	0,00520	0,00015	0,00000	0,00000	0,00000	0,00000	0,00360	0,03980	0,35495	0,01485	0,00000		+	+
Tlalchihualica	0,00475	0,02065	0,00170	0,00000	0,00000	0,00860	0,00100	0,00230	0,42470	0,00085	0,02150	0,00945		+
Huehuetla	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	

Tabla III.36. Resultados del test exacto de diferenciación poblacional para las catorce comunidades estudiadas en el Estado de Hidalgo. En la hemimatriz inferior se representa en valor de p y la superior si el resultados es significativo (+) o no (-)

Criterio	Agrupación	Grupo	Fuente de variación	Porcentaje varianza	Índice fijación	P
Grupo	Otomíes Valle	El Buena, El Deca, San Andrés Dabothá	Dentro de poblaciones	93.23	Fst = 0.06773	0.0000*
	Otomíes Sierra	San Nicolás, Santa Mónica, Santiago	Entre poblaciones dentro grupos	3.76	Fsc =0.03872	0.0000*
	Nahuas Huasteca	Atlalco, Ixtaczoquico, Pahactla, Santa Catarina, Tecacahuaco, Texoloc, Tlalchihualica	Entre grupos	3.02	Fct = 0.03017	0.0068*
	Tepehuas	Huehuetla				
Valle-Sierra	Valle	El Buena, El Deca, San Andrés Dabothá	Dentro de poblaciones	92.75	Fst = 0.07247	0.0000*
	Sierra	San Nicolás, Santa Mónica, Santiago, Atlalco, Ixtaczoquico, Pahactla, Santa Catarina, Tecacahuaco, Texoloc, Tlalchihualica, Huehuetla	Entre poblaciones dentro grupos	5.31	Fsc =0.0541	0.0000*
			Entre grupos	1.93	Fct =0.01934	0.0508
Vertientes	Otomíes Valle	El Buena, El Deca, San Andrés Dabothá	Dentro de poblaciones	93.41	Fst = 0.06593	0.0000*
	Otomíes Sierra	San Nicolás, Santa Mónica, Santiago	Entre poblaciones dentro grupos	3.63	Fsc =0.03744	0.0000*
	Vertiente Valle	Atlalco, Pahactla, Santa Catarina, Tecacahuaco, Tlalchihualica	Entre grupos	2.96	Fct = 0.02960	0.0068*
	Vertiente Costa	Ixtaczoquico, Texoloc, Huehuetla				

Tabla III.37. AMOVAS calculadas para las catorce comunidades procedentes del estado de Hidalgo según varios criterios de clasificación. Los resultados significativos están indicados por un * ($p < 0,05$)

significativas, y tan solo en seis ocasiones estas no son significativas. Estos seis casos se corresponden a las comparaciones entre San Andrés Dabothá con El Buena ($p = 0,16600$) y el Deca ($p = 0,06470$), y entre estas dos últimas ($p = 0,26210$). Por otra parte Pahactla no presenta diferencias con Atlalco ($p = 0,07125$), Texoloc ($p = 0,35495$) y Tlalchihualica ($p = 0,42470$). Estos resultados son lógicos para las poblaciones Otomíes, ya que las tres (El Buena, El Deca y San Andrés Dabothá) son del mismo municipio, ahora bien, parecen no tener ninguna lógica con las procedentes de la Huasteca, donde cada una de ellas es de un municipio diferente.

Para facilitar la interpretación, los resultados de Fst se han representado en un MDS (Figura III.35). En la gráfica puede observarse una distribución superficial de las comunidades, indicativa de sus distancias genéticas. Aunque no existe una clara estructura sí que pueden diferenciarse las poblaciones Otomíes respecto a las Nahuas. Las primeras tienden a ocupar el cuadrante negativo del eje de las X, mientras que las Nahuas se distribuyen preferentemente en el cuadrante positivo de este mismo eje. Además, los índices de *stress* son aceptables, de hecho el valor límite para una representación de 14 objetos es 0,217, mucho mayor que el detectado en la gráfica que es de 0,144 (Sturrock and Rocha, 2000).

Para determinar la existencia de estructura genética en la región de Hidalgo, se han realizado tres AMOVAS atendiendo a diferentes criterios, uno cultural y dos geográficos (Tabla III.37). Al realizar una agrupación cultural, es decir, por grupos indígenas, la variabilidad entre grupos representa el 3,02, con $F_{ct} = 0.03017$ y significativa ($p = 0,0068$). Los resultados adquieren todavía mayor relevancia al repetir otras AMOVAS según criterios geográficos. Al clasificar a las poblaciones por grandes regiones orográficas, como valle y sierra, los resultados son no significativos ($F_{ct} = 0.01934$; $p = 0,0508$), aunque hay que matizar que la probabilidad está en el límite de la significación. Pero la geografía puede ser valorada de una forma más sutil, considerando por ejemplo que las vertientes de los grandes ejes montañosos tienen una particular relevancia. En este caso y como ya se ha comentado, la sierra Madre Oriental representa un límite geográfico pero también económico y cultural, el levante se abre hacia el mar, hacia la costa de Veracruz, mientras que el poniente se distribuye por infinidad de complejos valles que convergen en el valle del Mezquital. Al realizar la agrupación atendiendo a criterios geográficos pero matizando si las comunidades se disponen hacia el este o el oeste, los resultados son significativos ($F_{ct} = 0.02960$; $p = 0,0068$).

III.7. Relaciones interpoblacionales

Hasta ahora se han realizado análisis de la composición y diversidad de las poblaciones integradas en la tesis, prestando mayor atención a aquellos grupos en los que han podido abordarse aspectos de diversidad intrapoblacional, como Otomíes y Nahuas de la Huasteca, y diseccionando de forma especial las comunidades distribuidas en el estado de Hidalgo. En este apartado se analizarán las relaciones genéticas entre todas las poblaciones incluidas en la tesis, basadas en el análisis de la región D-Loop, HVI y HVII.

El resultado del test de diferenciación poblacional entre las ocho poblaciones (incluyendo los dos grupos Otomíes) mediante la región D-Loop, se muestra en la tabla III.38. Como puede apreciarse en todos los casos el índice de diferenciación es significativo, indicando que las diferencias genéticas entre las poblaciones son elevadas.

En la tabla III.39 se muestran los resultados al calcular F_{st} entre pares de poblaciones. Estos resultados se han utilizado para la creación de un MDS representado en la figura III.36.

	Pimas	Mayos	Huicholes	Nahuas	Otomíes Valle	Otomíes Sierra	Tepehuas	Mayas
Pimas		+	+	+	+	+	+	+
Mayos	0,000		+	+	+	+	+	+
Huicholes	0,000	0,000		+	+	+	+	+
Nahuas	0,000	0,000	0,000		+	+	+	+
Otomíes Valle	0,000	0,000	0,000	0,000		+	+	+
Otomíes Sierra	0,000	0,000	0,000	0,000	0,000		+	+
Tepehuas	0,000	0,000	0,000	0,000	0,000	0,000		+
Mayas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	

Tabla III.38. Resultados de la aplicación del test de diferenciación poblacional a los ocho grupos indígenas utilizando la región D-Loop. En la hemimatriz inferior se muestran los valores de significación (p), y en la superior el resultado del test; (+) diferencias significativas; (-) diferencias no significativas

La distancia genética, estimada por F_{st} , es significativa en la mayoría de los casos aunque se detectan dos excepciones: al comparar grupos Nahuas con Huicholes ($F_{st} = 0,01133$; $p = 0,14414$), y Huicholes con Otomíes del Valle ($F_{st} = 0,01096$; $p = 0,09009$). Los mayores valores de distancia genética son entre poblaciones Pimas y el resto, dato indicativo de la relación entre esta población del noroeste y el resto de grupos indígenas.

	Pimas	Mayos	Huicholes	Nahuas	Otomías Valle	Otomías Sierra	Tepehuas	Mayas
Pimas	-	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Mayos	0,29391	-	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Huicholes	0,36025	0,08746	-	0,14414	0,09009	0,00000	0,00000	0,00000
Nahuas	0,34393	0,07564	0,01133	-	0,00000	0,00000	0,04505	0,00000
Otomías Valle	0,27009	0,08050	0,01096	0,03573	-	0,00000	0,00000	0,00000
Otomías Sierra	0,29839	0,08545	0,03409	0,04745	0,03132	-	0,00000	0,01802
Tepehuas	0,41952	0,10110	0,03292	0,01387	0,05478	0,06431	-	0,01802
Mayas	0,46450	0,16094	0,10093	0,06986	0,10996	0,06120	0,08363	-

Tabla III.39. Resultados de las distancias genéticas F_{st} aplicadas a los grupos ocho grupos indígenas utilizando la región D-Loop. En la hemimatriz inferior se muestran los valores de F_{st} , en la superior el valor de p

Los resultados de las distancias genéticas se han representado en un MDS (Figura III.36) con un nivel de *stress* de 0,097, más elevado de lo permisivo para una representación de ocho elementos (Sturrock and Rocha, 2000). La distribución de los grupos indígenas muestra claramente una agrupación conformada por cinco poblaciones indígenas, Nahuas de la Huasteca, Huicholes, Tepehuas, Otomías del Valle y Otomías de la Sierra. Cercana a este gran *cluster* y en el mismo cuadrante, está la población Maya. Relativamente más alejados están los Mayos, y con una posición realmente llamativa dada su lejanía al resto de poblaciones, se encuentran los Pimas.

Este mismo test exacto de diferenciación poblacional, F_{st} y su representación en un MDS, se ha realizado utilizando de forma independiente las regiones HVI y HVII. Los resultados, son muy semejantes a los obtenidos al utilizar la región D-Loop completa. Respecto a HVI el test exacto de diferenciación poblacional muestra los mismos resultados (Tabla III.40).

Los valores de F_{st} , aunque en la misma dirección que los obtenidos para D-Loop, presenta algunas diferencias interesantes (Tabla III.41). Por ejemplo, las distancias no son significativas entre Nahuas y Huicholes ($F_{st} = 0,00478$; $p = 0,23423$), y entre Nahuas y Tepehuas ($F_{st} = 0,01374$; $p = 0,10811$). Otra diferencia interesante es la obtenida al comparar Huicholes con Otomías del Valle ($F_{st} = 0,00399$; $p = 0,27928$), y de la Sierra ($F_{st} = 0,00763$; $p = 0,13514$), que tampoco da diferencias significativas.

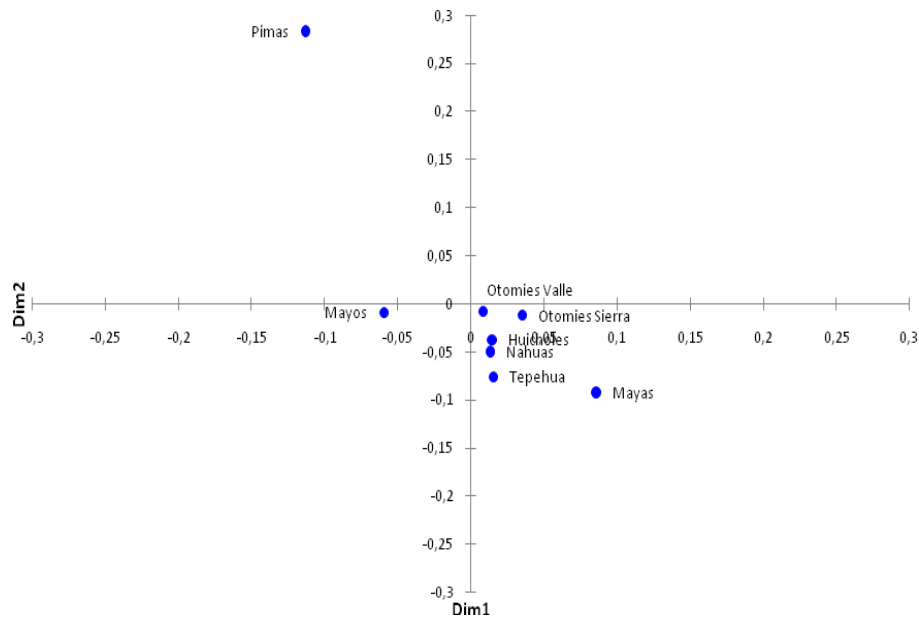


Figura III.36. MDS construido a partir de la distancia genética F_{st} y considerando la región D-Loop para las ocho poblaciones estudiadas ($stress = 0,097$)

	Pimas	Mayos	Huicholes	Nahuas	Otomies Valle	Otomies Sierra	Tepehuas	Mayas
Pimas		+	+	+	+	+	+	+
Mayos	0,000		+	+	+	+	+	+
Huicholes	0,000	0,000		+	+	+	+	+
Nahuas	0,000	0,000	0,000		+	+	+	+
Otomies Valle	0,000	0,000	0,000	0,000		+	+	+
Otomies Sierra	0,000	0,000	0,000	0,000	0,000		+	+
Tepehuas	0,000	0,000	0,000	0,000	0,000	0,000		+
Mayas	0,000	0,000	0,000	0,000	0,000	0,000	0,000	

Tabla III.40. Resultados de la aplicación del test de diferenciación poblacional a los ocho grupos indígenas utilizando la región HVI. En la hemimatriz inferior se muestran los valores de significación (p), y en la superior el resultado del test; (+) diferencias significativas; (-) diferencias no significativas

	Pimas	Mayos	Huicholes	Nahuas	Otomías Valle	Otomías Sierra	Tepehuas	Mayas
Pimas	-	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000	0,00000
Mayos	0,29131	-	0,00901	0,00000	0,00000	0,00000	0,00000	0,00000
Huicholes	0,35809	0,06338	-	0,23423	0,27928	0,13514	0,09910	0,00901
Nahuas	0,32839	0,07272	0,00478	-	0,00000	0,00000	0,10811	0,00901
Otomías Valle	0,26501	0,06656	0,00399	0,02960	-	0,07207	0,00000	0,00000
Otomías Sierra	0,28780	0,07224	0,00763	0,03112	0,01202	-	0,00000	0,00000
Tepehuas	0,43413	0,12458	0,01990	0,01374	0,05040	0,04940	-	0,03604
Mayas	0,48584	0,15594	0,04693	0,02735	0,07503	0,05655	0,02876	-

Tabla III.41. Resultados de las distancias genéticas Fst aplicadas a los ocho grupos indígenas utilizando la región HVI. En la hemimatriz inferior se muestran los valores de Fst, en la superior el valor de significación (p)

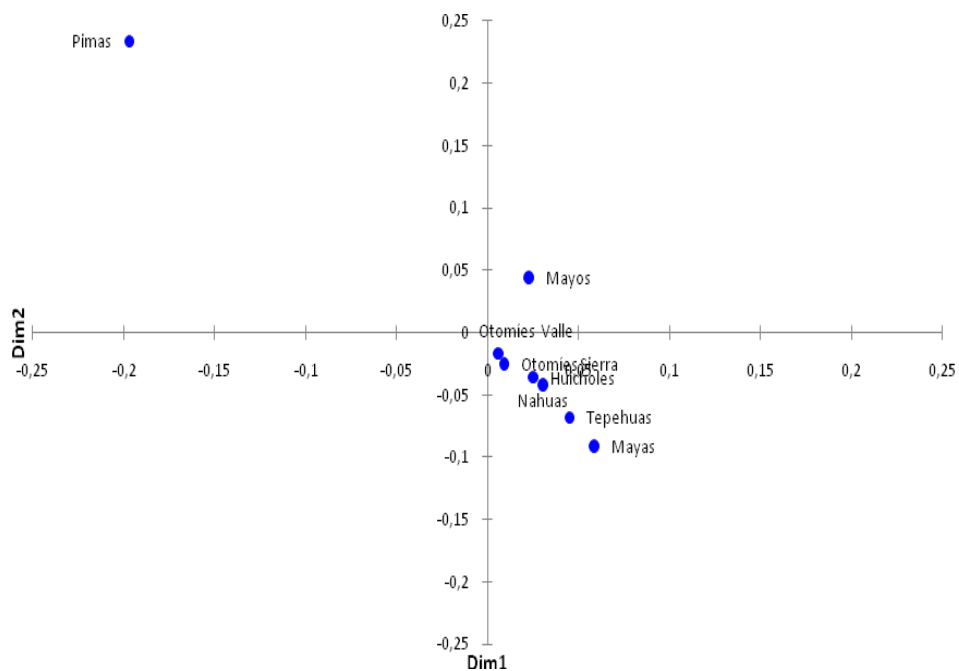


Figura III.37. MDS construido a partir de la distancia genética Fst y considerando la región HVI para las ocho poblaciones estudiadas (*stress* = 0,137).

A pesar de las diferencias detectadas al utilizar D-Loop y HVI, los resultados del MDS son muy semejantes. En la figura III.37 se representa el MDS construido a partir de HVI. Aunque el *stress* es ligeramente superior al calculado para el de la región D-Loop, los resultados son muy parecidos. Se diferencian claramente la población Pima, muy

alejada del resto, y los Mayos. Los Mayas parecen agruparse con estos, aunque ligeramente diferenciados de los otros cinco grupos.

Respecto al análisis de HVII los resultados del test exacto de diferenciación se muestran en la tabla III.42. A diferencia del resto de análisis, en este caso se han detectado dos poblaciones que no presentan diferencias significativas, Otomíes del Valle y de la Sierra ($p = 0,06780$).

Respecto a las distancias genéticas, la población Huichol no presenta diferencias significativas con tres grupos indígenas; Nahuas ($F_{st} = 0,00918$; $p = 0,19820$), Otomíes del Valle ($F_{st} = 0,01031$; $p = 0,22523$) y Otomíes de la Sierra ($F_{st} = 0,00990$; $p = 0,15315$). Igual que con los test exactos de diferenciación, los dos grupos Otomíes tampoco presentan diferencias significativas en sus distancias genéticas ($F_{st} = 0,00069$; $p = 0,34234$) (Tabla III.43).

	Pimas	Mayos	Huicholes	Nahuas	Otomíes Valle	Otomíes Sierra	Tepehuas	Mayas
Pimas		+	+	+	+	+	+	+
Mayos	0,000		+	+	+	+	+	+
Huicholes	0,000	0,000		+	+	+	+	+
Nahuas	0,000	0,000	0,000		+	+	+	+
Otomíes Valle	0,000	0,000	0,000	0,001		-	+	+
Otomíes Sierra	0,000	0,000	0,000	0,000	0,067		+	+
Tepehuas	0,000	0,000	0,000	0,000	0,000	0,000		+
Mayas	0,000	0,000	0,001	0,000	0,000	0,000	0,000	

Tabla III.42. Resultados de la aplicación del test de diferenciación poblacional a los ocho grupos indígenas utilizando la región HVII. En la hemimatriz inferior se muestran los valores de significación (p), en la superior el resultado del test; (+) diferencias significativas; (-) diferencias no significativas

El MDS (Figura III.38) construido con las distancias F_{st} al estudiar la región HVII presenta una distribución muy semejante a la obtenida con D-Loop y HVI. En este caso se diferencia sobre todo la población Pima del resto, seguida de Mayos y Mayas, que parecen alejarse del *cluster* formado por el resto de grupos indígenas. Aunque el índice de *stress* está por encima de los valores admitidos para dos dimensiones y ocho poblaciones (0,071) (Sturrock and Rocha, 2000).

	Pimas	Mayos	Huicholes	Nahuas	Otomíes Valle	Otomíes Sierra	Tepehuas	Mayas
Pimas	-	0,00000	0,00000	0,00000	0,00000	0,00901	0,00000	0,00000
Mayos	0,17776	-	0,01802	0,00000	0,00000	0,00000	0,00000	0,00000
Huicholes	0,21370	0,04277	-	0,19820	0,22523	0,15315	0,04505	0,02703
Nahuas	0,32997	0,07240	0,00918	-	0,03604	0,02703	0,01802	0,00000
Otomíes Valle	0,27807	0,03741	0,01031	0,01564	-	0,34234	0,00000	0,00000
Otomíes Sierra	0,28725	0,05323	0,00990	0,01341	0,00069	-	0,00000	0,00901
Tepehuas	0,31187	0,06032	0,02881	0,02671	0,04169	0,02902	-	0,00000
Mayas	0,35245	0,12280	0,04860	0,04746	0,08508	0,06226	0,05853	-

Tabla III.43. Resultados de las distancias genéticas Fst aplicadas a los ocho grupos indígenas utilizando la región HVII. En la hemimatriz inferior se muestran los valores de Fst, en la superior el valor de significación (p)

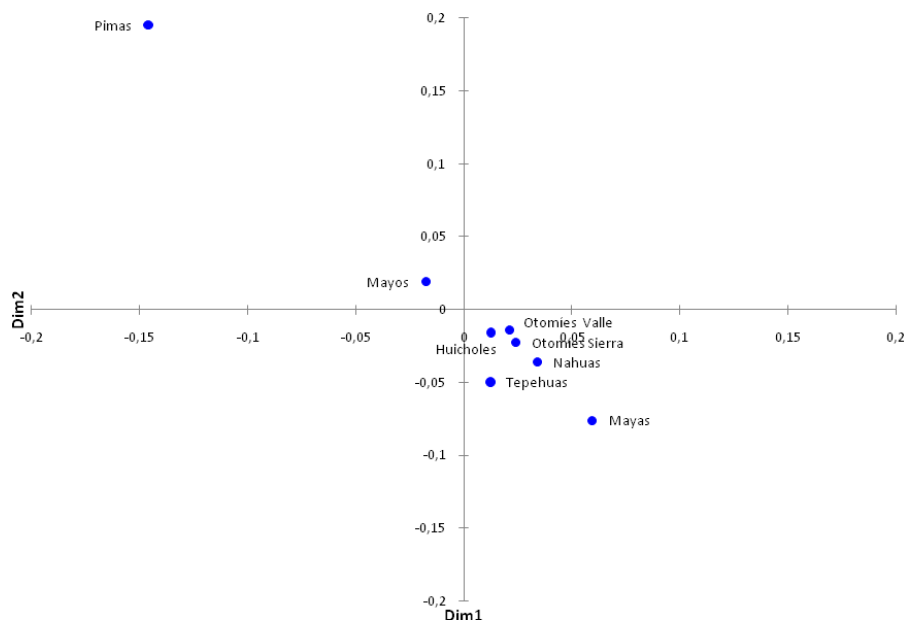


Figura III.38. MDS construido a partir de la distancia genética Fst y considerando la región HVII para las ocho poblaciones estudiadas ($stress = 0,117$).

Las *pairwise differences* para todas las poblaciones y construidas a partir de la región D-Loop están representadas en las figuras III.A y III.B. Las mismas distribuciones para las regiones HVI y HVII y para todas las poblaciones se representa en el capítulo VII. Anexo (Figuras VII.13 a VII.26).

Con el objetivo de ahondar más en las relaciones y diferenciación genética entre los distintos grupos indígenas, se han realizado AMOVAs atendiendo a distintos criterios de clasificación. Dado que la información aportada por la región control, HVI y HVII esencialmente es la misma, estos cálculos se han realizado en base a las secuencias completas de la región D-Loop.

Se han realizado seis AMOVAS (Tabla III.44) utilizando criterios geográficos, históricos y culturales. Para los criterios geográficos se ha utilizado una distribución que atiende a una separación entre “Norte”, “Centro” y “Sur”. Esta agrupación está en concordancia con las regiones culturales expresadas por la geografía mexicana y detallada en el apartado I.6. Regiones culturales del capítulo I. Introducción. Los resultados para esta agrupación, a pesar de cierta lógica geográfica detectada en los MDS, son no significativos, con una diversidad entre grupos del 8,77 y un $F_{st} = 0,08868$ ($p = 0,09384$).

En base al apartado mencionado, se han realizado agrupaciones que se han llamado “Áreas Culturales I” y “Áreas Culturales II”, utilizando la ubicación geográfica del grupo indígena y el área cultural a la que pertenece. Las poblaciones estudiadas representan a las diferentes áreas con una identidad cultural propia, e incluyen “Norte”, “Occidente”, “Centro”, “Costa” y “Sur”. Estas agrupaciones se han hecho de dos maneras diferentes con el objetivo de ahondar en la complejidad de la Huasteca. Como se puede ver en la tabla III.44, en ambas se han mantenido los grupos indígenas incluidos en las agrupaciones “Norte”, “Occidente” y “Sur”, y en las categorías “Centro” y “Costa” en las que se han variado los grupos incluidos. Otomíes del Valle, y Otomíes de la Sierra, Tepehuas y Nahuas respectivamente, en el caso de las “Áreas Culturales I”; y Otomíes del Valle, Otomíes de la Sierra y Nahuas, y Tepehuas respectivamente, en el caso de “Áreas Culturales II”. Los resultados para ambas clasificaciones son significativos en todos los niveles de agrupación (Tabla III.44). En ambos casos los niveles de diversidad representados para cada agrupación son muy semejantes, con valores que oscilan entre 84,11 y 86,35 para la variación dentro de las poblaciones, y 8,87 y 8,82 para la variación entre grupos.

Criterio	Agrupación	Grupo	Fuente de variación	Porcentaje varianza	Índice fijación	P
Geográfico	Norte	Pimas, Mayos	Dentro de poblaciones	84,11	Fst = 0,15889	0,00000*
	Sur	Mayas	Entre poblaciones dentro de grupos	7,02	Fsc = 0,07705	0,00000*
	Centro	Huicholes, Nahuas de la Huasteca, Otomíes Valle, Otomíes Sierra, Tepehuas	Entre grupos	8,87	Fct = 0,08868	0,09384
Áreas Culturales I	Norte	Pimas	Dentro de poblaciones	86.35	Fst = 0.13648	0.00000*
	Occidente	Mayos, Huicholes	Entre poblaciones dentro de grupos	4.72	Fsc = 0.05187	0.00000*
	Centro	Otomíes Valle	Entre grupos	8.92	Fct = 0.08923	0.04985*
	Costa Sur	Otomíes Sierra, Tepehuas, Nahuas Mayas				
Áreas Culturales II	Norte	Pimas	Dentro de poblaciones	85.77	Fst = 0.14225	0.00000*
	Occidente	Mayos, Huicholes	Entre poblaciones dentro de grupos	4.66	Fsc = 0.05157	0.00000*
	Centro	Otomíes Valle, Otomíes Sierra, Nahuas	Entre grupos	9.56	Fct = 0.09562	0.03128*
	Costa Sur	Tepehuas Mayas				
Lenguas	Yuto-Nahua	Pimas, Huicholes, Mayos, Nahuas	Dentro de poblaciones	89.36	Fst = 0.10644	0.00000*
	Otomangue	Otomíes Sierra, Otomíes Valle	Entre poblaciones dentro de grupos	15.58	Fsc = 0.14846	0.00000*
	Totonaco-Tepehua Maya	Tepehua Mayas	Entre grupos	-4.93	Fct = -0.04935	0.58847
Regiones indígenas	Norte	Pimas	Dentro de poblaciones	86.26	Fst = 0.13744	0.00000*
	Mayo	Mayos	Entre poblaciones dentro de grupos	3.77	Fsc = 0.04191	0.00000*
	Huicot	Huicholes	Entre grupos	9.97	Fct = 0.09970	0.07820
	Huasteca	Otomíes Sierra, Nahuas, Tepehuas				
	Otomí Yucatán	Otomíes Valle Mayas				
Historia	Chichimecas	Pimas, Mayos	Dentro de poblaciones	87.83	Fst = 0.12166	0.00000*
	Aztecas	Nahuas, Tepehuas	Entre poblaciones dentro de grupos	9.37	Fsc = 0.09641	0.00000*
	Tarascos	Huicholes	Entre grupos	2.79	Fct = 0.02794	0.24047
	Otomí	Otomíes Valle, Otomíes Sierra				
	Mayas	Mayas				

Tabla III.44. Resultados de las AMOVA realizadas para distintas agrupaciones. El valor de significación (p) está marcado con un asterisco cuando indica diferenciación

Las probabilidades asociadas a los Fct son significativas, tanto para “Áreas Culturales I” (Fct = 0.08923; p = 0,04985), como para “Áreas Culturales II” (Fct = 0.09562; p = 0.03128).

La AMOVA con agrupaciones lingüísticas, atendiendo a los criterios del INALI (INALI 2008) y detallados en el apartado I.8. Clasificaciones lingüísticas no ha resultado significativa.

El análisis estableciendo agrupaciones en las llamadas zonas indígenas definidas por el CoNaDePI y detallado en el apartado I.10. Zonas y grupos indígenas analizados tampoco es significativa (Fct = 0.09970-, p = 0,07820).

Finalmente, se ha creado una clasificación considerando el pasado histórico de los grupos indígenas estudiados en esta tesis, cuyos resultados tampoco han sido significativos (Fct = 0,02794; p = 0,24047).

Por último se han determinado los haplotipos únicos en cada población para la región D-Loop, HVI y HVII. Respecto a la región D-Loop completa, la población que ha presentado un mayor número de haplotipos únicos ha sido la Nahua de la Huasteca con 74, seguida de los Otomíes de la Sierra con 39 y por los Mayos con 31. Para este mismo parámetro las poblaciones con menor número de haplotipos exclusivos han sido los Tepehuas (18) y los Pimas (16). En la tabla VII.2 del capítulo VII. Anexo pueden consultarse los detalles de estos análisis. En ellos se especifica el número de haplotipos únicos y su referencia, que puede consultarse en la base de datos Haplotipos.xls incluida en el CD (Hoja D-Loop).

Este mismo parámetro, para HVI (Tabla VII.3, VII. Anexo), arroja datos muy parecidos aunque con un valor numérico lógicamente menor. La población con mayor número de haplotipos únicos sigue siendo la Nahua (34), seguida de los Otomíes de la Sierra (16) y del Valle (15). Las poblaciones con un valor menor volvieron a ser los Tepehuas (6) y los Pimas (3). Se presentan los datos de la misma forma que para la D-Loop aunque en este caso para consultar la composición polimórfica de cada haplotipo puede consultarse en la base de datos Haplotipos.xls incluida en el CD (Hoja HVI).

Los haplotipos únicos para HVII dan resultados ligeramente distintos. En este caso la población con un número mayor de variantes mitocondriales únicas, son los Nahuas (16), seguidos de Mayos (10) y Mayas (9). Los grupos indígenas con valores más pequeños han sido los Otomíes del Valle (5), Tepehuas (5) y los Pimas (3). Para

consultar los haplotipos puede consultarse en la base de datos Haplotipos.xls incluida en el CD (Hoja HVII).

La filosofía para presentar los haplotipos compartidos es la misma. Las tablas se han clasificado en el apartado Anexos VII, conteniendo; tabla VII.5 los datos referentes a D-Loop, tabla VII.6 los de la región HVI y la tabla VII.7 los de HVII. La consulta de los haplotipos completos se puede realizar en la base de datos ya mencionada (Haplotipos.xls) en las pestañas D-Loop, HVI y HVII para las regiones respectivas.

Los haplotipos compartidos entre las poblaciones para la región D-Loop son muy escasos. En un hecho lógico ya que como se ha comentado anteriormente, existe gran cantidad de variantes polimórficas únicas, lo que impide que se compartan haplotipos. El mayor número de haplotipos compartido es entre Nahuas y Otomíes del Valle (8). Los Nahuas comparten bastantes haplotipos con Otomíes de la Sierra (7) y estos últimos con Otomíes del Valle (7). Las poblaciones que menos haplotipos comparten son Mayas, aunque tienen dos iguales a Nahuas y Pimas.

La región HVI es quizás la más interesante de analizar por dos motivos. El primero de ellos es la que tradicionalmente se han utilizado para caracterizar y estudiar poblaciones. El segundo aspecto es porque a partir de esta información se generarán los *networks*, incluidos en el apartado siguiente. La información aportada por esta región es de todas formas muy semejante a la aportada por D-Loop aunque con un rango de haplotipos compartidos algo mayor. Las poblaciones con mayor rango de variantes mitocondriales compartidas son Nahuas: 14 con Otomíes de la Sierra, 13 con Otomíes del Valle, 9 con Mayas y 8 con Tepehuas. Es destacable que con Pimas solo comparten 2 haplotipos. El resto de valores es semejante a los ya expuestos para D-Loop.

Respecto a HVII completa, se puede apreciar que la población Nahua es la que comparte mayor número de haplotipos con el resto de poblaciones. De hecho, comparte 7 con Otomíes del Valle y de la Sierra, y 6 con Mayas. Los dos grupos Otomíes también comparten 7 haplotipos. Tepehuas comparten 6 haplotipos con Otomíes del Valle. Se aprecia cierta carencia en la aparición conjunta entre las poblaciones del altiplano y Pimas y Mayas, y llama la atención la escasez de haplotipos compartidos entre Huicholes con la mayor parte de poblaciones.

III.8. Networks

Los *networks* son un sistema de representación que permite determinar, haplogrupo a haplogrupo, la relación de variantes mitocondriales que hay entre diferentes poblaciones (Bandelt et al., 1995). Para evitar una red de haplotipos muy compleja y de difícil interpretación, se ha utilizado únicamente HVI. Como complemento a esta información en el apartado VII. Anexos pueden consultarse las distribuciones *pairwise differences* para cada uno de los haplogrupos (Figuras VII.27, VII.28. VII.29 y VII.30).

El haplogrupo más complejo y también el más frecuente en las poblaciones estudiadas, es A2. Por este motivo el *network* es bastante complejo (Figura III.39), aunque pueden destacarse algunos datos claros. Existe un nodo central formado por cinco poblaciones: Nahuas de la Huasteca, Tepehuas, Otomíes del Valle y de la Sierra, y Mayas. De este nodo parten variantes haplotípicas en forma de estrella entre las que destaca la formada por la mutación 16.111. Constituye un nuevo nodo representado por las mismas poblaciones que el anterior con una excepción, en lugar de la población Maya está representado el grupo Huichol. Un tercer nodo que parte del central y está caracterizado por la variante 16.032, está compuesto por todas las poblaciones excepto por Pimas y Otomíes del Valle. Tomando nuevamente el nodo definido como central, el polimorfismo 16.335 define a una nueva variante haplotípica que comparten únicamente Nahuas y Otomíes de la Sierra. También el polimorfismo 16.329 define una nueva agrupación, constituida por los dos grupos Otomíes y Nahuas. Al margen de otros nodos interesantes, uno de los aspectos más destacables es que en la mayor parte de ellos, al menos en los más numerosos, siempre aparece la población Nahua y en muchas otras ocasiones una o ambas poblaciones Otomíes.

El *network* para el haplogrupo B2 es bastante más sencillo que el anterior en cuanto a la complejidad de interrelaciones (Figura III.40). Aparece un nodo central en el que están representadas todas las poblaciones excepto Pimas y Tepehuas. El polimorfismo 16.182 define un nodo secundario constituido casi completamente por poblaciones Nahuas y Pimas, y en menor proporción por Huicholes y Mayas. A partir de este último y por aparición de la variante 16.111, aparece una agrupación formada exclusivamente por Mayas. También se diversifican otros haplotipos constituidos casi exclusivamente por la población Nahua. La variante 16.295 desde el nodo central, define una nueva agrupación constituida por Otomíes de la Sierra, Nahuas y Tepehuas.

El haplogrupo C1 presenta una clara distribución en estrella a partir de un único nodo central (Figura III.41). En esta agrupación se representan todas las poblaciones excepto Mayos y Huicholes. El polimorfismo 16.295 agrupa a individuos procedentes de regiones del norte, es decir, exclusivamente formado por Pimas y Mayos. La variante 16.051 caracteriza exclusivamente a variantes presentes en poblaciones Otomíes.

El haplogrupo D1 es aparentemente el menos complejo (Figura III.42). Presenta un nodo central donde aparecen cinco poblaciones, todas excepto Pimas, Huicholes y Mayas. Tres haplotipos secuenciales parten de él por los polimorfismos 16.142, 16.292 y 16.305. Todos ellos están exclusivamente representados por el grupo Mayo. Por último comentar que la variante 16.274 forma un nodo constituido por poblaciones Otomíes.

III.9. Barreras geográficas vs barreras genéticas

Una vez abordada la relación genética entre las poblaciones incluidas en la tesis, se propone situarlas en el contexto poblacional de los grupos indígenas de la región. Para ello se han recopilado datos referentes a poblaciones indígenas que proceden desde Estados Unidos hasta Panamá ubicadas entre 40,51° y 7,00° de latitud. Se han considerado estas poblaciones, y no únicamente las pertenecientes al ámbito mesoamericano estrictamente, para determinar si existe una continuidad genética más allá de la región de estudio propuesta en este trabajo. Mesoamérica tiene continuidad, geográfica y cultural hacia el norte, en los actuales EEUU. Hacia el sur ocurre lo mismo, ya que la región Maya tiene su continuidad hasta Guatemala y hacia regiones más meridionales.

Desgraciadamente de la mayor parte de poblaciones publicadas solamente se dispone de las frecuencias de los haplogrupos amerindios, y no hay información de los haplotipos. Esta circunstancia quita potencia a los análisis estadísticos, ya que las comparaciones se tendrán que realizar en base a las frecuencias de los cuatro haplogrupos amerindios, más una quinta clasificación categorizada como "Otros". De todas formas, la utilización de métodos estadísticos multivariantes permitirá realizar una aproximación probabilística ajustada. La tabla VII.8 del apartado VII. Anexos recoge el nombre de la población, su procedencia (comunidad, estado, país), la frecuencia de las cinco categorías para los haplogrupos, las coordenadas geográficas

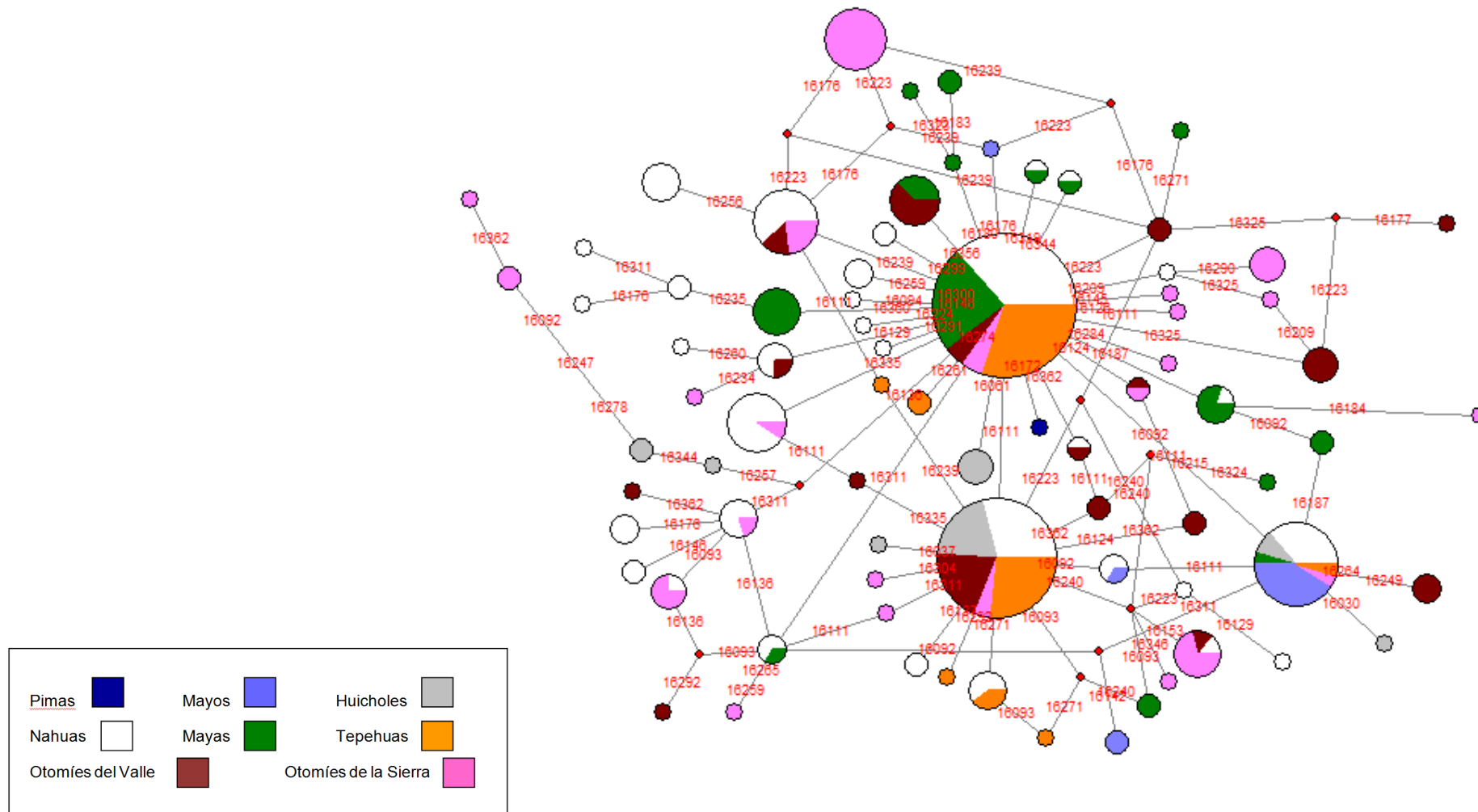


Figura III.39. *Network* del haplogrupo A2 construido a partir de HVI

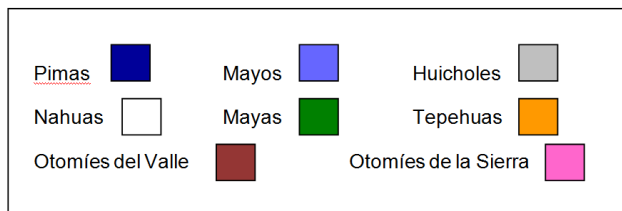
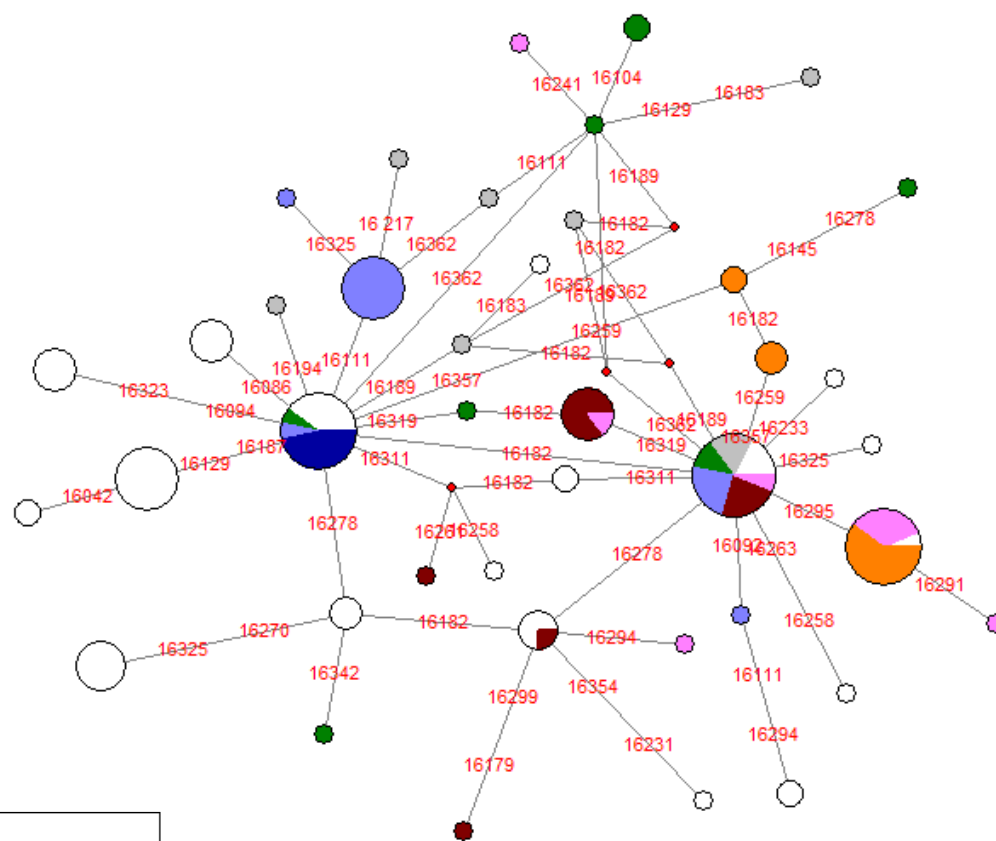


Figura III.40. Network del haplogrupo B2 construido a partir de HVI

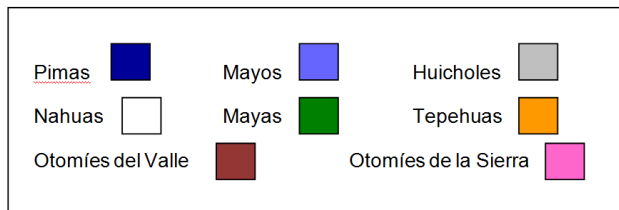
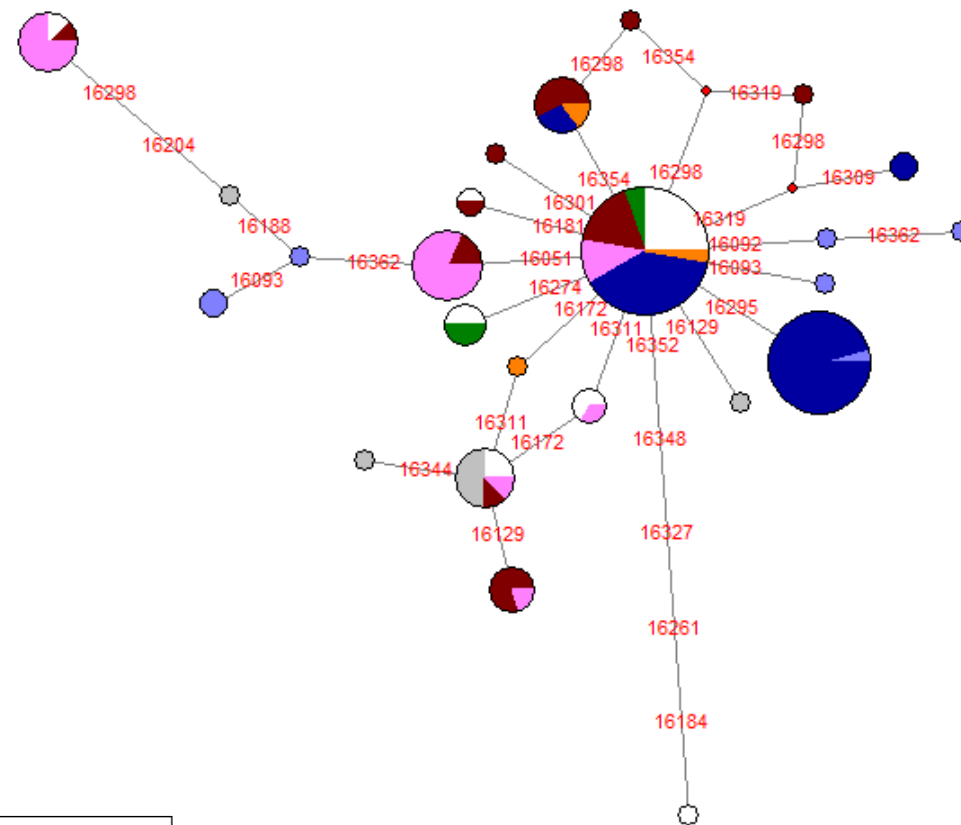


Figura III.41. *Network* del haplogrupo C1 construido a partir de HVI

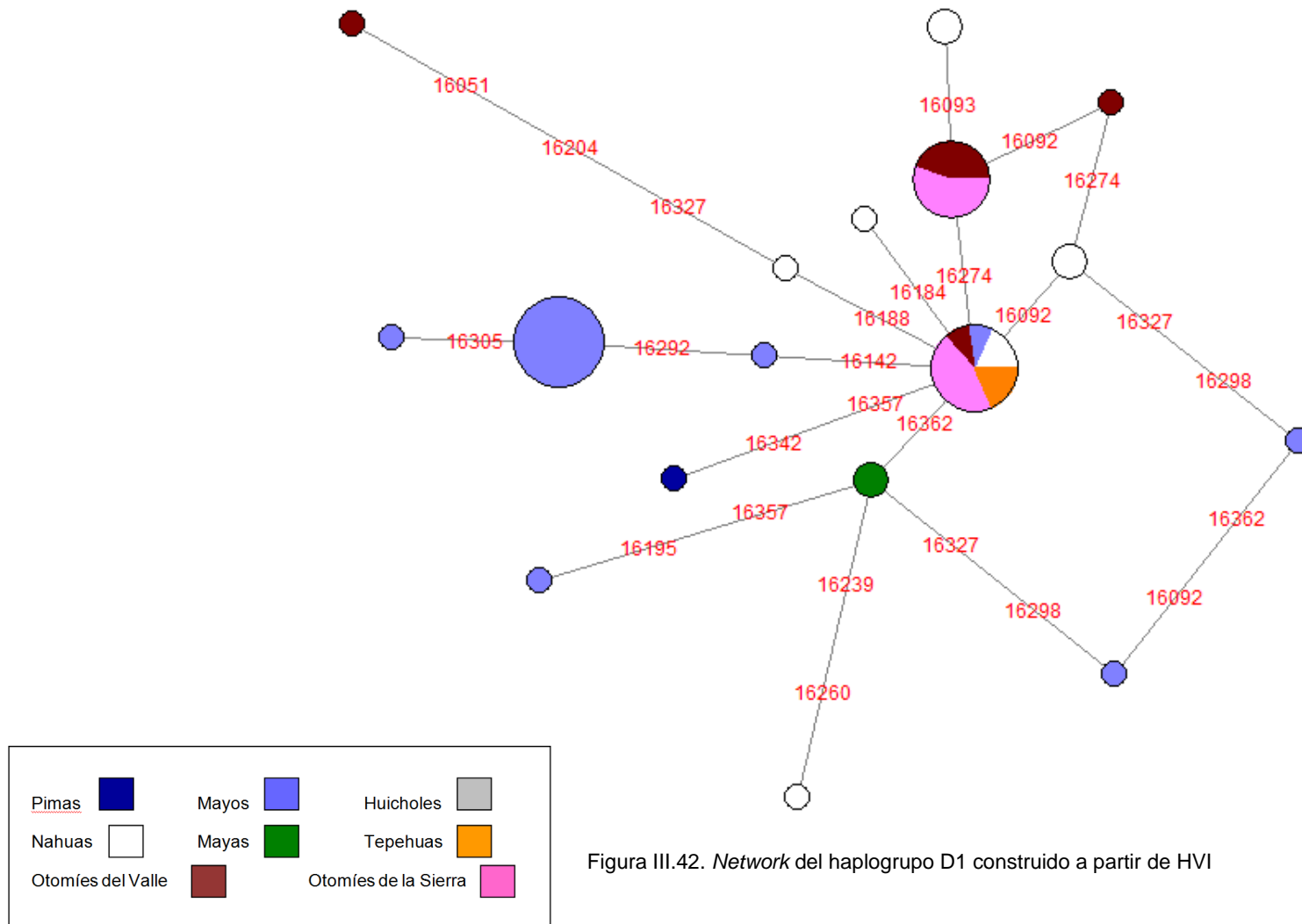


Figura III.42. *Network* del haplogrupo D1 construido a partir de HVI

(latitud y longitud), así como la referencia a la publicación de la que proceden. Los criterios de inclusión utilizados han sido la pertenencia a la región geográfica y un tamaño muestral superior a 20. Muchos autores han clasificado los haplogrupos como A, B, C y D, por lo que se utilizará esta nomenclatura sin distinguir A2, B2, C1 y D1.

Para los grupos indígenas cuyas muestras proceden de varias comunidades, se ha tomado un punto de referencia geográfico, asumiendo que los errores cometidos por esta desviación serán mínimos. En total se compararán 35 poblaciones, 27 de las cuales han sido publicadas con anterioridad. De ellas, seis proceden de Estados Unidos, dos de Costa Rica y dos más de Panamá. El resto proceden de México.

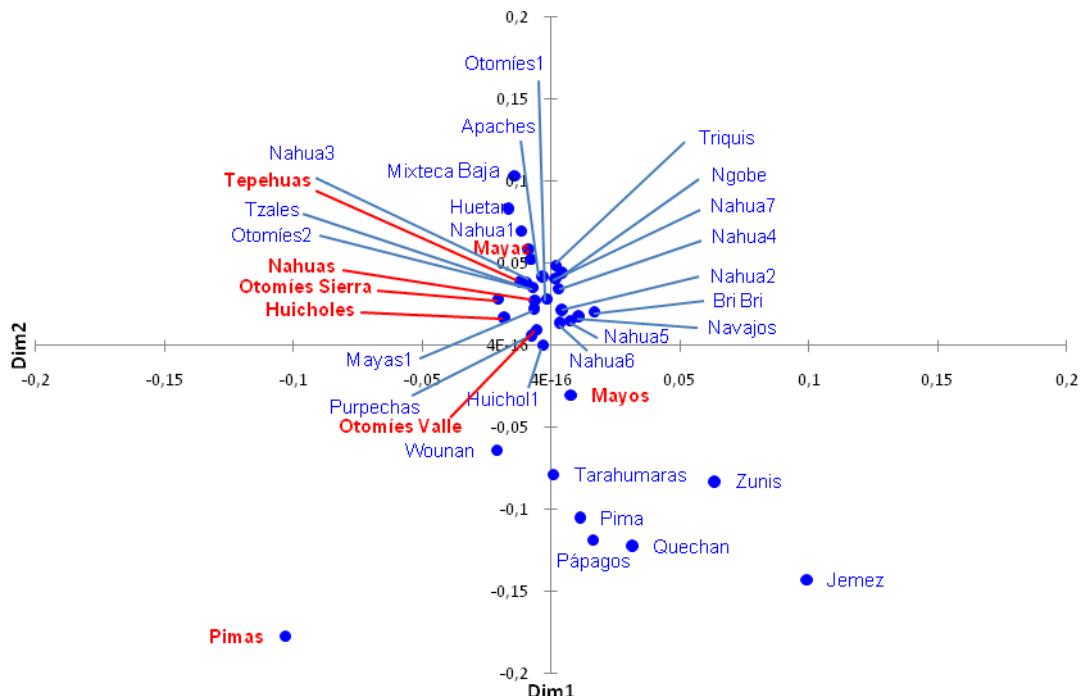


Figura III.43. MDS que representa la posición de las poblaciones estudiadas (en color rojo) junto con otros grupos indígenas de la región geográfica (*stress* = 0,250)

Los resultados del MDS basados en la distancia genética de Nei se muestran en la figura III.43. Aunque los valores de *stress* (0,250) son elevados, son perfectamente asumibles, ya que para esta cantidad de poblaciones y para dos dimensiones el límite de *stress* no debe de ser superior a 0,342 (Sturrock and Rocha, 2000). Las poblaciones estudiadas en la tesis se representan en color rojo para resaltarlas del resto. Se pueden observar dos *cluster* bien definidos. Uno de ellos engloba claramente

a las poblaciones procedentes de Aridoamérica con la sola excepción del grupo Wounan, procedente de Panamá. La población Pima estudiada en este trabajo se diferencia claramente del resto, ocupando una posición marginal dentro del *cluster* norteño. El otro *cluster* está configurado por el resto de poblacionales procedentes de Mesoamérica y Centroamérica, aunque también incluye a Apaches y Navajos, que proceden del sur de Estados Unidos. Las poblaciones integradas en esta agrupación están muy concentradas, lo que implica que sus diferencias no son muy acusadas. Aparentemente no presentan un patrón muy definido aunque sí que se puede apreciar que la totalidad de las poblaciones Otomíes, incluyendo los del Valle y de la Sierra, se distribuyen en un mismo *subcluster* que ocupa el cuadrante superior izquierdo. Otro aspecto a tener en cuenta es la falta de un modelo para las poblaciones Nahuas, que parecen dispersarse sin ningún orden aparente dentro del *cluster*.

En un intento de comprender en profundidad la distribución de las poblaciones, se ha aplicado el programa Barrier. Este programa, como ya se ha explicado en el apartado II.11. Software y determinaciones estadísticas del capítulo II. Material y Métodos, detecta la existencia de barreras a partir de datos, en este caso, genéticos. Los resultados se expresan como fronteras que representan rupturas en la continuidad de las frecuencias genéticas de los haplogrupos, y suelen acompañarse de un valor numérico que indica su intensidad. De hecho, es una reconstrucción geográfica de una matriz de distancias (F_{st}) con 1000 repeticiones.

En la figura III.44 se puede observar la representación geométrica como resultado de modelar los datos genéticos y geográficos. Los valores numéricos junto a un punto indican el orden de la población, en este caso decreciente según la latitud y que coincide con el orden de las poblaciones en la tabla VII.8 del capítulo VII. Anexos. Las líneas de color negro representan una reconstrucción geométrica en torno al valor geográfico (latitud y longitud) de cada población. Las líneas rojas expresan la existencia de barreras, su grosor es proporcional a la intensidad de la barrera y queda expresado mediante un número también de color rojo. Aunque la revisión de la figura expresa de forma clara la topografía genética, continuidad y rupturas, de la distribución de los haplogrupos. En la figura III.45 se ha representado la misma información sobre un mapa físico de la región. En este caso se ha utilizado una imagen del área estudiada y sobre ella se han ubicado las poblaciones, atendiendo a sus referencias geográficas. Sobre ella se han reconstruido y superpuesto las barreras genéticas más importantes. Las líneas gruesas representan las barreras con una intensidad

comprendida entre el 75 y el 100%. La línea punteada expresa también barreras pero de intensidades menores, en este caso comprendidas entre 50 y 75%.

Estas representaciones se han generado a partir de las cinco frecuencias disponibles, pero para profundizar en la distribución de los haplogrupos se han generado cinco mapas más (Figuras III.46, III.47, III.48, III.49 y III.50) que representan las barreras existentes, considerando independientemente cada uno de los haplogrupos. En estos casos las limitaciones estadísticas son muy importantes, ya que para cada una de ellas puede representar únicamente una variable. Esta está expresada mediante ausencia o presencia del haplogrupo. Por este motivo las fronteras se han representado sin estimar su intensidad.

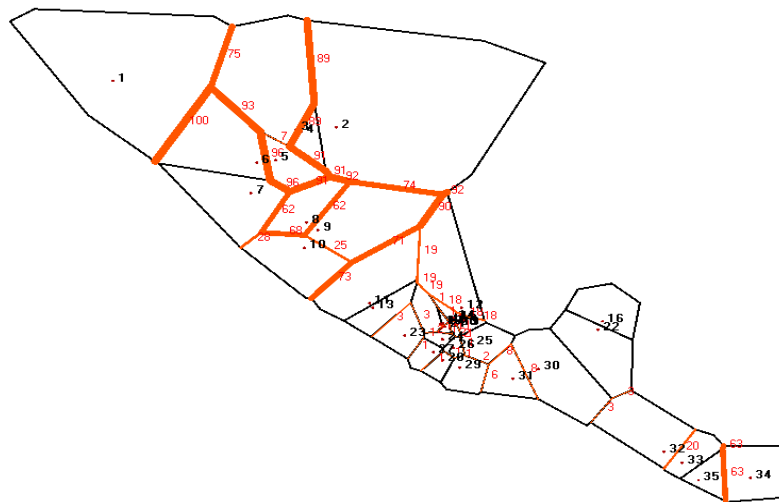


Figura III.44. Representación obtenida a través del programa Barrier. Las poblaciones se indican por un número negro que coincide con el orden de la tabla VII.8 del capítulo VII. Anexos. Las líneas rojas indican barreras genéticas, o discontinuidades en la distribución de las frecuencias de los haplogrupos, y van acompañadas por un número que indica su intensidad

A partir del análisis de la figura III.45 se puede observar claramente la existencia de una barrera genética que diferencia las poblaciones del norte del resto de grupos indígenas. Concretamente esta ruptura genética diferencia a grupos norteamericanos, representados por Quechan, Navajos, Jemez, Zunis y Apaches del resto. Dentro de este grupo norteño destacan los Apaches, es una población que se distribuye en latitudes menores que el resto y que parece desplazar la frontera hacia el sur.

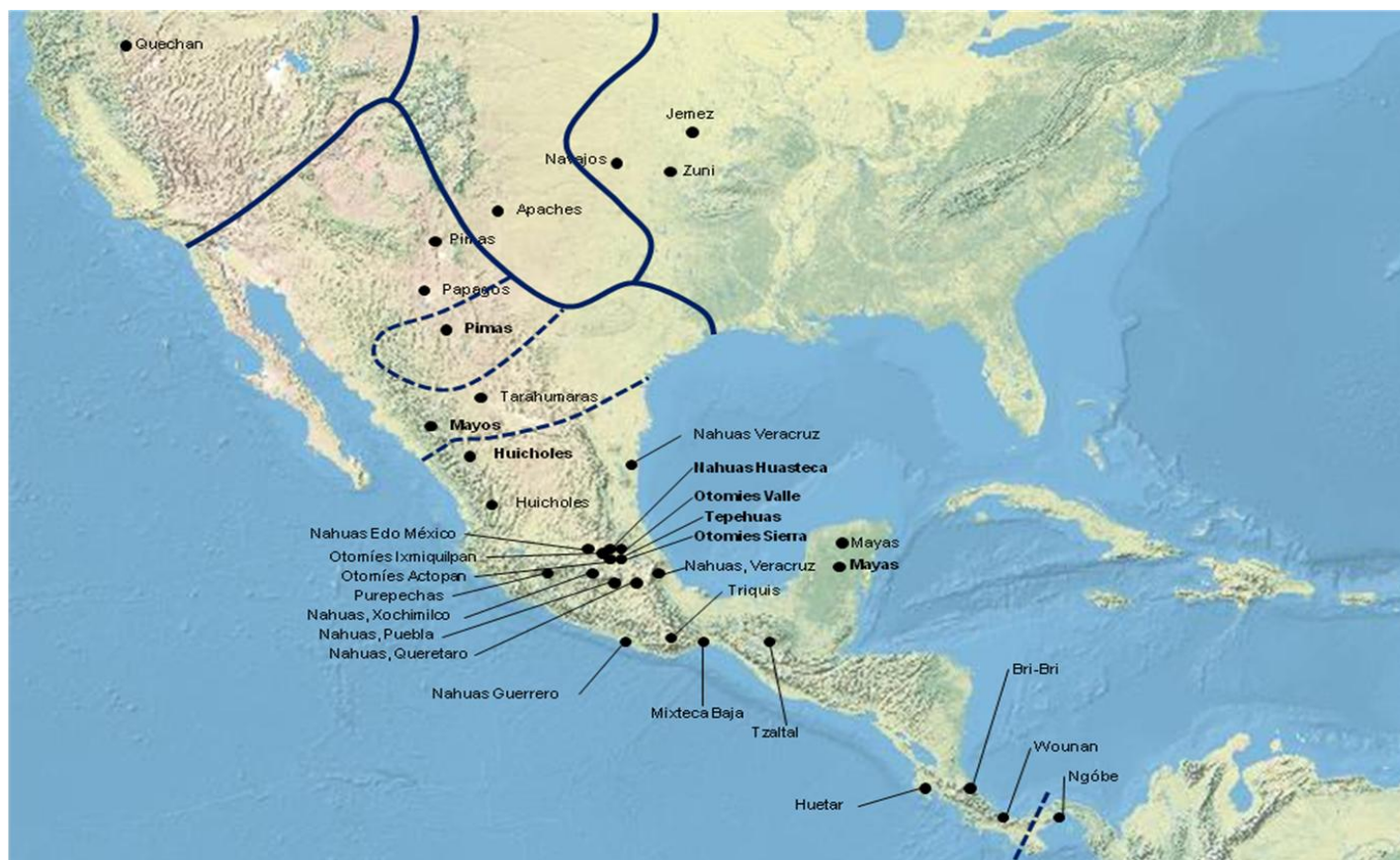


Figura III.45. Representación resumida de la información obtenida a través del programa Barrier. Las poblaciones inéditas, y estudiadas en la tesis, se representan en negrita. Las líneas azules continuas expresan barreras genéticas de intensidad comprendidas entre 75 y 100%. Las líneas azules discontinuas expresan barreras genéticas de intensidad comprendidas entre 5 y 75%. Para más detalles consultar el texto.

Con un rango de intensidad menor se pueden describir tres fronteras más. Una de ellas delimita exclusivamente a la población Pima estudiada en el presente trabajo del resto de poblaciones que se distribuyen en su entorno cercano. En la figura III.44 se puede apreciar como el valor de la barrera que delimita a esta población es superior al 60%. Es una población diferenciada de Pápagos, Mayos, Tarahumaras e incluso de grupos Pimas procedentes de otros trabajos. Con una intensidad comprendida entre 71 y 90% aparece otra barrera que separa Mesoamérica del resto de poblaciones del norte. Este límite separa a todas las poblaciones septentrionales hasta Tarahumaras y Mayos del resto de grupos de distribución mesoamericana. En la región meridional, en torno al actual Panamá, se puede describir otra barrera con una intensidad superior al 60%. Entre esta última y la anterior, se distribuyen la totalidad de poblaciones mesoamericanas y centroamericanas. Es decir, definen un espacio ocupado por la totalidad de los grupos mesoamericanos más alguna población de Centroamérica.

En las figuras III.46 a III.50 se representan los resultados de Barrier aplicados a cada uno de los haplogrupos. Respecto al haplogrupo A podemos apreciar que la mayor parte de disrupciones se encuentran en el norte, por encima de la región habitada por los Huicholes (Figura III.46). También se detecta una barrera en el sur y una en Mesoamérica. Esta última afecta a la población numerada como 24 que corresponde a los Nahuas de Xochimilco, distribuidos en torno a México D.F. Este grupo, a diferencia del resto de poblaciones cercanas, presenta una frecuencia de haplogrupo A relativamente elevada (72,1%).

La representación de la frecuencia del haplogrupo B también llama la atención, ya que de forma general se parece bastante a la del haplogrupo A, sobre todo en lo referente a Mesoamérica (Figura III.47). En este caso las fronteras vuelven a concentrarse en el norte y en las regiones más meridionales.

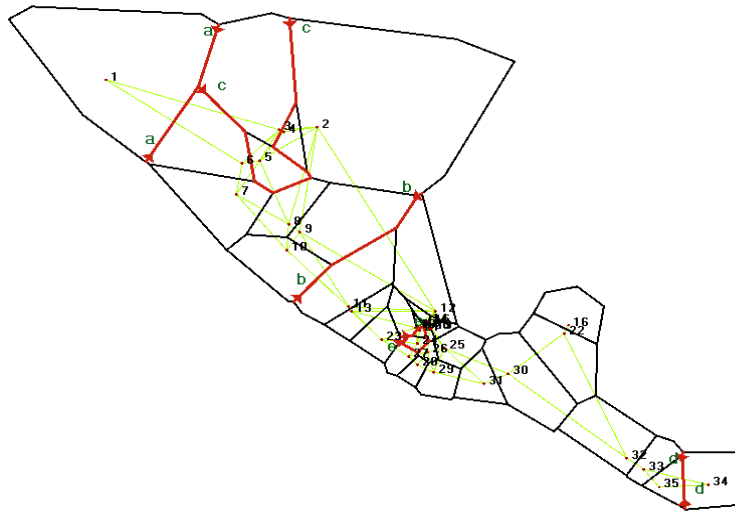


Figura III.46. Representación obtenida a través del programa Barrier según las frecuencias del haplogrupo A. Las poblaciones se indican con un número negro que coincide con el orden de la tabla VII.8 del capítulo VII. Anexos. Las líneas rojas indican barreras genéticas, o discontinuidades en la distribución de las frecuencias de los haplogrupos, y van acompañadas por un número que indica su intensidad.

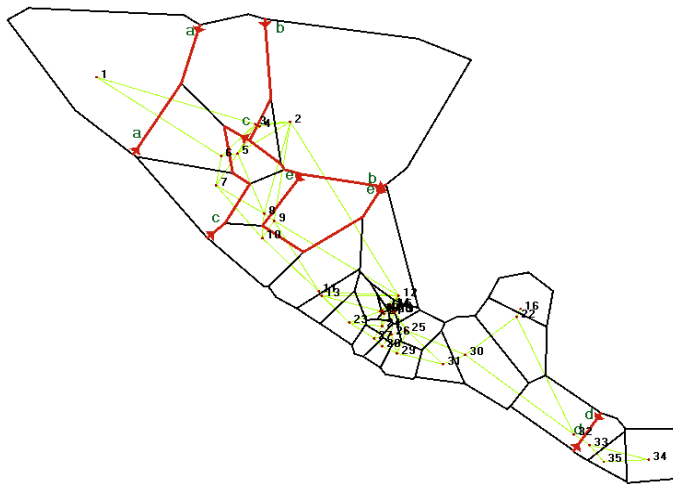


Figura III.47. Representación obtenida a través del programa Barrier según las frecuencias del haplogrupo B. Las poblaciones se indican con un número negro que coincide con el orden de la tabla VI.8 del capítulo VI. Anexos. Las líneas rojas indican barreras genéticas, o discontinuidades en la distribución de las frecuencias de los haplogrupos, y van acompañadas por un número que indica su intensidad.

El mismo panorama, con diferencias concretas, se puede observar en la representación de las frecuencias del haplogrupo C (Figura III.48). Nuevamente el

norte vuelve a ser el que concentra las interrupciones, a excepción de una barrera muy meridional que afecta a la región de Panamá.

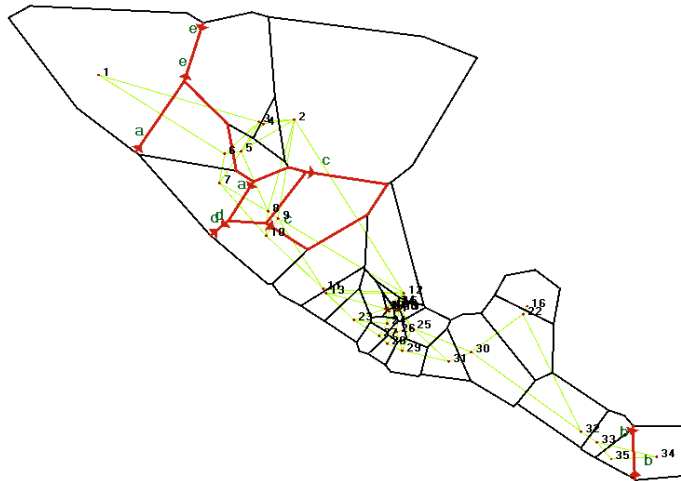


Figura III.48. Representación obtenida a través del programa Barrier según las frecuencias del haplogrupo C. Las poblaciones se indican con un número negro que coincide con el orden de la tabla VII.8 del capítulo VII. Anexos. Las líneas rojas indican barreras genéticas, o discontinuidades en la distribución de las frecuencias de los haplogrupos, y van acompañadas por un número que indica su intensidad.

La distribución de las frecuencias del haplogrupo D presenta ligeras diferencias respecto al resto (Figura III.49). En este caso se detectan barreras que diferencian las regiones del poniente respecto al levante, y además se detectan dos discontinuidades en el sur, una propia de la región centroamericana y otra que afecta justo al límite sur de la península de Yucatán.

También se ha realizado un análisis a partir de las frecuencias de los haplogrupos clasificados como “Otros” (Figura III.50). Los resultados hay que analizarlos con mucha cautela ya que se desconoce si estos haplogrupos podrían ser de origen amerindio, como el X, o bien se tratan de haplogrupos originarios de Europa y/o África. De todas formas se observa una barrera que diferencia la costa norte de México hasta la región sureste de Estados Unidos.

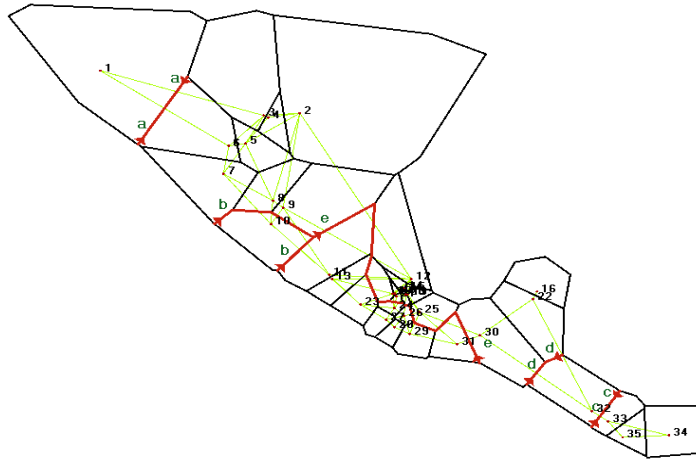


Figura III.49. Representación obtenida a través del programa Barrier según las frecuencias del haplogrupo D. Las poblaciones se indican con un número negro que coincide con el orden de la tabla VII.8 del capítulo VII. Anexos. Las líneas rojas indican barreras genéticas, o discontinuidades en la distribución de las frecuencias de los haplogrupos, y van acompañadas por un número que indica su intensidad.

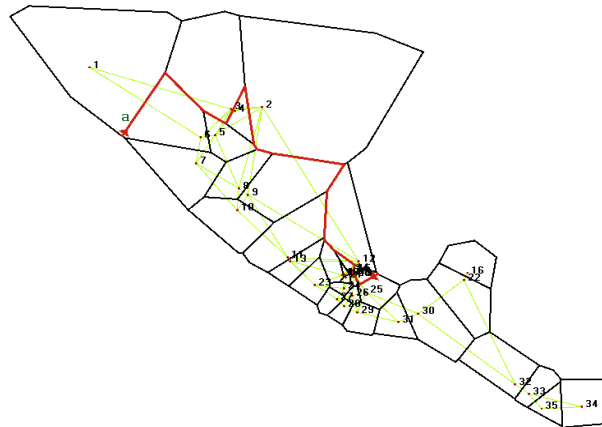


Figura III.50. Representación obtenida a través del programa Barrier según las frecuencias de los haplogrupos indefinidos clasificados por los autores como "Otros". Las poblaciones se indican con un número negro que coincide con el orden de la tabla VI.8 del capítulo VI. Anexos. Las líneas rojas indican barreras genéticas, o discontinuidades en la distribución de las frecuencias de los haplogrupos, y van acompañadas por un número que indica su intensidad.

IV. DISCUSIÓN

IV.1. Consideraciones sobre el muestreo

El trabajo de diseño hecho previo a las campañas de muestreo en las comunidades ha demostrado ser adecuado en el sentido global, así como en el particular para cada una de las regiones trabajadas. Para cada una de ellas los diseños funcionaron en el trabajo *in situ*, ya que en todos los casos se pudo trabajar del modo planificado obteniéndose un número de muestras lo suficientemente elevado como para realizar el análisis de los grupos. El muestreo directo en las comunidades permitió explicar los objetivos del estudio en cada comunidad, para que la donación de la muestra fuera con conocimiento de causa.

En algunos casos sin embargo, la planificación del muestreo tuvo que ser modificada, por la extrañeza y desconfianza que causaba la presencia de extranjeros en la comunidad o por hechos ajenos a nuestra presencia. A excepción de la campaña en zona Huichol y Mayo, se alargaron los días de campaña con el fin de permitir el flujo de información a toda la comunidad desde Presidencia Municipal, así como la interacción con los habitantes. Igualmente en todos los casos se procuró permanecer el máximo tiempo posible antes de empezar la toma de muestras, y se hicieron varios grupos de trabajo para diluir nuestra presencia entre investigadores mexicanos. Para facilitar estos aspectos, se aprovechó la organización de las mismas comunidades indígenas para estructurar el muestreo, por ejemplo, alojamiento e intendencia en casas de miembros de la comunidad.

Por otro lado, el muestreo directo tiene la gran ventaja de la visión de la estructura social y organización de las comunidades de un mismo grupo indígena, lo que permitió la obtención de información que en algunos casos resultó de gran utilidad para la creación de las bases de datos y complementar las encuestas. Igualmente la relación con los habitantes de cada comunidad indígena, permitió la resolución por parte de los participantes, de todo tipo de dudas durante los días de campaña acerca del estudio que se estaba realizando.

El modelo de encuesta utilizado fue actualizado para cada una de las campañas, optimizándose los apartados, la codificación *in situ*, el número de personas implicadas en la toma de muestras y en la realización de las encuestas. Se prestó especial atención a los topónimos de cada región así como a los apellidos característicos de cada zona, aspectos especialmente complejos en las zonas Pima y Maya. Igualmente es importante destacar los apartados referentes a familiares incluidos en la encuesta,

esenciales en la posterior selección de los individuos que serían incluidos definitivamente en el análisis. La información sobre la lengua indígena fue primordial puesto que mostraba la pertenencia del individuo al grupo, aunque en algunos casos fue un aspecto muy complejo de determinar. Por la situación social de los indígenas en México, es frecuente que los individuos nieguen hablar una lengua indígena ya que pone de manifiesto su origen, siendo esta razón de discriminación o menosprecio. Esta cuestión fue evidente durante las campañas en el trabajo realizado casa por casa o en las escuelas, para lo que se utilizaron distintas estrategias con el fin de resolverlo y obtener la información.

Estos aspectos fueron de especial relevancia en las campañas realizadas en escuelas (Sabán en Mayas y Pimas) en las que se prestó especial atención, ya que la presencia de individuos emparentados (especialmente hermanos y primos) en distintos cursos es muy habitual. En comunidades pequeñas y aisladas suele haber únicamente una escuela y las familias son numerosas, por lo que es frecuente encontrar cuatro o cinco hermanos y dos o tres primos para cada familia. Es esencial tener esto en cuenta durante el desarrollo del muestreo y a la hora de crear las bases de datos, puesto que lo normal es que el tamaño del muestreo efectivo sea entre un 30% y un 40% menor del real efectuado.

En el caso de la comunidad Maya de Tihosuco, el tamaño muestral se vio muy reducido debido a un suceso acaecido durante la campaña en la segunda de las comunidades mencionadas. Después de que la convocatoria fuera hecha a través de los altavoces de la comunidad por el Presidente Municipal, se congregaron la casi totalidad de familias con todos los niños, generando un muestreo con gran cantidad de individuos emparentados. Por esta razón y únicamente en este caso, se trabajó exclusivamente con las muestras procedentes de la comunidad de Sabán.

Para cada una de las bases de datos generadas mediante las encuestas, se creó un archivo en el que fue incluida toda la información recogida con la que, una vez ordenadas las muestras en el laboratorio y previo al inicio del trabajo en el mismo, se hizo la selección. En algunos casos y debido a la antigüedad de las muestras, se prefirió trabajar con la totalidad disponible, haciéndose la selección posteriormente. Este es el caso de los grupos indígenas Otomí, Tepehua y Nahuatl del Estado de Hidalgo.

Para hacer la selección se trabajó en primer lugar con los apellidos de los individuos, descartando todos aquellos que mostraran parentesco materno, así como todos los asignados como familiares (en todos los casos se había incluido al menos el nombre de pila en las encuestas), permitiendo una selección rápida. La lengua indígena fue tomada en cuenta también como criterio de selección en aquellos casos en que la procedencia de los ancestros del individuo no era de la región, por ejemplo, individuos hablantes de lengua Maya procedentes de otros estados. Se tuvieron en cuenta otros aspectos de exclusión como el caso concreto de un individuo pelirrojo y con ojos claros que, a pesar de hablar lengua Pima, se consideró que presentaba aportes genéticos externos a la comunidad de El Kipur.

Los porcentajes de efectividad del muestreo fueron variables, aunque con valores bajos para los grupos Mayas y Pimas, debido al elevado porcentaje de individuos emparentados en el muestreo. En el caso de la población Pima además, es importante tener en cuenta el bajísimo número de individuos pertenecientes a este grupo, únicamente con 741 hablantes en todo el país (INEGI 2000).

El tipo de muestreo diseñado para esta tesis, aunque está basado en los mismos principios y presenta claros puntos en común con la mayoría de trabajos realizados en otras regiones de Mesoamérica, tiene características distintas que lo hacen más exhaustivo y pormenorizado, reflejando la realidad genética de cada uno de estos grupos. Se muestrearon comunidades enteras, en la mayoría de los casos varias por municipio y en más de uno cuando fue posible, tratando de que todos los individuos participaran. Se evitó elegirlos al azar, así como trabajar con pocos individuos, detallando las comunidades exactas en las que se trabajó, al contrario que en otros estudios (Ibarra-Rivera et al., 2008; Martínez-Cortés et al., 2010; Peñaloza-Espinosa et al., 2007; Sandoval et al., 2009).

IV.2. Consideraciones sobre las muestras

Es importante mencionar que el número de participantes en las comunidades de Capomos y de San Sebastián de Teponahuastlán, Mayos y Huicholes respectivamente, en las que se extrajo sangre total fue menor, debido probablemente al efecto que tiene un muestreo más invasivo para los participantes. En otras comunidades, Pimas o Mayas por ejemplo, en que se muestreó cabello o epitelio bucal, el número de participantes fue superior. Independientemente de estas

consideraciones, el material, el trabajo de campo y la infraestructura necesaria para cada uno de los casos es absolutamente diferente, así como las condiciones de conservación de las muestras y los procesos de laboratorio asociados a cada uno de ellos. En el caso de las campañas en las que se obtuvo sangre fue necesario disponer de un contenedor refrigerado para la conservación de las muestras, y en el caso del muestreo de cabello fue imprescindible trabajar con tubos tipo eppendorf esterilizados para el depósito del material biológico. Aunque ambos casos tienen una clara desventaja en este sentido, en el caso de la sangre la cantidad de ADN obtenida es muy elevada y fue posible gracias al apoyo por el IMSS; en el caso del cabello sin embargo, fue una estrategia adaptada a las circunstancias de las comunidades que desconfiaban de la recogida y conservación de determinado tipo de muestras biológicas por parte del equipo de investigadores. Finalmente y tratándose el epitelio bucal de una muestra que apenas es visible en el algodón de un hisopo, se decidió adoptar esta estrategia para la campaña en comunidades Mayas y posteriores, por su sencillez y por su rendimiento en cuanto a cantidades de ADN.

En cuanto a la eficiencia de la metodología en el laboratorio, se realizaron un total de 716 extracciones con los distintos métodos. De estas, las 518 muestras de cabellos procedentes del estado de Hidalgo (Nahuas, Otomíes y Tepehuas) se hicieron con el método Chelex[®], 91 muestras de sangre de los grupos Huichol y Mayo, y 58 de epitelio bucal de la población Maya, se hicieron con extracción orgánica con fenolcloroformo isoamílico, y las 49 muestras de la población Pima se extrajeron con el kit ChargeSwitch Forensic DNA Purification de Invitrogen.

Cada una de estas muestras de ADN ya extraído, se amplificó según lo descrito anteriormente con unos porcentajes de error para todos los grupos cercanos al 5% a excepción de las PCR de la población Maya, en la que el porcentaje aumenta hasta el 23%. El valor tan elevado para la no amplificación se considera asociado al método de extracción, ya que el fenol es un conocido inhibidor de la PCR (Erlich 1989). Una insuficiente purificación podría haber dejado residuos en los extractos, degradando en parte la muestra y/o interfiriendo en el proceso de amplificación, explicando el elevado porcentaje de errores. Las amplificaciones de ADN procedente de sangre, aunque como se ha comentado suponen un problema añadido durante el muestreo por lo invasivo, tuvieron una efectividad del 100%. Las muestras extraídas con Chelex[®] dieron buenos resultados y aunque es un método muy cuestionado (Goldberg 1996; Prager et al., 1998; Saltonstall et al., 1998; Steffens and Roy, 1998), se obtuvieron bandas de intensidad correcta en su visualización y secuencias claras en su análisis.

Fue necesaria la resolución de un problema que se planteó para algunas muestras de los grupos Nahua (Na216, Na218, Na257, Na265, Na275 Na307 Na326, Na334, Na336, Na357), Tepehua (T467, T470, T558, T560, T561, T664) y Otomíes de la Sierra (OS473, OS482, 490, OS494). En las amplificaciones de los fragmentos II con la pareja de *primers* 16.555L- 611H no se obtuvieron bandas para ninguno de ellos, a pesar de haberse observado bandas intensas en los fragmentos I y estando seguros, por lo tanto, de que la cantidad de ADN presente en los extraídos era suficiente para dar lugar a una amplificación correcta.

Se consideró la posibilidad de que la ausencia de banda fuera debida al no anillamiento de uno de los *primers* en el ADN molde, por lo que se decidió trabajar dividiendo el fragmento II en otros dos como se realizó para la población Huichol, es decir con las parejas de *primers* 16.380L- 285H y 172L- 599H (Figura II.5). Una vez obtenidas las secuencias se revisaron las mismas para tratar de encontrar la mutación, no pudiéndose detectar en el extremo correspondiente al primer 16.555L. En el extremo opuesto, el *primer* utilizado en estos fragmentos más pequeños no incluye la región de anillamiento del 611H, por lo que aunque se obtuvieron las secuencias, no se pudo resolver la incidencia y hubo que suponer que era en esa región dónde se encontraba la mutación.

IV.3. Haplogrupos y haplotipos detectados

La asignación de los haplogrupos para cada uno de los haplotipos detectados se ha realizado trabajando con diversas publicaciones (Achilli et al., 2008; Behar et al., 2007; O'Rourke and Raff, 2009; Perego et al., 2009, 2010; Tamm et al., 2007), y recursos en línea (Mitosearch, Mitoweb, Genographic Project y Phylotree). En algunos casos sin la integración de todas las herramientas, la determinación del haplogrupo hubiera sido compleja, pues la gran diversidad de los individuos analizados complicó este proceso. Sin embargo, y pese a la gran cantidad de fuentes consultadas, un individuo quedó sin clasificar.

Para cada uno de ellos además, es importante tener en cuenta la antigüedad de los haplogrupos estimada por los distintos autores (Bonatto and Salzano, 1997a; Silva et al., 2002; Torroni et al., 1994d), ya que nos puede ayudar a situar el contexto sobre el que ubicaremos los grupos indígenas incluidos en esta tesis. Las estimaciones de antigüedad están hechas con diversas muestras en base al análisis de distintos

marcadores o regiones analizadas, basándose una en RFLP (Torrioni et al., 1994d), otra en el análisis de HVI (Bonatto and Salzano, 1997a) y la otra en un fragmento de 8.800 pb de la región codificante del ADNmt (Silva et al., 2002). Aunque algunos de estos estudios han encontrado valores bajos de diversidad para el haplogrupo B, y por lo tanto un menor tiempo al antecesor común más cercano, la mayor parte detectan valores de diversidad y tiempos estimados similares para los cuatro haplogrupos, dándose una antigüedad entre 15.000 y 28.000 AAP para todos ellos.

El análisis de los haplogrupos presentes en los grupos indígenas estudiados en esta tesis, sitúa las poblaciones en la región mesoamericana. En el caso de la población Pima encontramos una mayoría de haplogrupo C (81,6%) y B (14,3%), teniendo los haplogrupos A2 y D1 la mínima representación de 1 individuo (2,04%), lo que está en concordancia con pequeñas diferencias con lo reportado respecto a este grupo hasta la actualidad (Torrioni et al., 1993b; Lorenz and Smith, 1996; Malhi et al., 2003; Sandoval et al., 2009). Otros grupos cercanos a esta misma región, aunque situados un poco más al norte como Pápagos y Seris, presentan porcentajes similares a los aquí obtenidos (Malhi et al., 2003). En el caso de los Mayos, que tienen porcentajes más equilibrados para los cuatro haplogrupos detectados (A2 26,6%, B2 30,9%, C1 12,7% y D1 32,7%), no hay semejanzas con la población Pima ni con otras más cercanas como Purépechas o Tarahumaras, aunque estas últimas sin embargo, sí presentan semejanza con Pimas (Peñaloza-Espinosa et al., 2007; Sandoval-Ramirez et al., 2006; Sandoval et al., 2009). En este caso sin embargo, cabe tener en cuenta que los resultados para Tarahumaras son variables a pesar de existir cuatro trabajos, resultando que a dos de estas referencias la población Pima incluida en esta tesis es similar (Peñaloza-Espinosa et al., 2007; Sandoval-Ramirez et al., 2006), y con otras dos hay diferencias importantes (Kemp et al., 2010; Sandoval et al., 2009). En el primer caso es posible que se trate de las mismas muestras o de muestras conseguidas en las mismas poblaciones, ya que las referencias geográficas dadas por los autores son coincidentes. También es esencial tener en cuenta que el número de individuos para alguno de estos grupos es muy bajo (en algunos casos inferior a 20), por lo que podríamos estar viendo desviaciones debidas al bajo número de individuos incluidos en el muestreo.

La población Huichol presenta mayoritariamente haplogrupo A con un 55,6%, seguida en menor proporción por el B y el C (25% y 16,7% respectivamente), y un único individuo de haplogrupo D (2,8%), siendo parecida a lo ya reportado por Sandoval para este mismo grupo indígena (Sandoval et al., 2006), aunque no coincidente con

otros estudios (Kemp et al., 2010; Peñaloza-Espinosa et al., 2007), a pesar de que como ya se ha comentado, la procedencia de las muestras de dos de estas publicaciones parece ser la misma (Sandoval et al., 2006; Peñaloza-Espinosa et al., 2007). Tampoco hay concordancia con muestras de poblaciones Coras muy cercanas a la región Huichol (Kemp et al., 2010).

Respecto a los grupos Nahuas, hay gran cantidad de estudios realizados en poblaciones de distintas zonas, y aunque los porcentajes encontrados en los haplogrupos son parecidos a las poblaciones de Zitlala (Guerrero), Actopan (México), Xochimilco (México) y Necoxtla (Puebla) (Peñaloza-Espinosa et al., 2007), hay diferencias con otros estudios de muestras de distinta procedencia u otras poblaciones (Guardado-Estrada et al., 2009; Lorenz and Smith, 1996; Peñaloza-Espinosa et al., 2007; Sandoval et al., 2009). Esta gran diversidad estaría justificada por la historia biológica de este grupo procedente del norte, considerado descendiente de los antiguos Mexicas. Su amplia distribución en la zona central de México, sería consecuencia de su dispersión y conquista de la región. En algunos casos sin embargo, habría que considerar los posibles errores asociados a los bajos valores de *n*.

Los escasos estudios realizados en el grupo indígena Otomí no dejan mucho espacio para las comparaciones, aunque las proporciones entre los dos grupos incluidos en esta tesis son similares entre ellas y con las reportadas en sendos estudios de los años 2006 y 2007 (Peñaloza-Espinosa et al., 2007; Sandoval-Ramirez et al., 2006; Sandoval et al., 2009). Los grupos Otomíes estarían caracterizados por una mayoría de haplogrupo A, seguida de valores intermedios para B y C (aunque un poco mayor para C en las muestras incluidas en esta tesis), y residual para el haplogrupo D.

La población Tepehua es la última población de la región Huasteca incluida en este trabajo, y en ella se ha detectado un patrón parecido a los Nahuas de la misma región, es decir mayoría de A (62,3%), seguida de B (24,6%), y con mínima representación para los haplogrupos C y D (6,9% y 3,4% respectivamente). Estas semejanzas podrían estar indicando cierta relación con los grupos Nahuas debido a la cercanía geográfica en la sierra Otomí-tepehua. Este aspecto se detallará más adelante.

Por último la población Maya presenta mayoría de haplogrupo A (74,1%) con representación mínima del resto de haplogrupos (B 15,5%, C 6,9% y D 3,4%), y aunque con pequeñas diferencias es parecida a otras reportadas para esa misma

región (González-Olivier et al., 2001; Schurr et al., 1990; Sandoval et al., 2009; Torroni et al., 1992).

La mayor parte de los haplogrupos detectados en las poblaciones indígenas estudiadas pertenecen a los cuatro linajes más frecuentes en nativos americanos (A2, B2, C1 y D1). No se ha detectado ninguna variante del haplogrupo C (C4c), X (X3a) o D (D4h3) presentes en otras poblaciones estudiadas recientemente (Kemp et al., 2007; Kitchen et al., 2008; Malhi et al., 2010; Perego et al., 2009; Sandoval et al., 2009; Tamm et al., 2007). También se han localizado haplogrupos que podrían ser de origen europeo en la sierra Otomí-tepehua (tres H, un V y un T.), lo que demuestra la existencia de cierto grado de mestizaje en las poblaciones indígenas, a pesar de lo abrupto de la región y su supuesto aislamiento, así como el absoluto y estricto control que se ha utilizado para la selección de las muestras. Para el resto de grupos no se ha detectado mestizaje, a pesar de estar muchos de ellos ubicados en regiones más accesibles del altiplano mexicano. Este resultado es reflejo de la historia de la región Huasteca, que pudo ser zona de paso entre el centro del país y la costa del Atlántico. Este hecho es indicativo de que los donantes con haplogrupos de ADNmt europeos en vez de amerindios, estaban perfectamente integrados en las poblaciones indígenas, y eran originarios al menos en las últimas generaciones de la región muestreada, además de hablantes de lengua indígena en su gran mayoría. Es decir que el mestizaje se produjo en un periodo como mínimo anterior a tres generaciones, teniendo al menos 100 años de antigüedad, en caso contrario se habrían detectado al analizar los datos incluidos en las encuestas. De todas formas, gracias a las bases de datos poblacionales y a la integración de la información, la procedencia de los ancestros de los donantes, así como otros aspectos culturales, lingüísticos y familiares, podemos dar una explicación a la presencia de estos individuos.

De cualquier manera y aunque los modelos mencionados se cumplen en la mayoría de los casos, los análisis realizados en esta tesis dibujan un panorama más detallado para Mesoamérica. Para el haplogrupo A, la distribución es uniforme si trabajamos con diversas poblaciones de la región mesoamericana, y se detecta correlación entre las frecuencias de las ocho poblaciones incluidas en este trabajo y latitud y longitud. Para el haplogrupo B no se detecta correlación con ninguna de las componentes geográficas, pero en cambio para el haplogrupo C se detecta una correlación significativa con la latitud. Esto significa que la frecuencia de esta variante aumenta al desplazarnos hacia el sur. En los tres casos las distribuciones se ajustarían a lo descrito por Lalueza (Lalueza et al., 1997), pero no con lo aportado por otros estudios

(Guardado-Estrada et al., 2009; Malhi et al., 2003; Merriwether et al., 1995; Salzano 2002). Por último, la distribución de la variante D no presenta correlaciones al analizar las poblaciones incluidas en este trabajo, y su presencia en la región mesoamericana es distinta a lo descrito hasta ahora.

Como introducción a la descripción de los haplotipos definidos en cada población y su distribución en las distintas regiones del ADNmt estudiadas, es importante tener en cuenta el número de haplotipos detectados para cada uno de los haplogrupos, ya que es también una medida de la diversidad existente. Para el haplogrupo A se han detectado 77 haplotipos diferentes en 313 individuos de los 626 estudiados que presentan este haplogrupo, 40 en 134 individuos para el B, 24 en 123 muestras para el C y solamente 17 en 50 individuos para el D. Estos valores son un poco bajos si los comparamos con los detectados en otros estudios (Kemp et al., 2010), aunque debe tenerse en cuenta en este estudio estamos trabajando con poblaciones pertenecientes todas ellas, a excepción de una, a la región mesoamericana, mientras que los datos procedentes de la publicación incluyen numerosas poblaciones del norte del continente americano.

El número de haplotipos únicos que hallamos en cada una de las poblaciones es también indicativo de la diversidad de cada una de ellas. Al analizar estos valores entre las ocho poblaciones incluidas en esta tesis, los más altos son para población Nahua en todos los casos (D-Loop, HVI y HVII, valor absoluto/porcentaje) presentando 74/38,95%, 34/17,89% y 16/8,42% haplotipos únicos respectivamente, seguido de Otomíes de la Sierra (39/42,86%, 16/17,58%, 7/7,69%) y Mayos (31/56,36%, 14/25,45%, 10/18,18%). Las poblaciones con valores más bajos son Pimas (16/32,65%, 3/6,12%, 3/6,12%) y Tepehuas (18/33,96, 6/11,32%, 5/9,43%), lo que apunta una vez más lo complejo de estas poblaciones o el posible efecto del pequeño tamaño muestral. Con valores intermedios tenemos las poblaciones Otomíes del Valle (26/32,10%, 15/18,52%, 5/6,17%), Mayas (24/54,55%, 14/31,82%, 9/20,45%) y Huicholes (21/60%, 13/37,14%, 7/20%).

Los grupos indígenas con un menor número de haplotipos únicos en relación a haplotipos totales son los grupos son Otomíes del Valle, Pimas y Tepehuas, que a pesar de tener tamaños muestrales que están en la línea de las poblaciones Maya y Mayo, presentan únicamente un 32,10%, 32,65% y 33,96% de haplotipos no compartidos. El número de haplotipos únicos con bajas frecuencias podría estar indicando deriva genética y endogamia. Una población en expansión y de rápido

crecimiento demográfico, acumula variantes que no aparecen en otras poblaciones y que tampoco habrían tenido oportunidad de crecer en número dentro de la población. Estos bajos valores podrían estar indicando que estas poblaciones no solamente no han tenido una gran expansión en su historia biológica reciente, a pesar de lo indicado por otros resultados, sino que además no se han mestizado, lo que habría hecho aumentar el valor de estos haplotipos únicos (Fuselli et al., 2003). Sin embargo en poblaciones como Huicholes, Mayos y Mayas, estos valores son bastante elevados (60%, 56,36% y 54,55% respectivamente), indicando probablemente el efecto de una expansión poblacional y de procesos de mestizaje. Estos valores elevados muestran un sustrato genético bastante diverso, que podría estar reflejando cierto mestizaje e indirectamente antigüedad de las poblaciones. En el caso de la población Maya, la gran expansión de la que fue protagonista el imperio en la península del Yucatán y en toda la región al sur de esta, conllevó importantes movimientos de población y probablemente fenómenos de mestizaje asociados a la misma (Ruz 2006; Serrano Carreto et al., 2006). Es curioso que la población Otomí del Valle, a pesar de presentar un tamaño muestral similar al del grupo de la Sierra, tenga un número de haplotipos únicos bastante inferior (32,10% frente a 42,86%).

Si centramos nuestra atención en el análisis para la región HVI, y considerando que acumula la mayor parte de diversidad de la región control, hay que comentar que presenta únicamente 148 haplotipos distintos en 620 secuencias, lo que representa solamente un 23,87%, casi la mitad que con D-Loop. Para esta región el número de haplotipos compartidos entre los grupos es mayor, pero sigue reflejando la tendencia observada con región control completa. La población Pima comparte una media de dos o tres haplotipos con el resto de poblaciones, y en el caso de Mayos, Huicholes y Mayas, la media aumenta a tres o cuatro. Sin embargo, hay una clara semejanza entre todas las poblaciones del centro mesoamericano; Nahuas comparten trece haplotipos con Otomíes del Valle, catorce con los de la Sierra y ocho con Tepehuas. Es importante destacar los nueve que comparte con el grupo indígena Maya por la posible relación entre ambos apuntada en algunos trabajos (Martínez-Cortés et al., 2010; Solanes Carraro and Vela Ramírez, 2005; Valle Esquivel and Hernández Alvarado, 2006). Otomíes del Valle comparten también catorce haplotipos con los de la Sierra y seis con Tepehuas. Y finalmente Tepehuas, comparten siete haplotipos con Otomíes de la Sierra.

En referencia a los haplotipos definidos por toda la región D-Loop los valores de coincidencia son más bajos. Entre Pimas y Mayos no se comparte ningún haplotipo,

mientras que la primera población comparte uno con Nahuas, Otomíes del Valle y Tepehuas; la población Mayo comparte a su vez dos haplotipos con Huicholes debido probablemente a su cercanía geográfica, y otros dos con Nahuas de la Huasteca y solamente uno con Otomíes de la Sierra. La población Huichol presenta tres haplotipos compartidos con Nahuas, y uno con cada una de las siguientes poblaciones: Otomíes del Valle, Otomíes de la Sierra y Tepehuas. La población Maya comparte dos haplotipos con Nahuas, y uno con Otomíes del Valle y con cada una de las otras poblaciones de la Huasteca (Otomíes de la Sierra y Tepehuas). Y es precisamente en esta región Huasteca, dónde hallamos los valores más elevados para haplotipos compartidos, que podrían estar reflejando la cercanía geográfica con Otomíes del Valle y la llamada Sierra Otomí-tepehua, y poniendo de manifiesto la relación genética que existe entre todas ellas. Nahuas comparten ocho haplotipos con Otomíes del Valle, siete con Otomíes de la Sierra y cinco con Tepehuas. Otomíes del Valle comparten también siete haplotipos con Otomíes de la Sierra y tres con Tepehuas; y finalmente Tepehuas, comparten un único haplotipo con Otomíes de la Sierra a pesar de su cercanía.

El análisis de HVII no sigue el patrón comentado en el párrafo anterior y el número de haplotipos compartidos disminuye a una media de cuatro, presentando un mínimo de dos y un máximo de siete entre los distintos grupos.

La historia biológica de los grupos Nahuas es primordial para comprender lo que estamos viendo, si nos centramos en los haplotipos compartidos entre poblaciones. La procedencia de esta población y su actual localización, explican en parte esta elevada diversidad y semejanza con el resto de grupos. Con todos ellos comparten algún haplotipo, principalmente con las poblaciones con las que conviven en la sierra Otomí-tepehua, pero también con aquellas con las que tuvieron algún tipo de relación. Este último caso sería el de las poblaciones Mayo y Huichol, con las que probablemente tuvieron contacto en su viaje desde el mítico Aztlán del que procederían, situado en el norte del país, hasta su actual ubicación en el altiplano mexicano. Precisamente su presencia en esta región y la convivencia con los grupos de esta zona, habría facilitado el intercambio de cierta carga genética y de ahí, la homogeneidad observada entre los grupos de la Sierra. No hemos de perder de vista que podría tratarse también de un sustrato ancestral, compartido por los grupos que conviven en esta región.

Los *network* construidos para cada uno de los cuatro haplogrupos representan gráficamente los haplotipos compartidos entre las distintas poblaciones, siendo el

tamaño de los nodos proporcional al número de secuencias. La morfología que presentan es estrellada y en algunos casos muy ramificada, presentando 34 mutaciones en el *network* del haplogrupo A2 y 35 en el del B2. Para cada uno de ellos además, encontramos un nodo principal y otro secundario unido por una única mutación en cada caso. Ambas redes recogen la mayor parte de la diversidad y reflejan nuevamente cierta homogeneidad entre los grupos del altiplano central, sobre todo entre las poblaciones de la Huasteca, reflejado por la cercanía entre Otomíes, Tepehuas y Nahuas, que son las poblaciones a las que pertenecen los haplotipos integrados en los nodos principales.

En el *network* del haplogrupo A2 las mutaciones entre los nodos principales están situadas en las posiciones 16.111 entre el primer y segundo nodo, y la 16.092 entre el primero y el tercero. La coalescencia estimada es de 8.408 y 5.223 años respectivamente. Es llamativa la presencia en los dos nodos principales de haplotipos Nahuas, Otomíes y Tepehuas. El principal se diferencia del secundario en que en el primero hay una presencia importante de población Maya y ausencia del resto, mientras que en el otro hay una importante presencia de población Huichol y ausencia del resto. El haplogrupo A parece indicar que de un sustrato genético ancestral compuesto por haplotipos detectados en las actuales poblaciones Nahua, Otomí y Tepehua, se habría disgregado un grupo del que procederían los actuales Huicholes. Posteriormente, ese mismo sustrato inicial se habría disgregado en los actuales Mayas. Por este motivo Mayas y Huicholes no comparten ningún haplotipo en estos nodos. Esta separación se habría producido como máximo hace unos 8.408 años. Los tres haplotipos para HVI compartidos entre ambos grupos (el haplotipo 2, el 32 y el 42; Tabla VI.6 del capítulo VII. Anexos), se comparten también con el resto de poblaciones incluidas en esta tesis y están presentes en otras poblaciones de México y Estados Unidos (datos no publicados obtenidos al volcar los haplotipos en Mitosearch).

Para el haplogrupo B2, el nodo central está formado por haplotipos actualmente presentes en todas las poblaciones estudiadas excepto Tepehuas y Pimas. Estas dos poblaciones presentan nodos derivados y no coincidentes entre ellos. En primer lugar, y con una coalescencia estimada de 9.459 años, el nodo derivado a través de la mutación 16.295 es exclusivo de Nahuas, Otomíes de la Sierra y Tepehuas, probablemente fruto del mestizaje en la Sierra Otomí-tepehua por la cercanía geográfica de los grupos (Figura II.2). En segundo lugar y con una coalescencia semejante, otro nodo derivado caracterizado por el polimorfismo 16.182, y característico por su presencia en población Nahua, Maya y Pima. Derivados de este

nodo secundario, surgen otros presentes exclusivamente en población Nahua y Mayo. Estos resultados indican nuevamente un sustrato compartido entre la práctica totalidad de las poblaciones estudiadas, a excepción de la población Pima que parece tener un sustrato genético característico. Es llamativa la amplia disgregación de los haplotipos presentes en la población Mayo.

El *network* del haplogrupo C1 es claramente estrellado con un nodo central en el que se encuentran individuos de casi todas las poblaciones. Es especialmente interesante el nodo secundario caracterizado por el polimorfismo 16.295, con una coalescencia estimada de 7.866 años y que forma una agrupación exclusiva de población Pima. Estos resultados están en concordancia con lo visto hasta ahora para este grupo, apoyando una historia relativamente distinta al resto de poblaciones incluidas en esta tesis.

Por último, el *network* del haplogrupo D1 es el menos complejo, ya que este es el menos representado en las poblaciones estudiadas. Se observa un nodo principal en el que están presentes todas las poblaciones excepto Huicholes, Pimas y Mayas. Es destacable el nodo secundario formado por un haplotipo presente exclusivamente en la población Mayo a través de las mutaciones 16.142, 16.292 y 16.305. A partir del nodo central y a través del polimorfismo 16.274, aparece otro nodo secundario exclusivamente Otomí y del que a su vez, aparece un haplotipo presente en población Nahua por la mutación 16.093. Esta última información apoyaría la idea de nahuatlización de las poblaciones de la Sierra.

En los cuatro casos se puede considerar que los diagramas tienen un patrón en forma de estrella que indicaría una expansión poblacional, como ya han reportado otros estudios en amerindios (Achilli et al., 2008; Guardado-Estrada et al., 2009; Harpending et al., 1998). En los extremos de las ramas y caracterizados por los nodos más pequeños se situarían los haplotipos únicos para cada grupo indígena. Esta morfología reticulada es concordante con las distribuciones *pairwise differences* para cada uno de los haplogrupos que presentan una gran uniformidad en sus perfiles unimodales y centrados entre 2 y 3 diferencias, lo que indica también una expansión poblacional (Figuras VI. 27 a VI.30 en el capítulo VII. Anexos), (Harpending et al., 1998; Mishmar et al., 2003; Di Rienzo and Wilson, 1991; Rogers and Harpending, 1992; Schneider and Excoffier, 1999).

El análisis de los *network* nos da una idea de la complejidad de la historia de esta región mesoamericana, aunque en todos ellos se observa un modelo general en el que

el sustrato genético basal es compartido. Se detectan fenómenos de diferenciación poblacional, como el caso de Huicholes y Mayas detectado para el haplogrupo A2, prueba del elevado fraccionamiento poblacional, producto de la colonización de hábitats específicos. La colonización de microrregiones habría facilitado la fragmentación poblacional así como su diversificación. La elevada heterogeneidad cultural que estaría en concordancia con esta hipótesis, lo que implicaría la diferenciación genética (Wang et al., 2008). Sobre este panorama se habrían producido movimientos poblacionales posteriores, absorciones y aculturizaciones, sobre todo por parte del grupo Nahua, lo que justificaría su alta heterogeneidad genética y el elevado número de haplotipos compartidos con el resto de poblaciones incluidas en esta tesis.

Un caso excepcional en este escenario lo constituye la población indígena Pima, que presenta una interesante configuración de sus individuos en cada uno de los *network*. En el del haplogrupo A2, presenta un único individuo que deriva del nodo principal y por lo tanto procede de la misma diversidad global compartida con la mayoría de individuos. Para el del haplogrupo B2, los siete individuos de la población que presentan este haplogrupo comparten haplotipo, y se sitúan en el nodo secundario. Para el *network* del haplogrupo C1, gran parte de los individuos están representados en el nodo principal o a una mutación de este. Esto estaría indicando en primer lugar, que esta población comparte el sustrato genético con el resto de poblaciones. En segundo lugar, que su historia demográfica ha sido ligeramente distinta al resto de poblaciones estudiadas, caracterizándose por una mayor estabilidad poblacional, sin las expansiones características del resto de grupos. La estrategia adoptada por los grupos del norte y fenómenos como la deriva genética o la endogamia asociada a esta, tendrían como consecuencia los resultados que observamos. Todo esto estaría en concordancia con lo ya comentado para este grupo y la historia reciente del mismo.

IV.4. Análisis por grupos indígenas

Habitualmente los estudios de ADNmt se centran exclusivamente en el análisis de HVI, por los valores más elevados de diversidad que presenta, y por considerarse lo suficientemente discriminante en cuanto a asignación de haplogrupos. Sin embargo, en los últimos años se ha tendido a secuenciar la región D-Loop completa para este tipo de estudios (Álvarez-Iglesias et al., 2007; Kemp et al., 2010; Sandoval et al., 2009), ya que dos ADNmt que compartan el mismo haplotipo para HVI no

necesariamente son iguales para HVII ni comparten el mismo haplogrupo (Álvarez-Iglesias et al., 2007). En este trabajo y por lo exhaustivo del análisis que se quería realizar, se ha secuenciado toda la región D-Loop y se han hecho los análisis de forma independiente para cada una de las regiones hipervariables y la región control completa. Los resultados muestran que a pesar de que los análisis con región D-Loop completa y HVI están en concordancia, los datos obtenidos para una región mayor permiten interpretar los resultados en base a una mayor información, siendo conveniente el estudio de la D-Loop completa frente a la región HVI sola o en combinación con HVII (Salas et al., 2005, 2007).

Es importante tener en cuenta que el número de haplotipos es variable en función de la región de ADN secuenciada, por lo que los valores de diversidad en este sentido son obviamente distintos. Para D-Loop el número de haplotipos diferentes en 605 secuencias es de 278, lo que representa prácticamente el 46%, valor mucho más elevado que el obtenido en otro trabajo que incluye 689 secuencias y 323 haplotipos distintos, representando alrededor del 30% (Kemp et al., 2010). Debemos tener en cuenta que en este último estudio muchas de las muestras están localizadas geográficamente fuera de Mesoamérica, y por lo tanto no habrían estado inmersas en el escenario de expansión poblacional que tuvo lugar en esta área. Los valores para estos porcentajes sin embargo, se reducen a más de la mitad cuando evaluamos las regiones hipervariables por separado, siendo casi el 24% para HVI y cercano al 13% para HVII.

En lo referente a análisis de diversidad y teniendo en cuenta únicamente los datos de haplotipos amerindios para las poblaciones, los índices calculados son llamativos por lo elevado de algunos de los parámetros como la diversidad haplotípica H , que presenta valores más elevados para la región D-Loop y parecidos para HVI, pero significativamente menores para HVII. El número de sitios polimórficos (S) y el número medio de nucleótidos diferentes (M), presentan valores elevados para D-Loop pero menores para las regiones hipervariables. La representación gráfica del número de sitios polimórficos de la región control dibuja claramente esta distribución (Figura III.6), indicando una mayor variabilidad en HVI. La diversidad nucleotídica π , sin embargo tiene valores bastante parecidos tanto para D-Loop como para las regiones hipervariables. Frente a este panorama y vistos los valores en las distintas poblaciones, como ya se ha comentado, es preferible el análisis de la región D-Loop, siempre que se pueda, frente a cualquiera de las otras por la mayor diversidad reflejada en ella (Álvarez-Iglesias et al., 2007; Salas et al., 2005, 2007).

Al revisar de forma global estos índices para las ocho poblaciones estudiadas, se observan los valores más elevados en las poblaciones localizadas en el centro de la región mesoamericana, quedando desmarcadas con valores inferiores las poblaciones Pima y Maya, además de la Tepehua. Esto podría ser debido a las diferencias entre las áreas y a su distancia a la región central mesoamericana, donde tuvieron lugar los grandes movimientos de población y el encuentro de los diferentes grupos y culturas. Valores más elevados de diversidad para Mesoamérica y menores para las regiones al norte y al sur, han sido también reportados por otros autores (Salas et al., 2009). En el caso de la población Maya debe tenerse en cuenta que está en una región con una historia biológica distinta del resto, aunque tuvo igualmente grandes expansiones poblacionales y movimientos relacionados con el intercambio comercial durante el esplendor de este imperio entre el 2.000 AAP y el 1.546 ADP. Sin embargo, es importante no perder de vista que los distintos conflictos y la llegada de poblaciones del altiplano ocasionaron su declive (Ruz 2006; Solanes Carraro and Vela Ramírez, 2005). Para el grupo Pima es importante considerar que es una población procedente originariamente del oeste de Canadá, por lo que su composición genética sería ligeramente diferente a las poblaciones del centro (Hope 2006; Muñoz Orozco s/f). Igualmente no hay que perder de vista que el número total de hablantes de lengua Pima es inferior a mil individuos, por lo que la diversidad poblacional podría verse reducida por un proceso endogámico asociado a deriva genética o efecto fundador. Por último el grupo indígena Tepehua presenta también valores bajos para la diversidad haplotípica, lo que podría estar poniendo de manifiesto el número reducido de individuos del grupo (por debajo de 10.000 para todo el país), y/o el aislamiento de estas poblaciones en la sierra Otomí-tepehua, que podría haber desembocado en un proceso de deriva genética. De cualquier manera estas poblaciones procederían de grupos del noreste de los que podrían haberse disgregado, quedando aislados en su actual ubicación (Williams García 1963). Esto estaría de acuerdo a valores de diversidad para otras poblaciones incluidas en otros estudios, aunque con diferencias para grupos indígenas con bajos valores muestrales (Sandoval et al., 2009). Hay que tener en cuenta la posibilidad de que estas desviaciones sean debidas también en el presente estudio, a estos bajos valores del número de muestras.

Los test de diferenciación poblacional junto con las distancias genéticas F_{st} representadas en el espacio a través de MDS, dan resultados concordantes con lo explicado anteriormente, apoyando la homogeneidad de la región mesoamericana central y pudiendo esbozarse incluso el mapa de México sobre la distribución de las

poblaciones. Esta configuración general incluiría un grupo de poblaciones indígenas procedentes del altiplano mexicano, de la Huasteca y del occidente de México. Ligeramente diferenciadas se encontrarían las poblaciones Mayas del sudeste, aunque las poblaciones situadas en el norte son las que se diferencian mucho más del resto, concretamente Mayos y sobre todo Pimas, esta última siempre con valores elevados en los test de diferenciación y posiciones extremas en el MDS respecto al resto. La cercanía del grupo Huichol con las poblaciones del altiplano central corroboraría lo mostrado por el *network* del haplogrupo A2, que apuntaba una sustrato genético compartido entre todas las poblaciones.

Las dos regiones geográficas consideradas en esta tesis, Mesoamérica y Oasisamérica, están bien definidas geográfica y culturalmente. Seis poblaciones se distribuyen en la primera, mientras que la única población representante de Oasisamérica es el grupo indígena Pimas. Dentro de Mesoamérica los estudios históricos, etnohistóricos y arqueológicos permiten diferenciar áreas naturales y culturales muy bien definidas, y el comportamiento demográfico de las poblaciones puede estar marcado por las diferentes estrategias de subsistencia que se desarrollaron en cada una ellas. Igualmente deben considerarse otros aspectos como el desarrollo agrícola y tecnológico, las relaciones culturales y comerciales, e incluso la historia política precolombina que tuvo lugar.

Este modelo de supervivencia diferente entre las dos grandes regiones pudo afectar a la estructura y composición genética de los grupos aborígenes en cada una de las zonas. Aunque también es correcto plantearse estudios intrarregionales, sobre todo en Mesoamérica. Es posible que las diferencias geográficas, históricas, ambientales, económicas y culturales dentro de una misma área, pudieran ser lo suficientemente significativas como para determinar modelos genéticos diferentes. Por ejemplo, una comparación superficial entre la región Maya y la Huasteca, pone de manifiesto que a pesar de la aparente homogeneidad cultural existente por la pertenencia de ambas a Mesoamérica, las diferencias pueden ser importantes.

La región Maya es una región selvática asentada en planicies que favorecieron el comercio y la agricultura, mientras que la Huasteca está inmersa en la Sierra Madre Oriental, una de las regiones montañosas más abruptas de toda Mesoamérica. A pesar de esta ubicación y la complejidad orográfica que presenta, fue el camino natural entre el centro del altiplano mexicano, representado por Tenochtitlan, y su salida natural al mar hacia el Golfo de México, Veracruz. La Huasteca además, pudo

representar un refugio poblacional importante, donde gran número de poblaciones se habrían resguardado ante la expansión de culturas del altiplano, como los Aztecas inicialmente y la presión de los colonizadores europeos después. Sea como fuere, un cometido del presente estudio es determinar cómo ha influido esta riqueza geográfica, biológica, cultural e histórica, sobre la conformación genética de las poblaciones que las habitaron.

Por último, y en cuanto a la distribución *pairwise differences*, indica una expansión poblacional cuando analizamos D-Loop, lo que estaría en consonancia con los *network* para los distintos haplogrupos, así como con el exceso de polimorfismos en baja frecuencia y de ahí el elevado número de haplotipos (Harpending et al., 1998; Mishmar et al. 2003; Di Rienzo and Wilson, 1991; Rogers and Harpending, 1992; Schneider and Excoffier, 1999). El perfil unimodal de estas distribuciones sin embargo podría estar mostrando mestizaje o cuellos de botella (Babalini et al., 2005; Brazek et al., 2001; Rogers and Harpending, 1992). Quedaría patente de cualquier manera un deficiente efecto de la selección purificadora, a pesar de que el test de Tajima (D) es significativo para el análisis de esta región, lo que indicaría una ruptura con la neutralidad genética. El estudio de estos mismos indicadores para HVI y HVII, esbozan un escenario distinto, indicando expansión poblacional para ambas, aunque sin presión selectiva para HVI y sí para HVII.

Las AMOVAs entre los ocho grupos estudiados se estructuraron en seis categorías diferentes para tratar de establecer las relaciones entre los grupos frente a distintos criterios. Estas categorías se hicieron según varias agrupaciones. En primer lugar con un criterio geográfico, con un resultado negativo y hasta cierto punto lógico, ya que muchas veces las agrupaciones por distancia no coinciden con la realidad orográfica de la zona, hallándose barreras entre poblaciones muy cercanas (Barbujani and Goldstein, 2004; Simoni et al., 2000b). En segundo lugar, los troncos lingüísticos representados por los ocho grupos en base a la clasificación atendiendo a los criterios del INALI (INALI 2007, 2008), tampoco ha resultado positiva en el análisis del ADNmt, encontrándose en la misma línea que otros trabajos realizados en el continente (Kemp et al., 2010; Sandoval et al., 2009). La clasificación por regiones indígenas, se ha construido con los criterios de una de las instituciones con mayor trayectoria en el campo indigenista, así como en el estudio de sus culturas y su historia, como es el CoNaDePI. Esta institución propone áreas de distribución e influencia de las distintas regiones, cuyo resultado tampoco es significativo. En penúltimo lugar, la historia de las poblaciones. Para algunos autores (Solanes Carraro and Vela Ramírez, 2005), los

pueblos indígenas actuales son los herederos, intelectuales y biológicos, de las grandes culturas mesoamericanas que se desarrollaron, y por este motivo se ha creado una clasificación que relaciona cada uno de los grupos indígenas con un pasado cultural de gran influencia en la historia de la región. Los grupos Nahuas, por ejemplo, serían los descendientes de los Aztecas (Solanes Carraro and Vela Ramírez, 2005; Valle Esquivel 2006), o las poblaciones Huicholes de la cultura Tarasca del occidente mexicano (Neurath 2003; Solanes Carraro and Vela Ramírez, 2005). Los resultados obtenidos tampoco han sido significativos. Ninguna de estas agrupaciones tuvo un valor de p superior a 0,05, por lo que podemos concluir que ninguna de ellas ha generado unas diferencias lo suficientemente importantes entre los grupos para que sean significativas.

Por último, las AMOVAs por áreas culturales se han trabajado en base a dos agrupaciones distintas variando los grupos incluidos en cada categoría. Ambas son significativas y en base a ellas podemos descartar, por los resultados obtenidos en las categorías mencionadas anteriormente, que estas diferencias se deban a distancia geográfica, lengua indígena, región indígena o historia de las poblaciones. La significación de las AMOVAs nos está indicando que sin variar los grupos de la región “Norte”, “Occidente” y “Sur”, modificando los grupos incluidos en “Costa” y “Centro”, las diferencias son significativas. Esta agrupación no ha sido descrita hasta el momento, aunque es probablemente la más intuitiva, ya que está basada en las distintas grandes civilizaciones y culturas que dominaron cada una de estas regiones.

La agrupación “Norte” estaría representada únicamente por el grupo Pima, y englobaría grupos cazadores-recolectores que no tuvieron éxito al tratar de adaptar una estrategia sedentaria a causa de las condiciones climáticas de la región. Esta zona ocuparía una franja continental complementaria en una parte con el área considerada “Occidente”, que forma otra de las agrupaciones considerada en las AMOVAs. La categoría “Occidente” incluiría las poblaciones Mayo y Huichol, centrándose en una franja de territorio limitada por la costa del Pacífico. Esta zona estaría caracterizada por las relaciones con otras regiones, sobre todo gracias al intercambio marítimo de algunos productos, o fenómenos de migración costera como consecuencia de este comercio, y al desarrollo de la urbe de Teuchitlán, que dominó el área durante parte del período Clásico. En la región “Sur”, quedaría englobada la población Maya, y se correspondería con la zona centrada en la península de Yucatán, en la que surgieron gran cantidad de ciudades de gran tamaño y complejidad.

Las tres agrupaciones anteriores se han mantenido para las dos AMOVAs de Áreas Culturales, y las categorías “Centro” y “Costa” se han variado para comprender mejor las relaciones en estas zonas.

La región “Centro” engloba una zona de gran riqueza que propició el surgimiento de grandes culturas, por ejemplo durante el Clásico, como Cuicuilco, Cholula, que construyó los mayores monumentos de toda la región central de México, y Teotihuacán que llegó a tener gran influencia en toda Mesoamérica. Esta región fue escenario de numerosos movimientos migratorios, por lo que las manifestaciones culturales fueron compartidas entre las culturas ubicadas en esta zona. Los Mexicas, procedentes del norte y antecesores de los Nahuas (Solanes Carraro and Vela Ramírez, 2005; Valle Esquivel 2006), dominaron la región en el Posclásico, lo que acabó dando lugar a la fundación de Tenochtitlan. En esta categoría y para “Áreas Culturales I”, se incluye a Otomíes del Valle, y para “Áreas Culturales II”, a Otomíes del Valle, de la Sierra y a la población Nahua.

La región “Costa” consideraría el golfo de México, englobando la vertiente atlántica de la Sierra Otomí-tepehua hasta el litoral. Esta zona estuvo poblada por grupos nómadas muy eficaces que explotaban todos los recursos, tanto recolección como caza y pesca, y que adoptaron finalmente estrategias sedentarias, lo que conllevaría el surgimiento de grandes culturas como la Huasteca y la Olmeca durante el Preclásico, y la Totonaca durante el Posclásico. A la que se cree que fue la capital de esta última civilización, Cempoala, llegaban los Mexicas para obtener recursos de la zona y fue la primera gran ciudad que encontraron los españoles a su llegada al continente americano. En esta categoría y para “Áreas Culturales I”, se incluye a Otomíes de la Sierra, Tepehuas y Nahuas, y para “Áreas Culturales II”, únicamente a Tepehuas.

Como se puede observar en las agrupaciones hechas para las AMOVAs (Tabla III.44), el grupo indígena Otomíes del Valle se ubica siempre en el grupo “Centro”, mientras que el grupo “Costa” siempre contiene al menos al grupo Tepehua. Por lo tanto, son las poblaciones Nahua y Otomíes de la Sierra las que basculan entre ambas agrupaciones. La razón para esta distribución es sencilla para las poblaciones fijas, puesto que tanto Otomíes del Valle como Tepehuas pertenecen claramente a las categorías mencionadas, mientras que en los otros casos son discutibles. Así, en la categoría “Áreas Culturales” se ha considerado la posibilidad de que todas las poblaciones de la Huasteca pertenezcan a la región del golfo. Como segunda alternativa, la clasificación de las poblaciones distribuidas hacia el poniente como del

“Centro”, y las situadas hacia el levante, abiertas hacia el golfo de México y Veracruz, como de la “Costa”, representada también por la población Tepehua. En primer lugar, la población Nahuatl ubicada en la Huasteca, ha sido ampliamente cuestionada en cuanto a su pertenencia a este grupo, apuntando algunos autores su relación con el grupo indígena Téenek (o Huasteco) de ascendencia Maya. Este grupo sin embargo tiene una localización geográfica desviada con respecto al resto de grupos, encontrándose ampliamente distribuido en la zona entre la sierra y el litoral (Martínez-Cortés et al., 2010; Solanes Carraro and Vela Ramírez, 2005; Valle Esquivel and Hernández Alvarado, 2006). Por otro lado, el grupo Otomí de la Sierra, tiene una ubicación relativamente cercana a la comunidad indígena Tepehua, por lo que una posible relación genética por cercanía o afinidad cultural no es descartable. De cualquier manera, y puesto que ambas posibilidades son significativas, no puede excluirse la opción de que los Nahuas sean genéticamente Téenek y estén situados en el límite montañoso de su distribución. Igualmente con el grupo Otomí de la Sierra, habría que considerar que tanto la posibilidad de una mayor cercanía con el otro grupo Otomí en el valle del Mezquital, como cierto mestizaje con la franja costera, y una semejanza considerable a ese *pool* genético junto con Tepehuas y Nahuas, sean factibles vistos los resultados de la AMOVA.

La región Huasteca es una región con unas características geográficas únicas y con un carácter cultural excepcional en el entorno que la rodea. Representa a la vez una barrera natural entre el altiplano central y el golfo, pero también ha constituido un paso ancestral, al ser el camino más factible entre el Atlántico y el centro del continente. Ha sido refugio de comunidades marginales sometidas a la presión demográfica, militar y cultural de otras culturas en expansión (Solanes Carraro and Vela Ramírez, 2005; Valle Esquivel 2006; Valle Esquivel and Hernández Alvarado, 2006). Por este motivo, y dado que los especialistas no se ponen de acuerdo en su clasificación, en estos análisis se ha intentado desvelar parte de esta incertidumbre.

IV.5. Otomíes del Valle vs Otomíes de la Sierra

Se han estudiado dos poblaciones Otomíes separadas por unos 50Km, y ubicadas a ambos lados de la Sierra Madre Oriental. Una de ellas comprende varias comunidades del Valle del Mezquital, en el municipio de El Cardonal (Estado de Hidalgo), con una población hablante de lengua indígena de 8.862 individuos. La otra agrupa varias comunidades de la llamada sierra Otomí-tepehua, en los municipios de Tenango de

Doria y San Bartolo Tutotepec (Estado de Hidalgo), con un total de 11.498 hablantes de lengua indígena Otomí. La distribución de esta segunda localización geográfica es colindante con el sur de la Huasteca, estando compartida con dos grupos incluidos en esta tesis, Tepehuas y Nahuas, y dos que no están incluidos, Totonacos y Téenek. Las dos poblaciones de Otomíes hablan la misma lengua, aunque presentan variantes dialectales asociadas a los lugares geográficos en que se ubican (INALI 2007, 2008).

La comparación de los índices de diversidad tanto si consideramos D-Loop, HVI y HVII al separar en los dos grupos, o en los tres municipios las muestras en base a su procedencia, muestran valores más elevados para el grupo y municipios de la Sierra. Esto podría explicarse por la convivencia de estos grupos con otras poblaciones indígenas, que probablemente ha producido cierto mestizaje con el consecuente aumento de la diversidad. Las comunidades Otomíes de la Sierra se localizan geográficamente cercanas a Huehuetla, la comunidad indígena Tepehua, y también a las comunidades Nahuas incluidas en este estudio (Figura II.2). Sin embargo, es importante tener en cuenta que las comunidades Otomíes del Valle, están más aisladas en la zona de Ixmiquilpan (Figura II.2). El test de Tajima no es significativo para los grupos ni para ningún municipio. Ambas poblaciones así como los tres municipios considerados presentan distribuciones *mismatch* con perfiles multimodales, lo que estaría indicando que son poblaciones en expansión. Los valores para F_{st} y test de diferenciación son significativos entre los tres municipios mediante el análisis con D-Loop. Con las regiones hipervariables los resultados muestran algunas variaciones.

La composición de haplogrupos amerindios entre ambos grupos es muy similar, por lo que a pesar de las diferencias que encontramos para otros indicadores, este resultado estaría reflejando un origen común, lo que está apoyado por otros autores incluso dando una fecha estimada de separación de 700 AAP (Ochoa 1989). Es importante comentar que en las poblaciones de la sierra se detectaron dos haplogrupos europeos, por lo que estas poblaciones no estuvieron tan aisladas como se pensó al planificar el muestreo. De cualquier manera, está relativamente cercana a Cempoala (Estado de Veracruz), que fue la primera gran ciudad que encontraron los españoles a su llegada al continente americano, lo que podría explicar estas variedades. Estos resultados apoyarían lo ya reportado por otros autores en poblaciones de esta región, que detectaron un haplogrupo L en un individuo de una población Nahua mestizada del golfo de México (Aguirre-Beltrán 1989; Lisker 1981; Lisker et al., 1996; Peñaloza-Espinosa et al., 2007; Zavala et al., 1982).

A pesar de la lengua común y el territorio adyacente entre ambos, no constituyen argumentos suficientes para los miembros de cada grupo para aceptar su cercanía como etnia, recurriendo a elementos subjetivos para establecer una separación de otros hablantes de la lengua (Collin Harguindeguy 2006; Ramsay 2003). Los patrones de distribución de haplogrupos en las comunidades presentan muchas variaciones, y las diversidades para las comunidades muestran que las más cercanas a los otros grupos con los que conviven en la Huasteca, El Deca y Santiago, son más elevadas. A pesar de eso, la representación gráfica mediante el MDS de las comunidades, muestra una agrupación clara de cada una de ellas en dos zonas que se corresponden con los Otomíes del Valle y de la Sierra. Esta representación gráfica debe valorarse con cautela por el elevado valor de *stress* que presenta (Sturrock and Rocha, 2000). Los test de diferenciación entre comunidades no son significativos en el caso de los Otomíes del Valle, pero sí entre la mayoría de las de la Sierra. Tampoco las AMOVAs realizadas según criterios administrativos y geográficos son significativas, lo que nos está indicando que si bien detectamos ciertas diferencias entre las comunidades, estas no son lo suficientemente elevadas para determinar subestructura genética.

Las crónicas históricas y la información etnográfica señalan esta subdivisión, pero la duda a resolver es si se trata realmente del mismo grupo o tienen orígenes distintos y han adoptado la misma cultura. Para iniciar esta discusión, en primer lugar comentar que el propio nombre “Otomí” tiene origen náhuatl (Hekking 1995), su origen aproximado está datado por algunos autores entre los 4.000 y los 1.000 AAP (Moreno Alcántara et al., 2006). La historia del grupo Otomí es una historia de movimientos de población que se inicia con su llegada al altiplano central de México, dispersándose por la cuenca de México entre 1.050 y 1.250 ADP simultáneamente a Nahuas y Chichimecas. Estos grupos recorrieron el límite norte de Mesoamérica, y son considerados por algunos autores, los pobladores originarios antes incluso de la llegada de los hablantes de náhuatl (Carrasco 1950; Collin Harguindeguy 2006; Weitlaner 1941). La disgregación y los desplazamientos debido a conflictos armados con la Triple Alianza, que en el momento de su llegada dominaba el centro mesoamericano, habría poblado el valle y habría dado lugar a los actuales grupos Otomíes y Nahuas que mantendrían todavía paralelismos culturales debido a ese origen compartido (Carrasco 1950; Collin Harguindeguy 2006), o debido a la gran afinidad cultural con los grupos con los que han compartido territorio (Barrientos López 2006; Questa Rebolledo and Utrilla Sarmiento, 2006). De hecho, algunos trabajos incluso defienden el origen Chichimeca de estos grupos, proponiendo un bilingüismo

otomí/náhuatl de los mismos y una posterior disgregación (Collin Harguindeguy 2006; Cook de Lenhard 1981). Entre el siglo IX y XII los pueblos Nahuas habrían desplazado a las poblaciones del Valle hacia la Huasteca, lo que las habría hecho entrar en contacto con grupos Tepehuas y Téenek, reduciendo su territorio y provocando la mezcla de las mismas. Esta teoría cuadraría con lo que los resultados están reflejando, por ejemplo en los nodos principales de los *network* de todos los haplogrupos, se agrupan haplotipos compartidos entre Otomíes del Valle y de la Sierra. El número de haplotipos totales para HVI compartidos entre ellos (siendo cinco exclusivos Otomíes) es uno de los dos valores más elevados que hallamos entre los grupos estudiados en esta tesis, siete de los cuales están distribuidos por todo el continente americano, en localizaciones como California o Texas en Estados Unidos, o Puerto Rico, Venezuela, Brasil, El Salvador, Nicaragua, Guatemala, Colombia, Honduras, Chile, Perú o Cuba (datos no publicados y detectados por comparación con la herramienta Mitosearch). Además cuando analizamos el número de haplotipos compartidos definidos por D-Loop, aunque este número desciende a siete, cuatro siguen siendo exclusivos de grupos Otomíes, lo que sigue reflejando el sustrato biológico compartido entre ellos. Ambos resultados nos muestran de forma indirecta el pasado de estos haplotipos compartidos con otras poblaciones americanas, y consecuentemente la antigüedad de los grupos estudiados.

La explicación alternativa a la de las sucesivas oleadas migratorias, propone un asentamiento original en los valles de México y Toluca, desde donde partieron grupos de hablantes de Otomí en busca de las zonas de frontera para huir del sometimiento de los grupos conquistadores. En ese caso las diferencias entre los asentamientos Otomíes remitirían a una degradación cultural por el aislamiento de los centros de cultura y el contacto prolongado con grupos de menor desarrollo. Efectivamente, los momentos de auge para grupos Otomíes, tuvieron como centro el altiplano. El llamado reino de Xaltocan por ejemplo, logró una amplia extensión, cuya posterior fractura aisló parte del territorio y propició la decadencia de sus áreas periféricas.

La primera hipótesis expuesta es la más aceptada, pero debe tenerse en cuenta que hay dos posibles orígenes para las poblaciones Otomíes. El primero de ellos serían los grupos Chichimecas del norte de México, como ya se ha comentado, habrían sido cazadores-recolectores que al conquistar la zona central de Mesoamérica se sedentarizaron. Y la otra hipótesis los asocia a grupos procedentes del golfo de México en la región Olmeca. Para contrastar estos planteamientos deberíamos utilizar poblaciones de referencia, tanto de la zona Chichimeca como del área Olmeca, y

concluir a cuál serían más semejantes lo que nos podría dar información sobre su origen.

Si los dos grupos Otomíes fueran genéticamente diferentes podríamos encontrar dos posibles explicaciones. Por un lado que fueran dos grupos con sustratos biológicos distintos y que uno de ellos, por absorción o imposición adoptase la cultura Otomí. Como ya se ha mencionado la afinidad cultural con grupos con los que han convivido es muy marcada (Barrientos López 2006; Questa Rebolledo and Utrilla Sarmiento, 2006). Y por otro lado que originariamente hubieran sido el mismo grupo, y que al separarse en el pasado no hubiera habido flujo genético entre ellos, de modo que podríamos explicar esas diferencias por el pasado biológico reciente de ambos grupos tras su separación aproximada hace unos 700 años (Ochoa 1989). Los resultados obtenidos tras los análisis realizados apoyan esta segunda alternativa, ya que los grupos son parecidos genéticamente, lo que parece apuntar a un origen común. En este caso y al ser el sustrato genético el mismo, las diferencias culturales y lingüísticas que estamos viendo podrían ser debidas al tiempo que hace que se fraccionó su área de distribución, probablemente por la colonización o presión de otros grupos, como por ejemplo Nahuas. No hemos de olvidar que podríamos estar viendo aculturización de uno de los grupos y absorción por parte del otro, ya fuera voluntario o impuesto, es una opción que debe tenerse en cuenta, especialmente en el caso del grupo Otomí de la Sierra.

IV.6. Relaciones internas entre comunidades Nahuas

En el estudio de la población Nahua se han incluido individuos procedentes de cuatro municipios de la Sierra Madre Oriental, en tres de ellos se han trabajado dos comunidades y en uno, únicamente una comunidad. Todas ellas están ubicadas alrededor de un eje imaginario noreste-sudoeste y son hablantes de náhuatl. La distribución de esta segunda localización geográfica es colindante con el este de la Huasteca, estando compartida con tres grupos incluidos en esta tesis, Otomíes del Valle y de la Sierra y Tepehuas, y dos que no están incluidos, Totonacos y Téenek.

Los cuatro municipios tienen un patrón de porcentajes de haplogrupos amerindios similar, siendo los más diversos Xochiatipan y Atlapexco. Los test de diferenciación poblacional son significativos para todos los municipios, excepto entre Atlapexco y Yahualica que se encuentran muy cercanos geográficamente. Destaca Huejutla de

Reyes que, más retirada en la sierra, presenta diferencias con el resto. La representación del MDS dispone los municipios más cercanos en el mismo cuadrante, y Huejutla de Reyes alejada y en otro cuadrante, lo que sería consistente con lo ya comentado.

La presencia de grupos Nahuas en la Huasteca es fruto de migraciones históricas que han conferido a este pueblo características singulares, diferenciándolos de otros hablantes de náhuatl del país. El poblamiento de la región se habría iniciado hace 2.000 AAP con varios grupos Mayas emigrando hacia el río Pánuco y estableciéndose en la región del golfo de México (Williams García 1953). Los actuales grupos Téenek separados de los Mayas hace 3.000 años habrían llegado también en esa migración (Valle Esquivel and Hernández Alvarado, 2006). La llegada consecutiva de otros grupos a esa zona, dio lugar al escenario tan diverso y especial, con relaciones estrechas entre los grupos que actualmente encontramos en la Huasteca (Valle Esquivel 2006).

Se reconocen dos incursiones Nahuas en la región, la primera hacia el 800 ADP en el periodo Epiclásico, y la segunda como resultado de la expansión Mexica durante el siglo XV, que habría nahuatlizado parte de las poblaciones de la zona (Escobar Ohmstede 1998) aunque habrían adoptado también tradiciones del área (Valle Esquivel 2003). Sin embargo, hay autores que defienden que los Nahuas que se establecieron en el sur de la Huasteca en la primera migración debían ser en gran parte Téenek nahuatlizados; en la segunda, los Mexicas procedentes de algún punto en el noroeste, atravesaron toda Mesoamérica hasta llegar a la región central del actual México donde fundaron Tenochtitlan, que llegaría a ser la ciudad más importante de toda Mesoamérica y que junto con Tlacopan y Texcoco, formaría la llamada Triple Alianza en el año 1.430. Después de varios ciclos de conquistas sucesivas a lo largo de 100 años, dominaban grandes regiones mesoamericanas, sometiendo y controlando a los pueblos que vivían en ellas, llegando a ser más de 400 en su época de mayor esplendor. Esta oleada se superpuso seguramente a la antigua población indígena de la región, consiguiendo imponer su lengua pero no su cultura. La estrategia militar Mexica se debió ver fortalecida por estos asentamientos, ya que la ocupación de esta parte meridional de la Huasteca, abrió una puerta hacia la conquista sistemática de la franja litoral entre la sierra y el Atlántico (Stresser-Péan 2008; Williams García 1953).

Esta casi ubicuidad Nahua en el altiplano mesoamericano se ve reflejada en los *networks* construidos para los cuatro haplogrupos amerindios. Es importante tener en cuenta que, como se ha visto en el apartado anterior, la presencia de haplotipos Nahuas tanto en los nodos centrales como en gran cantidad de nodos derivados, pondría de manifiesto la antigüedad de su presencia en estas regiones y/o la heterogeneidad por las relaciones establecidas entre grupos o incluso la absorción de alguno de ellos parcial o totalmente. Los valores de diversidad para este grupo, seguido de Otomíes, son los más elevados de todas las poblaciones estudiadas en esta tesis.

La población Nahua de la Huasteca es una población singular. Su distribución en la Huasteca hidalguense y potosina, se localiza en una franja montañosa barrera natural entre el golfo de México y el altiplano central mexicano. Esta región es extremadamente abrupta y la comunicación entre valles cercanos, o incluso entre vertientes diferentes de una misma cordillera, puede resultar extremadamente compleja. Por ese motivo, el concepto de población homogénea de un mismo grupo indígena distribuida en esta región geográfica acepta consideraciones. Una de ellas es el posible grado de diferenciación que han podido experimentar los Nahuas en esta región. Debido a esto el muestreo se realizó siguiendo un detallado patrón, en el que se consideró como prioritaria la obtención de muestras de diferentes comunidades indígenas.

Al estudiar el grupo indígena a nivel de comunidad podemos encontrar un problema con el tamaño muestral para las poblaciones, aunque se asume que el número de muestras procedentes de algunas comunidades pueden ser relativamente bajo, se ha considerado interesante realizar esta prospección por lo excepcional del análisis: desde el máximo nivel de agrupación como es el grupo étnico, hasta el mínimo nivel de convivencia, que sería el de comunidad. La diversidad por comunidades es más elevada en las poblaciones de Texoloc y Yahualica, seguidas de Ixtaczoquico. Los valores más bajos son para Pahactla. Se detectan tres patrones de distribución de los porcentajes de los haplogrupos amerindios, el más extendido engloba las comunidades de Ixtaczoquico, Texoloc, Tlalchihualica y Atlalco, predominando el haplogrupo A2 seguido del B2, y con valores para C1 y D1 bajos y en algunos casos inexistentes, lo que estaría en la línea de otros grupos Nahuas reportados previamente en los estados de Veracruz, Puebla, Querétaro y Guerrero (Lorenz and Smith, 1996; Peñaloza-Espinosa et al., 2007). Pahactla y Santa Catarina tienen predominancia de A2 y valores parecidos para B2, C1 y D1, como en otra población Nahua de

Xochimilco (México D.F.) (Peñaloza-Espinosa et al., 2007). Y Tecacahuaco presenta un patrón totalmente distinto al resto, con mayoría de haplogrupo A2, seguido de B2 y C1, estando el haplogrupo D1 ausente. La representación del análisis de componentes principales, pone de manifiesto que las relaciones entre las distintas comunidades no necesariamente tienen que ver con los municipios o con los patrones mencionados de porcentajes de haplogrupos.

Los resultados para el resto de análisis no siguen esta misma línea, excepto en el caso de Tecacahuaco que se comporta como una comunidad *outlayer*, siendo además la que presenta menor número de haplotipos compartidos. Igualmente la comunidad de Santa Catarina presenta un elevado número de haplotipos únicos (15) y comparte únicamente tres. Estos valores son variables para el resto de comunidades. Llama la atención que Tecacahuaco no comparta ningún haplotipo con Texoloc, ya que es una comunidad relativamente cercana, y que sin embargo esta comparta tres con Santa Catarina. Lo mismo sucede entre Atlalco e Ixtaczoquico, que a pesar de estar muy cercanas no comparten ninguno. Los resultados para los test de diferenciación entre las poblaciones indígenas sí apoyan las diferencias de Tecacahuaco y Santa Catarina con respecto al resto, presentando significación casi para la mayoría de estos análisis (Tablas III.29 y III.30, y Figura II.2). Es importante considerar que, aunque las expansiones pueden generar patrones de diversidad caracterizados por el exceso de alelos raros, producido por el crecimiento rápido de la población y por lo tanto de haplotipos, se puede ver intensificado si hay un elevado nivel de flujo genético entre poblaciones (Fuselli et al., 2003).

Las AMOVAs realizadas entre las comunidades Nahuas se han organizado según tres criterios. En primer lugar la "Administrativa", teniendo en cuenta los municipios y las comunidades que pertenecen a cada una de ellas; en segundo lugar, según dos categorías, "Levante" y "Poniente", en las que se han distribuido las comunidades según la vertiente de la Sierra Madre Oriental en la que se encuentran. Y por último, el criterio de distribución con las agrupaciones "Norte", "Centro" y "Sur". En ningún caso las diferencias encontradas entre las categorías apoyan una subestructura genética entre las comunidades del grupo indígena Naha.

Aunque estos resultados ponen de manifiesto que las diferencias entre los *pool* genéticos de las comunidades no son lo suficientemente grandes como para que haya diferencias significativas, por lo que hemos visto hasta ahora la organización de las comunidades y las relaciones entre ellas responden a algún otro patrón que no se ha

detectado, ya que la distribución de los haplogrupos amerindios, así como lo relacionado con los haplotipos, no parece tener una explicación que podamos acotar con los datos obtenidos. Podría ser que lo que estamos encontrando tenga algo que ver con la posible confusión entre Nahuas y Téenek de filiación Maya (Martínez-Cortés et al., 2010; Solanes Carraro and Vela Ramírez, 2005; Valle Esquivel and Hernández Alvarado, 2006), puesto que como se ha mencionado con anterioridad la región Huasteca es una zona de distribución compartida entre ambos grupos indígenas. Entre las siete comunidades que estamos analizando podríamos tener algunas de filiación Nahua, otras de filiación Téenek o incluso, Téenek *nahualtizados* como se ha mencionado anteriormente, y que esto estuviera aumentando el número de haplotipos compartidos entre Mayas y estos Téenek *nahualtizados*. Este amplio abanico de posibilidades hace más complejo el análisis a este nivel.

En este sentido la revisión de haplotipos compartidos entre Nahuas y Mayas podría darnos alguna información puesto que lo podríamos considerar un indicativo de la semejanza, dándonos una idea de la cercanía gracias a la historia biológica compartida. Si estas poblaciones Nahuas fueran realmente Téenek o tuvieran relación con ellos, por la relación con los grupos Mayas, deberían mostrar alguna afinidad en cuanto al análisis de haplotipos, como efectivamente así es. Encontramos nueve haplotipos de HVI compartidos entre Nahuas y Mayas, de los cuales cuatro son exclusivos entre los grupos, y uno de ellos está distribuido por Cuba, Guatemala y Colombia, indicándonos la antigüedad del haplotipo (datos no publicados y detectados por comparación con la herramienta Mitosearch). Para la totalidad de D-Loop se detectan únicamente dos, lo que aunque no es el elevado valor que hallábamos entre los grupos Otomíes, nos está señalando la relación entre los grupos. Sin embargo en este caso, estos dos haplotipos están presentes en el resto de grupos detectados en la Sierra Otomí-tepehua, por lo que sería descartable un origen Maya de estos grupos Nahuas, al menos en lo que se refiere a esta población, a no ser que los grupos Téenek se hubieran mezclado con todas las poblaciones de la Huasteca, quedando diluida la carga genética compartida y originaria de los grupos Mayas.

Para tratar de acotar un poco más las relaciones entre las comunidades Nahuas y la posible presencia de sustrato genético Maya, se han rastreado los haplotipos compartidos de manera única entre ambos grupos. La situación para D-Loop muestra únicamente dos haplotipos, y aunque estos están compartidos por el resto de los grupos de la Sierra Otomí-tepehua, se han revisado las comunidades Nahuas a las que pertenecen los haplotipos coincidentes. Para HVI se han rastreado igualmente

estos haplotipos y se han ubicado los individuos. El resultado de este seguimiento no da un patrón claro de distribución, por lo que no podemos acotar tampoco con estos datos, la posible relación entre estos grupos.

IV.7. Relaciones entre los grupos de Hidalgo

Previamente a hacer un análisis detallado de la problemática de los grupos indígenas en el estado de Hidalgo, es necesario hacer dos apreciaciones. Por una parte recordar que es uno de los estados con mayor composición indígena de México, por lo que el estudio de estos grupos representa un elevado porcentaje de la historia biológica de una de las regiones más diversas de América. Este aspecto adquiere más valor si consideramos que el muestreo se proyectó y desarrolló con sumo cuidado, manteniendo como premisa esencial la representación detallada de la composición genética de estos grupos. Por otra parte hay que recordar como ya se ha comentado, que se utiliza el término estado de Hidalgo sin criterio poblacional, ya que se trata de una asignación administrativa. Se hace referencia a él para facilitar su ubicación geográfica. El interés real del estudio de esta región es su elevada representación indígena, su compleja historia, su abundante diversidad cultural y su variada orografía. Un marco perfecto e inmejorable para que se aborde desde la genética, la historia biológica de la región.

Tras el análisis entre grupos Otomíes y entre las distintas comunidades y municipios Nahuas, siguiendo esa línea lógica y para poder comprender la globalidad que encontramos, es esencial considerar la totalidad de grupos indígenas que están presentes en el estado de Hidalgo. Como se ha visto hasta ahora, las diferencias entre los grupos indígenas no son significativas, los dos grupos Otomíes a pesar de hablar variantes dialectales distintas no presentan subestructura genética, y las poblaciones Nahuas aunque presentan variaciones considerables en sus composiciones genéticas, tampoco están estructuradas para ninguno de los criterios considerados. En este escenario vamos a introducir además el grupo Tepehua, para tratar de integrar toda la información disponible. En cuanto a los grupos que están incluidos en esta tesis y después de los dos apartados anteriores, es importante situar a los indígenas Tepehuas antes de profundizar en un análisis global de esta zona.

El grupo Tepehua está emparentado con indígenas Totonacos que migraron desde el centro de México hacia el golfo, donde se instalaron hacia el 800 ADP. Los posteriores

movimientos de población y la separación de ambos grupos, desplazaron a los primeros hasta la sierra Otomí-tepehua, localización que siguen manteniendo en la actualidad. Según Stresser-Péan el poblamiento de la Huasteca se produjo por etapas y en el orden siguiente: Téenek; Totonacos y Tepehuas, que habrían llegado en la misma época, y Nahuas y Otomíes, cuyo orden no está claro (Stresser-Péan 2008).

La visión global de los grupos indígenas del estado de Hidalgo incluidos en este trabajo, se distribuyen en un MDS con un *stress* aceptable (Sturrock and Rocha, 2000) y siguiendo cierta lógica dentro de lo esperado. Las comunidades Otomíes del Valle y de la Sierra, se sitúan más o menos agrupadas y en posiciones cercanas. En la parte central de los ejes se sitúan la mayor parte de comunidades Nahuas, con Santa Catarina desplazada hacia las poblaciones Otomíes, y Tecacahuaco desplazada al otro extremo del eje y aislada del resto. En ambos casos y vistos los datos revisados en el apartado anterior, esta ubicación apoya las diferencias anteriormente mencionadas. Por último, es importante tener en cuenta que Huehuetla, la comunidad Tepehua, se ubica en el límite inferior a la zona de las comunidades Nahuas (Figura II.2).

La comparación basada en distancias F_{st} entre las catorce comunidades indígenas incluidas en esta tesis, nos da una visión un tanto confusa de las semejanzas y diferencias entre las mismas, aunque destacan las comunidades que se diferencian de todo el resto como Tecacahuaco, población Nahua que como ya se ha visto tiene características especiales respecto al resto. San Nicolás y El Buena, comunidades Otomíes de la Sierra y del Valle respectivamente, aún siendo las más lejanas entre las seis poblaciones Otomíes se diferencian de todo el resto excepto entre ellas. Destaca también Santa Mónica, comunidad Otomí de la Sierra, que se diferencia de todas las otras comunidades excepto de Santiago, población más cercana de su mismo grupo indígena; y Santa Catarina, población Nahua ubicada en el norte de la Sierra. El test exacto de diferenciación poblacional nos dibuja un panorama de total diferenciación entre las distintas comunidades, excepto entre Pahactla, Atlalco, Texoloc y Tlalchihualica, situadas en la parte central de la zona de distribución del grupo Nahua. Sin embargo llama la atención que estando Ixtaczoquico y Tecacahuaco en esta misma área, estas si se diferencien del resto (Figura II.2).

Las AMOVAs realizadas entre las catorce comunidades indígenas estudiadas del estado de Hidalgo, se han construido en base a tres criterios. En primer lugar por grupo indígena, agrupando las comunidades en Nahuas, Otomíes del Valle y de la

Sierra, y Tepehuas. Como segundo criterio se ha adoptado la orografía, agrupando en la categoría “Valle” las tres comunidades de Otomíes situadas en la planicie hidalguense, y en “Sierra” al resto. Por último, se ha utilizado un tercer criterio que considera las vertientes de la sierra en que se localizan las comunidades, “Vertiente Valle” y “Vertiente Costa”.

La primera agrupación entre grupos indígenas es significativa. Este resultado es inesperado si tenemos en cuenta que el análisis entre los grupos Otomíes, no lo era y sin embargo sí lo es al incluir el resto de grupos con los que conviven en el estado de Hidalgo. Esto nos está indicando que sí hay una subestructura entre los distintos grupos, lo que clarifica en cierta manera el escenario en esta zona. Es una conclusión de gran interés ya que los trabajos realizados hasta el momento en esta región o en regiones aledañas, no han detectado diferencias significativas entre troncos lingüísticos, que en esencia es el parámetro que tradicionalmente se ha utilizado para designar a los grupos indígenas (Kemp et al., 2010; Sandoval et al., 2006). El porqué de estas diferencias entre los valores presentados y los propuestos por otros autores, posiblemente habría que buscarlos en el estricto control que se ha realizado de las muestras en el presente trabajo.

La segunda AMOVA con las agrupaciones “Valle” y “Sierra” no es significativa, probablemente debido a que la dinámica de las poblaciones incluidas en este estudio no obedece a un criterio tan genérico. Sobre todo si tenemos en cuenta que las siguientes agrupaciones sí son significativas, detectándose subestructura genética entre las vertientes de la Sierra, sin incluir los dos grupos de Otomíes que al fin y al cabo están aislados unos en el valle del Mezquital, y otros en la zona sur de la Sierra Otomí-tepehua. Esto dibuja una situación de cierto aislamiento e independencia de los dos grupos Otomíes, así como de las vertientes valle y costa. Estas dos últimas agrupaciones incluyen las comunidades Nahuas y Tepehua, según sus posiciones respecto a la máxima altura montañosa, incluyendo Huehuetla junto con Texoloc e Ixtaczoquico en la vertiente costera, y el resto de comunidades, todas Nahuas, en la vertiente del valle (Tabla III.31, Figura II.2). Esta configuración de las comunidades explica la diferenciación que vemos entre ellas, por lo que la localización geográfica en una u otra vertiente de la Sierra Madre Oriental puede considerarse una de las causas de esta estructuración genética, más que la organización por grupos indígenas. Esto pone de manifiesto que los patrones de relaciones entre las comunidades obedecen más a su ubicación geográfica, y como esta determina la comunicación entre ellas, que a la supuesta separación de grupos y diferencias entre ellos.

En cuanto a los haplotipos compartidos entre los cuatro grupos indígenas, los valores más elevados los encontramos entre los grupos Otomíes y el Nahuatl, compartiendo un número más bajo de haplotipos con la población Tepehua, tanto si analizamos HVI como si analizamos D-Loop, lo que estaría apoyando la historia distinta de este grupo con respecto a los otros dos y ya mencionada con anterioridad. Sin embargo, si buscamos cuáles son los haplotipos compartidos de forma única entre los grupos, obtenemos un poco más de información, ya que los Nahuatl comparten con Otomíes del Valle cuatro haplotipos para D-Loop y dos para HVI. Estos últimos haplotipos volcados en la base de datos Mitosearch son coincidentes únicamente para dos de ellos, que tienen una distribución amplia sobre todo en el sur del continente americano, Brasil, Puerto Rico, Venezuela y Perú, siendo coincidente para todos los puntos variables del haplotipo con la Momia Juanita de Perú (McKenney et al., 2007), lo que sería un buen ejemplo de continuidad entre poblaciones precolombinas y actuales (datos no publicados). Si comparamos Nahuatl con Otomíes de la Sierra, tres haplotipos son compartidos tanto para D-Loop como para HVI, y solamente para uno de estos últimos es coincidente con alguna muestra reportada previamente, distribuyéndose en Michoacán y en otras localizaciones no detalladas de México. Cuando comparamos los dos grupos Otomíes entre ellos, se detectan cuatro haplotipos compartidos para D-Loop y cinco para HVI, uno de los cuales se distribuye por California, Canadá y Texas en el norte del continente, y Chile, Puerto Rico, Guatemala, Colombia, Costa Rica, República Dominicana, Brasil y Venezuela, en el sur. Si comparamos Nahuatl con Tepehuas, solamente se encuentran dos coincidencias para D-Loop y una para HVI, que no está reportada previamente en la base de datos Mitosearch. Por último, ninguno de los grupos indígenas Otomíes tiene haplotipos compartidos únicamente con Tepehuas.

Estos datos y coincidencias de haplotipos compartidos únicamente entre algunos grupos, podrían estar indicando la antigüedad de las poblaciones, señalando también el tiempo que hace que se habrían separado. Si analizamos los datos según este punto de vista, las semejanzas entre Nahuatl y grupos Otomíes podría estar señalando un origen común, apuntado ya anteriormente en el norte de Mesoamérica y de origen Chichimeca. Los Otomíes habrían continuado como grupo hasta su separación hace alrededor de 700 años (Ochoa 1989), lo que mantendría más semejanzas entre ellos y de ahí el mayor número de haplotipos compartidos de forma única. En ambos casos, la distribución de algunos de estos haplotipos en el resto del continente pero especialmente en el sur, podría apoyar también que son haplotipos antiguos puesto

que han tenido más tiempo para dispersarse en esa parte del continente. Las diferencias que estos grupos presentan con el grupo Tepehua podrían señalar el menor tiempo que han compartido en sus actuales distribuciones, por lo que la falta de relación entre ellos habría dado lugar a esta configuración. De cualquier manera, es importante tener en cuenta que esto podría deberse al bajo valor muestral de la población Tepehua.

IV.8. Análisis por grupos indígenas en el contexto americano

Los porcentajes de los distintos haplogrupos detectados en las poblaciones incluidas en esta tesis, presentan algunas diferencias con los de los trabajos publicados hasta la actualidad, aunque en general son concordantes con el modelo general de distribución de los haplogrupos en el continente americano (Lalueza et al., 1997). Sin embargo, y aunque también responden en términos generales con el patrón reportado por Salzano (2002), que establece que las poblaciones mexicanas presentan elevadas frecuencias de los haplogrupos A2 y B2 frente a C1 y D1, no se cumple cuando consideramos los grupos indígenas individualmente, como tampoco se cumple el patrón reportado por Malhi con elevados valores de A2 y bajos para B2 (Malhi et al., 2007). Las poblaciones Mayos, Huicholes, Nahuas, Tepehuas y Mayas cumplen esta distribución, mientras que Pimas y Otomíes, del Valle y de la Sierra, presentan otros modelos.

La ubicación de las poblaciones incluidas en esta tesis en los ejes del MDS, engloba en la parte inferior (Figura III.42), las poblaciones del norte, correspondiente a Oasisamérica, aunque la población Pima estudiada en esta tesis está desplazada con respecto al resto. Esto podría estar causado por el bajo valor muestral y un posible efecto de deriva genética, lo que habría alterado las proporciones de los haplogrupos en la población. En la parte superior del MDS, se distribuyen las poblaciones de Mesoamérica y algunas del sur de Estados Unidos. Están agrupadas las poblaciones de la Sierra Otomí-tepehua, con otras de la región mesoamericana bastante cercanas e incluso con otros grupos localizados más al sur. Llama la atención la distribución dispersa de los grupos Nahuas por toda la parte superior de la figura, comprensible por el variado origen y procedencia de los grupos incluidos en el análisis. En realidad lo que indica esta figura es que no hay grandes diferencias entre las distintas poblaciones incluidas, por lo que se analizaron los datos con un *software* específico para tratar de acotar las relaciones y lógica entre estas poblaciones.

A través de la triangulación, el programa Barrier permite la representación geográfica de las distancias genéticas entre poblaciones asociadas a un *bootstrap*, que se traduce en la intensidad de las barreras. El uso de este programa nos permite visualizar la ruptura en la distribución de las frecuencias de los haplogrupos (Figuras III.44 y III.45). El estudio conjunto de los cinco haplogrupos muestra homogeneidad genética en Mesoamérica, los límites de esta continuidad se detectan hacia el sur, coincidiendo prácticamente con la frontera política entre Panamá y Colombia. Hacia el norte el panorama es más complejo, estableciéndose la barrera más evidente en el límite de las regiones definidas como Oasisamérica y Aridoamérica con el resto del continente (Figura I.6). Aunque no es el objetivo de esta tesis, este resultado aporta algo de información a los trabajos realizados sobre la expansión del maíz y la interrelación entre Mesoamérica y el llamado “southwest” (Kemp et al., 2010). Esta importante barrera apunta a la existencia de una falta de flujo génico entre estas dos regiones, por lo que la difusión de la cultura del maíz podría haberse realizado sin la necesidad de movimiento poblacional.

Una segunda barrera de menor intensidad y situada en latitudes inferiores a la anterior, separaría las poblaciones que tradicionalmente han desarrollado un estrategia de supervivencia basada hasta hace poco en la caza y la recolección. En este sentido, se intuyen las diferencias entre las estrategias de supervivencia y su efecto en el modelo demográfico. Las sociedades sedentarias mantienen tamaños demográficos elevados respecto a grupos con otras estrategias, lo que implicaría un efecto diferencial de las fuerzas evolutivas (sobre todo deriva génica), permitiendo una diferenciación genética entre ellas. Entre estas poblaciones de cazadores-recolectores destaca la existencia de una barrera alrededor de la población Pima. La falta de datos de poblaciones cercanas, impide determinar si este aislamiento es exclusivo de la población Pima, o si se trata de un fenómeno poblacional más amplio. No debe perderse de vista el bajísimo tamaño demográfico de esta población, al que se ha hecho referencia reiteradamente, haciéndola muy sensible al efecto de la deriva genética u otros fenómenos relacionados.

Es importante tener presente que la visión aportada por Barrier está basada en frecuencias de los haplogrupos para las poblaciones disponibles, lo que resulta menos resolutivo, con menor peso estadístico y profundidad, sobre todo si se compara con resultados procedentes de otros análisis basados en secuencias completas de la región control. En cualquier caso, se observa una elevada homogeneidad entre las poblaciones mesoamericanas, que probablemente es el resultado de la compleja

historia de esta región, en la que el desarrollo de grandes culturas y las relaciones entre ellas, facilitaron importantes movimientos demográficos.

Para tratar de comprender al detalle la composición de haplogrupos de las poblaciones estudiadas, se analizaron de forma independiente las barreras para cada uno de los haplogrupos (Figuras III.46 a III.49). Para el haplogrupo A la distribución es uniforme si trabajamos con diversas poblaciones de la región mesoamericana, y se han detectado barreras que marcarían una línea imaginaria en el límite de Mesoamérica con Aridoamérica por el norte, y otra en la frontera política entre Costa Rica y Panamá. Para los haplogrupos B y C, la distribución mostrada por Barrier estaría focalizada también en la zona central del continente, siendo bastante parecida que para el haplogrupo A. En los tres casos las distribuciones se ajustarían a lo descrito por Lalueza (Lalueza et al., 1997) pero no con lo aportado por otros estudios (Guardado-Estrada et al., 2009; Merriwether et al., 2005; Salzano 2002).

Por último, la distribución de la variante D en la región mesoamericana es distinta a lo descrito hasta el momento. Los resultados de Barrier esbozan una distribución costera del haplogrupo D, reflejando la presencia de la variante con un patrón distinto que sería concordante con un posible poblamiento paralelo a la costa del Pacífico (González-José et al., 2003; Kitchen et al., 2008; Perego et al., 2009). Esto estaría apoyado por la presencia de haplogrupo D en otras poblaciones de la región pacífica del continente (Bonilla et al., 2005; Lorenz and Smith, 1996; Malhi et al., 2003; Peñaloza-Espinosa et al., 2007; Sandoval-Ramirez et al., 2006; Torroni et al., 1993b). Si bien es cierto que el número de poblaciones incluidas en este análisis no es muy elevado, este hecho ya ha sido descrito por algunos autores para subvariantes entre las que se encuentra D4h3 (Kitchen et al., 2008; Perego et al., 2009).

Es importante no perder de vista en ningún momento que los resultados que estamos viendo están basados en el ADNmt, por lo que la discusión que se puede hacer de ellos y lo que nos indican, es la historia femenina de las poblaciones estudiadas. En este sentido los resultados obtenidos coinciden con otros trabajos que detectaron correlaciones significativas entre la geografía y el ADNmt. En este mismo trabajo, centrado en Mesoamérica y el término norteamericano “southwest”, los autores sí que detectaron correlación con cromosoma Y, por lo que apoyan la idea de que la historia masculina y femenina presenta desviaciones. Este resultado no es extraño, ya que parte de las comunidades estudiadas en esta tesis se encuentran ubicadas en la región Huasteca, en la que se han descrito patrones de residencia matrilocal. Algunos

autores a través de estudios etnográficos han reportado que las mujeres al casarse tienen el derecho de recibir una parcela de su progenitor cuando la pareja escoge residencia en su propia comunidad, aportando la vivienda con la dote (Kellog 2005). Es costumbre que los matrimonios sean exogámicos ya que los hombres suelen elegir mujeres con las que no están emparentadas y de una zona lejana (Carrasco 1950; Cervantes Escandón 1978). Además en el caso de esta tesis, la pormenorizada estrategia y minucioso muestreo realizado, ha permitido incluso determinar correlaciones geográficas más detalladas como la significación de la AMOVAs al considerar geografía en ambas vertientes de la Sierra (Tabla III.31), categorías “Levante” y “Poniente”. De hecho sí que se ha detectado un modelo diferencial en la distribución de las frecuencias de los haplogrupos en poblaciones de una u otra vertiente (Figura III.31), prueba del movimiento poblacional entre ellas.

V. CONCLUSIONES

La planificación del muestreo ha mostrado ser una parte fundamental de este trabajo, permitiendo la organización de las bases de datos y la selección de los individuos. Es importante destacar los siguientes puntos:

- El diseño del muestreo debe adaptarse al máximo a la región en la que se está trabajando y a la idiosincrasia de las comunidades indígenas, afectando lo mínimo imprescindible a las vidas de la población con el objetivo de que la campaña sea exitosa.
- El muestreo directo en las comunidades indígenas permite la visión detallada de la población analizada, así como la evaluación exacta de los individuos incluidos en el estudio para la obtención de la totalidad de datos necesarios para crear una base de datos completa.
- El muestreo en varias comunidades de un mismo grupo indígena permite el análisis de la variabilidad y estructura genética que hay en la población. Los resultados nos permiten el análisis del grupo de una manera integral, posibilitando la comprensión de los fenómenos teóricos que nos da la bibliografía.
- Debe prestarse especial atención en poblaciones indígenas de pequeño tamaño muestral por las desviaciones que pueden cometerse. Este aspecto es extrapolable a grupos indígenas con población total baja.

El trabajo en el laboratorio es una parte esencial en cualquier investigación, en este caso la complejidad al trabajar con muestras de orígenes distintos y obtenidas a lo largo de los años, han obligado a dar un importante giro para la optimización de los protocolos en el laboratorio. En referencia a esto, es importante destacar que:

- El tipo de muestra biológica elegida para el muestreo es determinante en el éxito de una campaña, pero deben considerarse también las consecuencias técnicas que eso puede tener en el laboratorio.
- Los protocolos de extracción elegidos para los distintos tipos de material biológico han de permitir la optimización de los resultados en el laboratorio, por lo que es esencial una correcta puesta a punto de los métodos y una conservación adecuada. En el caso del trabajo realizado con resina Chelex, ambos han demostrado ser determinantes.

- El estudio de D-Loop completa es preferible frente a la región HVI sola o en combinación con HVII. La interpretación de los datos en base a una mayor información, permite estudios más detallados y conclusiones más pormenorizadas.

El análisis profundo realizado en esta tesis, permite una aproximación a la historia de las poblaciones estudiadas, siendo concluyentes los siguientes aspectos:

- Las diversidades genéticas detectadas en las comunidades estudiadas son relativamente elevadas respecto a otros estudios, debido en gran parte al control estricto de los parámetros que provocan desviaciones en estas estimaciones.
- Las frecuencias de los haplogrupos amerindios se ajustan a lo esperado en regiones de estas latitudes para el continente americano.
- El número de haplogrupos no amerindios detectados es bajo, como consecuencia del riguroso control realizado con las encuestas de la población.
- La distribución de los haplogrupos europeos en la Sierra Otomí-tepehua muestra que, esta región supuestamente aislada, presenta más tendencia al mestizaje que el resto. Esto apunta a que esta zona representó una conexión entre el altiplano central y el litoral atlántico.
- Todos los grupos indígenas estudiados presentan un modelo de expansión demográfica a excepción de Pimas.
- Mesoamérica se caracteriza por cierta homogeneidad genética, aunque a través del estudio de la región control del ADNmt es posible una aproximación más detallada de la historia de cada una de las poblaciones.
- En la población Pima se detecta un modelo general distinto al resto de grupos incluidos en esta tesis, caracterizado sobre todo por una baja diversidad. Esta situación es debida a un pasado cazador-recolector que implica una estructura demográfica diferente a la mesoamericana, facilitando el efecto de la deriva genética y probablemente intensificado por un total poblacional muy bajo.
- El grupo Mayo presenta una posición intermedia entre Mesoamérica y Oasisamérica.

- El otro representante del occidente mexicano, el grupo indígena Huichol, constituye una escisión del sustrato genético mesoamericano posterior a la detectada para el grupo Maya.
- La población Nahua es la población más diversa y heterogénea de Mesoamérica, y el sustrato genético que presenta es el reflejo de su compleja historia. Es la población más antigua y comparte con los grupos de la sierra, Otomíes y Tepehuas, gran parte del sustrato genético ancestral.
- Se han detectado tres modelos distintos en la distribución de las frecuencias de los haplogrupos en las comunidades indígenas Nahuas, indicando flujo genético diferencial entre ellas.
- Los grupos Otomíes tienen un origen común. Las diferencias detectadas entre ellos son debidas al tiempo transcurrido desde su separación, y al flujo genético del grupo de la sierra con otras poblaciones de la Huasteca.
- El sustrato genético Tepehua presenta semejanzas con grupos Otomíes y Nahuas, evidenciando el origen común de estos grupos. De cualquiera manera es el único grupo Totonaco presente en la sierra e incluido en esta tesis, lo que se refleja en sus características genéticas.
- El sur mesoamericano representado por el grupo indígena Maya, se caracteriza por una elevada diversidad fruto de los movimientos de población asociados a la historia de la región, aunque el sustrato genético que presenta procede del grupo ancestral. La escisión fue previa al grupo Huichol.
- La presencia de haplotipos compartidos entre grupos Nahuas y Mayas, es justificable con la difusión de genes de estos últimos a través de la población Téenek.
- Se ha determinado que un modelo matrilocal tiene un efecto favorable sobre la subestructuración genética de las poblaciones estudiadas.

Una vez tratadas las poblaciones de forma individual y aunque el estado de Hidalgo no constituye una entidad poblacional, se considerarán conjuntamente los grupos habitantes de esta región.

- Un análisis pormenorizado de los grupos presentes en el valle del Mezquital y de la sierra Otomí-tepehua, permite describir una estructura genética asociada a la distribución geográfica de las comunidades, detectándose un patrón relacionado a las vertientes litoral y continental de la sierra.
- El criterio lingüístico no ha sido determinante en el comportamiento de los grupos ni en las relaciones establecidas entre ellos, no siendo un parámetro fundamental en la subestructuración genética.
- La Huasteca constituye una entidad geográfica que permite el intercambio y el contacto entre los distintos grupos indígenas que la habitan. Esta circunstancia ha permitido que sea la orografía la que regule el flujo genético entre las comunidades, aunque se mantiene una estructura genética asociada a grupos indígenas.

La integración de la información genética obtenida en el estudio de los ocho grupos permite llegar a las siguientes conclusiones.

- Las regiones culturales definidas para Mesoamérica determinaron las dinámicas poblacionales de los grupos indígenas, estructurando el sustrato genético de cada uno de ellos, aún detectable hoy en día. Este resultado demuestra la importancia que los procesos históricos y culturales, unidos a los geográficos, han tenido en las interrelaciones poblacionales.
- Se ha detectado subestructura poblacional en el estado de Hidalgo asociada a geografía y a grupo indígena, y a las áreas culturales en el ámbito mesoamericano. Debe considerarse la importancia que la escala del análisis tiene sobre la detección de la estructura genética.
- Las frecuencias de los haplogrupos en el contexto mesoamericano, permite definir tres regiones. El límite más marcado se sitúa en el norte de Oasisamérica. Una segunda barrera de menor intensidad define una franja entre esta primera y las regiones más septentrionales de Mesoamérica, en ella destaca el grupo Pima. La homogeneidad domina el centro mesoamericano hasta prácticamente el inicio de Sudamérica.

VI. BIBLIOGRAFÍA

- Achilli A, Perego UA, Bravi CM, Coble MD, Kong QP, Woodward SR, Salas A, Torroni A, Bandelt HJ. 2008. The phylogeny of the four pan-American MtDNA haplogroups: implications for evolutionary and disease studies. *PLoS One*. 3(3):e1764
- Acosta J de. 1589. *Historia Natural y Moral de las Indias*. Sevilla.
- Adcock GJ, Dennis ES, Easton S, Huttley GA, Jermiin LS, Peacock WJ, Thorne A. 2001. Mitochondrial DNA sequences in ancient Australians: Implications for modern human origins. *Proc Natl Acad Sci USA*. 98(2):537-542
- Adelung JC, Vater JS. 1816. *Mithridates oder allgemeine Sprachenkunde mit dem Vater unser als Sprachprobe in bey nahe fünf hundert Sprachen und Mundarten. Dritten Theil, Dritte Abtheilung*. Vossischen Buchandlung, Berlín
- Adler R. 2000. Voices from the past. *New Scientist*. 165(2227): 36-40
- Adovasio JM, Pedler DR. 2004. in *Entering America: Northeast Asia and Beringia before the Last Glacial Maximum*, D. B. Madsen, Ed. (Univ. of Utah Press, Salt Lake City, UT), chap. 5
- Adovasio JM. 1993. The ones that will not go away. A biased view of pre-Clovis populations in the New World. In: *From Kostenki to Clovis: Upper Paleolithic-paleo-Indian adaptations* (ed. O. Soffer & N. D. Praslov), pp. 199–218. New York: Plenum
- Aguirre-Beltrán G. 1989. *La población negra de México: Obra antropológica. V.2 Ciudad de México*. Fondo de Cultura Económica. México
- Al Soud WA, Jonsson LJ, Radstrom P. 2000. Identification and characterization of immunoglobulin G in blood as a major inhibitor of diagnostic PCR. *J Clin Microbiol* 38: 345-350
- Al Soud WA, Radstrom P. 2001. Purification and characterization of PCR-inhibitory components in blood cells. *J Clin Microbiol*. 39: 485-493
- Álvarez-Iglesias V, Jaime JC, Carracedo A, Salas, A. 2007. Coding regions mitochondrial DNA SNPs: targeting east Asian and native American haplogroups. *Forensic Sci Int Genet* 1: 44-55
- Alves-Silva J, Silva Santos M, Guimaraes PEM, Ferreira ACS, Bandelt HJ, Pena SDJ, Ferreira Prado V. 2000. The ancestry of Brazilian mtDNA lineages. *Am J Hum Genet*. 67: 444-461
- Ameghino F. 1921. Le diprothomo d'après Schwalbe et d'après moi. *Anal. Mus. Nac. Buenos Aires*. 31:1-24
- Amos W, Hoffman JI. 2010. Evidence that two main bottleneck events shaped modern human genetic diversity. *Proc Biol Sci*. 277(1678):131-137
- Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M. H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J., Staden, R. and Young, I.G. 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290: 457-465
- Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genet* 23: 147

- Ankel-Simons F, Cummins JM. 1996. Misconceptions about mitochondria and mammalian fertilization: implications for theories on human evolution. *Proceedings National Academy Science USA*. 93: 13859-13863
- Arce JL, Macías JL, Vázquez-Selem L. 2003. The 10.5 ka Plinian eruption of Nevado de Toluca volcano, Mexico: Stratigraphy and hazard implications. *GSA Bulletin*. 115 (2): 230-248
- Arctander P. 1999. Mitochondrial recombination? *Science* 284:2090–2091
- Arnason U, Gullberg A, Xu X. 1996a. A complete mitochondrial DNA molecule of the white-handed gibbon, *Hylobates lar*, and comparison among individual mitochondrial genes of all hominoid genera. *Hereditas*. 124: 185-189
- Arnason U, Xu X, Gullberg, A. 1996b. Comparison between the complete mitochondrial DNA sequences of *Homo* and the common chimpanzee based on nonchimeric sequences. *J. Mol. Evol.* 42(2): 145-152
- Arnason U, Gullberg A, Janke A. 1998. Molecular timing of primate divergences as estimated by two nonprimate calibration points. *J Mol Evol.* 47(6):718-727
- Arnason U, Gullberg A, Burguete AS, Janke A. 2000. Molecular estimates of primate divergences and new hypotheses for primate dispersal and the origin of modern humans. *Hereditas*. 133(3):217-228
- Atkinson QD, Gray RD, Drummond AJ. 2008. mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. *Mol Biol Evol.* 25(2):468-474
- Awadalla P, Eyre-Walker A, Maynard Smith J. 1999. Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science* 286:2524–2525
- Babalini C, Martínez-Labarga C, Tolk HV, Kivisild T, Giampaolo R, Tarsi T, Contini I, Barac L, Jančićević B, Martinović Klarić I, Perićić M, Sujoldzić A, Villems R, Biondi G, Rudan P, Rickards O. 2005. The population history of the Croatian linguistic minority of Molise (southern Italy): a maternal view. *Eur J Hum Genet.* 13(8):902-912
- Bailliet G, Rothhammer F, Carnese FR, Bravi CM, Bianchi NO. 1994. Founder mitochondrial haplotypes in Amerindian populations. *Am J Hum Genet* 55:27-33
- Balaresque P, Bowden GR, Adams SM, Leung HY, King TE, Rosser ZH, Goodwin J, Moisan JP, Richard C, Millward A, Demaine AG, Barbujani G, Previderè C, Wilson IJ, Tyler-Smith C, Jobling MA. 2010. A predominantly neolithic origin for European paternal lineages. *PLoS Biol.* 8(1):e1000285
- Balding DJ. 1999. When can a DNA profile be regarded as unique? *Sci Justice.* 39(4):257-260
- Bamshad M, Wooding S, Salisbury BA, Stephens JC. 2004. Deconstructing the relationship between genetics and race. *Nat Rev Genet* 5:598-609
- Bandelt HJ, Forster P, Sykes BC, Richards MB. 1995. Mitochondrial portraits of human populations using median networks. *Genetics.* 141(2):743-753

- Bandelt HJ, Herrnstadt C, Yao YG, Kong QP, Kivisild T, Rengo C, Scozzari R, Richards M, Villems R, Macaulay V, Howell N, Torroni A, Zhang YP. 2003. Identification of Native American founder mtDNAs through the analysis of complete mtDNA sequences: some caveats. *Ann Hum Genet* 67: 512-524
- Barbetti M, Allen H. 1972. Prehistoric man at Lake Mungo, Australia, by 32,000 years BP. *Nature*. 3;240:46-48
- Barbujani G, Sokal RR. 1990. Zones of sharp genetic change in Europe are also linguistic boundaries. *Proc Natl Acad Sci USA*. 87(5):1816-1819
- Barbujani G, Bertorelle G, Capitani G, Scozzari R. 1995. Geographical structuring in the mtDNA of Italians. *Proc Natl Acad Sci USA*. 92(20):9171-9175
- Barbujani G, Magagni A, Minch E, Cavalli-Sforza LL. 1997. An apportionment of human DNA diversity. *Proc Natl Acad Sci USA*. 94(9):4516-4519
- Barbujani G, Goldstein D. 2004. Africans and Asians abroad: genetic diversity in Europe. *Annu Rev Genomics Hum Genet*. 5: 119-150
- Barbujani G. 1997. DNA variation and language affinities. *Am J Hum Genet*. 61(5):1011-1014
- Barrantes R, Smouse PE, Mohrenweiser HW, Gershowitz H, Azofeifa J, Arias TD, Neel JV. 1990. Microevolution in lower Central America: genetic characterization of the Chibcha-speaking groups of Costa Rica and Panama, and a consensus taxonomy based on genetic and linguistic affinity. *Am J Hum Genet*. 46(1):63-84
- Barrientos López G. 2006. Otomíes del Estado de México. Colección Pueblos del México contemporáneo. Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CoNaDePI). Programa de las Naciones Unidas para el desarrollo. México
- Barrot C, Sánchez C, Ortega M, González-Martín A, Brand-Casadevall C, Gorostiza A, Huguet E, Corbella J, Gené M. 2005. Characterisation of three Amerindian populations from Hidalgo State (Mexico) by 15 STR-PCR polymorphisms. *Int J Legal Med*. 119(2):111-115
- Batzer MA, Stoneking M, Alegria-Hartman M, Bazan H, Kass DH, Shaikh TH, Novick GE, Ioannou PA, Scheer WD, Herrera RJ, Deininger PL. 1994. African origin of human-specific polymorphic Alu insertions. *Proc Natl Acad Sci USA*. 91(25):12288-12292
- Behar DM, Hammer MF, Garrigan D, Villems R, Bonne-Tamir B, Richards M, Gurwitz D, Rosengarten D, Kaplan M, Della Pergolla S, Quintana-Murci L, Shorecki K. 2004. MtDNA evidence for a genetic bottleneck in the early history of the Ashkenazi Jewish population. *Eur J Hum Genet* 12(5): 355-364
- Behar DM, Rosset S, Blue-Smith J, Balanovsky O, Tzur S, Comas D, Mitchell RJ, Quintana-Murci L, Tyler-Smith C, Wells RS; Genographic Consortium. 2007. The Genographic Project public participation mitochondrial DNA database. *PLoS Genet*. 3(6):e104. Erratum in: *PLoS Genet*. 2007 14;3(9):1785
- Behar DM, Villems R, Soodyall H, Blue-Smith J, Pereira L, Metspalu E, Scozzari R, Makkan H, Tzur S, Comas D, Bertranpetit J, Quintana-Murci L, Tyler-Smith C, Wells

- RS, Rosset S. 2008. Genographic Consortium. The dawn of human matrilineal diversity. *Am J Hum Genet.* 82(5):1130-1140
- Beltrán, A.G. 1944. The slave trade in Mexico. *Hispanic Am Hist Rev* 24: 412-431
- Bodenteich A, Mitchell LG, Polymeropoulos MH, Merril CR. 1992. Dinucleotide repeat in the human mitochondrial D-loop. *Hum Mol Genet.* 1(2):140
- Boldurian AT. 1991. Folsom mobility and organization of lithic technology: A view from Blackwater Draw, New Mexico. *Plains Anthropologist* 36(137):281-295
- Bolnick DA, Shook BA, Campbell L, Goddard I. 2004. Problematic use of Greenberg's linguistic classification of the Americas in studies of native American genetic variation. *Am J Hum Genet*, 75: 519-523
- Bonatto SL, Salzano FM. 1997a. Diversity and age of the four major mtDNA haplogroups, and their implications for the peopling of the New World. *Am J Hum Genet* 61:1413-1423
- Bonatto SL, Salzano FM. 1997b. A single and early origin for the peopling of the Americas supported by mitochondrial DNA sequence data. *Proc Nat Acad Sci USA* 94: 1866-1871
- Bonilla C, Gutiérrez G, Parra EJ, Kline C, Shriver MD. 2005. Admixture analysis of a rural population of the state of Guerrero, Mexico. *Am J Phys Anthropol.* 128(4):861-869
- Bopp F. 1816. Ueber das Konjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprache. Francfort
- Bortolini MC, Zago MA, Salzano FM, Silva-Júnior WA, Bonatto SL, Da Silva MCBO, Weimer TA. 1997. Evolutionary and anthropological implications of mitochondrial DNA variation in African Brazilian populations. *Hum Biol* 69: 141-159
- Bortolini MC, Salzano FM, Thomas MG, Stuart S, Nasanen SP, Bau CH, Hutz MH, Layrisse Z, Petzl-Erler ML, Tsuneto LT, Hill K, Hurtado AM, Castro-de-Guerra D, Torres MM, Groot H, Michalski R, Nymadawa P, Bedoya G, Bradman N, Labuda D, Ruiz-Linares A. 2003. Y-chromosome evidence for differing ancient demographic histories in the Americas. *Am J Hum Genet* 73:524-539
- Bosch E, Calafell F, Perez-Lezaun A, Clarimon J, Comas D, Mateu E, Martinez-Arias R, Morera B, Brakez Z, Akhayat O, Sefiani A, Hariti G, Cambon-Thomsen A, Bertranpetit J. 2000. Genetic structure of north-west Africa revealed by STR analysis. *Eur J Hum Genet.* 8(5):360-366
- Bowler JM, Johnston H, Olley JM, Prescott JR, Roberts RG, Shawcross W, Spooner NA. 2003. New ages for human occupation and climatic change at Lake Mungo, Australia. *Nature.* 421:837-840
- Brakez Z, Bosch E, Izaabel H, Akhayat O, Comas D, Bertranpetit J, Calafell F. 2001. Human mitochondrial DNA sequence variation in the Moroccan population of the Souss area. *Ann Hum Biol.* 28(3):295-307
- Brandstätter A, Peterson CT, Irwin JA, Mpoke S, Koech DK, Parson W, Parson TJ. 2004. Mitochondrial DNA control region sequences from Nairobi (Kenya): inferring

- phylogenetic parameters for the establishment of a forensic database. *Int J Legal Med* 118: 294-306
- Brasel KE, Reif D. 1979. A procedure to generate Thiessen polygons. *Geographic Analysis*. 325: 31-36
- Briggs AW, Good JM, Green RE, Krause J, Maricic T, Stenzel U, Lalueza-Fox C, Rudan P, Brajkovic D, Kucan Z, Gusic I, Schmitz R, Doronichev VB, Golovanova LV, de la Rasilla M, Fortea J, Rosas A, Pääbo S. 2009. Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science*. 325(5938):318-321
- Bromham L, Eyre-Walker A, Smith NH, Maynard Smith J. 2003. Mitochondrial Steve: paternal inheritance of mitochondria in humans. *Trends Ecology Evolution*. 18:2-4
- Brown P. 2000. Australian Pleistocene variation and the sex of Lake Mungo 3. *J Hum Evol*. 38(5):743-749
- Brown WM. 1980. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc Natl Acad Sci USA* 77:3605-3609
- Brown WM, George M, Wilson AC. 1979. Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci USA*. 76: 1967-1971
- Brown WM, Prager EM, Wang A, Wilson AC. 1982. Mitochondrial DNA sequences in primates: tempo and mode of evolution. *J Mol Evol*. 18: 225-239
- Brown MD, Hosseini SH, Torroni A, Bandelt HJ, Allen JC, Schurr TG, Scozzari R, Cruciani F, Wallace DC. 1998. mtDNA haplogroup X: An ancient link between Europe/Western Asia and North America? *Am J Hum Genet*. 63(6):1852-1861
- Budowle B, Wilson MR, DiZinno JA, Stauffer C, Fasano MA, Holland MM, Monson KL. 1999. Mitochondrial DNA regions HVI and HVII population data. *Forensic Sci Int*. 12: 23-35
- Campbell L, Kaufman T, Smith-Stark T. 1986. Mesoamerica as a linguistic area. *Language* 62: 530-570
- Campbell L, Mithun M. 1979. *The languages of native America: Historical and comparative assessment*. Austin, University of Texas Press
- Campbell L. 1979. Middle american languages. En "The languages of native america: an historical and comparative assessment" Campbell and Mithun Editors. Austin, University of Texas
- Campbell L. 1997. *American Indian languages. The historical linguistics of native America*. Oxford, Oxford University
- Campbell MC, Tishkoff SA. 2008. African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annu Rev Genomics Hum Genet*. 9:403-433
- Cann RL, Brown WM, Wilson AC. 1982. Evolution of human mitochondrial DNA: a preliminary report. *Prog Clin Biol Res*. 103: 157-165
- Cann RL. 1984. Mitochondrial DNA variation in Australian aborigines: the spread of modern populations. *Acta Anthropogenet*. 1984;8(1-2):125-135

- Cann RL, Stoneking M, Wilson AC. 1987. Mitochondrial DNA and Human Evolution. *Nature* 325: 31-36
- Carrasco P. 1950. Los Otomíes: cultura e historia prehispánicas de los pueblos mesoamericanos de habla otomíama. Universidad Nacional Autónoma de México (UNAM), Instituto de Historia, México D.F., México
- Cavalli-Sforza LL, Piazza A, Menozzi P, Mountain J. 1988. Reconstruction of human evolution: bringing together genetic, archaeological, and linguistic data. *Proc Natl Acad Sci USA*. 85(16):6002-6006
- Cavalli-Sforza LL, Minch E, Mountain JL. 1992. Coevolution of genes and languages revisited. *Proc Natl Acad Sci USA*. 89(12):5620-5624
- Cavalli-Sforza LL, Piazza A, Menozzi P. 1994. *History and geography of Genes*. Princeton
- Cervantes Escandón, MT. 1978. Estructura Ocupacional y Organización Social de una comunidad Otomí del Estado de México: San Pedro Arriba, Temoaya, México. Tesis de Maestría, Escuela Nacional de Antropología e Historia (ENAH), México D.F., México
- Chakravarti A. 1999. Population genetics-making sense out of sequence. *Nat Genet*. 21: 50-60
- Chen X, Prosser R, Simonetti S, Sadlock J, Jagliello G, Schon EA. 1995. Rearranged mitochondrial genomes are present in human oocytes. *Am J Hum Genet*. 57: 239-247
- Chen Y-S, Olckers A, Schur TG, Kogelnik AM, Huoponen K, Wallace DC. 2000. mtDNA Variation in the South African Kung and Khwe-and Their Genetic Relationships to Other Africa Populations. *Am J Hum Genet*. 66: 1362-1383
- Cherni L, Fernandes V, Pereira JB, Costa MD, Goios A, Frigi S, Yacoubi-Loueslati B, Amor MB, Slama A, Amorim A, El Gaaied AB, Pereira L. 2009. Post-last glacial maximum expansion from Iberia to North Africa revealed by fine characterization of mtDNA H haplogroup in Tunisia. *Am J Phys Anthropol*. 139(2):253-260
- Chiaroni J, Underhill PA, Cavalli-Sforza LL.. 2009. Y chromosome diversity, human expansion, drift, and cultural evolution. *Proc Natl Acad Sci USA*. 106(48):20174-20179
- Coble MD, Loreille OM, Wadhams MJ, Edson SM, Maynard K, Meyer CE, Niederstätter H, Berger C, Berger B, Falsetti AB, Gill P, Parson W, Finelli LN. 2009. Mystery solved: the identification of the two missing Romanov children using DNA analysis. *PLoS One*. 4(3):e4838
- Collin Harguindeguy L. 2006. El caso de los exitosos otomíes de Temoaya. El Colegio de Tlaxcala, Tlaxcala, México
- Comas D, Calafell F, Mateu E, Pérez-Lezaun A, Bosch E, Martínez-Arias R, Clarimon J, Facchini F, Fiori G, Luiselli D, Pettener D, Bertranpetit J. 1998. Trading Genes along the Silk Road: mtDNA Sequences and the Origin of Central Asian Populations. *Am J Hum Genet* 63: 1824-1838

- Comas D, Calafell F, Benchemsi N, Helal A, Lefranc G, Stoneking M, Batzer MA, Bertranpetit J, Sajantila A. 2000. Alu insertion polymorphisms in NW Africa and the Iberian Peninsula: evidence for a strong genetic boundary through the Gibraltar Straits. *Hum Genet.* 107(4):312-319
- Comas D, Plaza S, Wells RS, Yuldaseva N, Lao O, Calafell F, Bertranpetit J. 2004. Admixture, migrations, and dispersals in Central Asia: evidence from maternal DNA lineages. *Eur J Hum Genet.* 12(6):495-504
- Cook de Lenhard C. 1981. Ciudad Netzahualcóyotl, Casa de la Cultura, Ayuntamiento de Netzahualcóyotl, México. México
- Crawford MH. 1992. *Antropología biológica de los indios americanos.* Ed. Mapfre. Madrid
- Crawford MH. 1998. *The origins of native Americans. Evidence from Anthropological Genetics.* Cambridge, Cambridge University Press.
- Curat M, Excoffier L. 2005. The effect of the Neolithic expansion on European molecular diversity. *Proc Biol Sci.* 272(1564):679-688
- Dansie A. 1997. Early Holocene burials in Nevada: overview of localities, research and legal issues. *Nev Hist Soc Q* 40:4–14
- Davies W. 2001. A very model of a modern human industry: new perspectives on the origins and spread of the Aurignacian in Europe. *Proceeding of the Prehistoric Society.* 67: 195-217
- De Laet SL (Ed). 1994. *The history of Humanity.* Routledge. Londres. Vol I.256-258
- DeGiorgio M, Jakobsson M, Rosenberg NA. 2009. Out of Africa: modern human origins special feature: explaining worldwide patterns of human genetic variation using a coalescent-based serial founder model of migration outward from Africa. *Proc Natl Acad Sci USA.* 106(38):16057-16062
- Denaro M, Blanc H, Johnson MJ, Chen KH, Wilmsen F, Cavalli-Sforza LL, Wallace DC. 1981. Ethnic variation in HpaI endonuclease cleavage patterns of human mitochondrial DNA. *Proc Nat Acad Sci* 78: 5768-5772
- Derenko MV, Grzybowski T, Malyarchuk BA, Czarny J, Miścicka-Sliwka D, Zakharov IA. 2001. The presence of mitochondrial haplogroup x in Altaians from South Siberia. *Am J Hum Genet.* 69(1):237-241
- Derenko MV, Grzybowski T, Malyarchuk BA, Dambueva IK, Denisova GA, Czarny J, Dorzhu CM, Kakpakov VT, Miścicka-Sliwka D, Woźniak M, Zakharov IA. 2003. Diversity of mitochondrial DNA lineages in South Siberia. *Ann Hum Genet.* 67(Pt5):391-411
- Derenko M, Malyarchuk B, Grzybowski T, Denisova G, Dambueva I, Perkova M, Dorzhu C, Luzina F, Lee HK, Vanecek T, Villems R, Zakharov I. 2007. Phylogeographic analysis of mitochondrial DNA in northern Asian populations. *Am J Hum Genet.* 81(5):1025-1041
- Di Rienzo A, Wilson A C. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc Natl Acad Sci USA,* 88: 1597- 1601

- Dillehay TD, Collins MB. 1988. Early cultural evidence from Monte Verde in Chile. *Nature*. 332: 150–152
- Dillehay, T.D. 2003. Palaeoanthropology: tracking the first Americans. *Nature*. 425(6953): 23-24
- Dillehay TD. 2004. Monte Verde. Un asentamiento humano del Pleistoceno tardío en el sur de Chile. Colección Serie Universitaria, LOM Ed. Universidad Austral de Chile, Santiago
- Dillehay TD, Rossen J, Maggard G, Stackelbeck K, Netherly P. 2003. Localization and possible social aggregation in the Late Pleistocene and Early Holocene on the north coast of Peru. *Quaternary Intern*. 109-110: 3-11
- Dillehay TD, Ramírez C, Pino M, Collins MB, Rossen J, Pino-Navarro JD. 2008. Monte Verde: seaweed, food, medicine, and the peopling of South America. *Science*. 320:784-786
- Dixon EJ. 2001. Human colonization of the Americas: Timing, technology and process. *Quaternary Science Reviews*, 21: 1-7
- Dornelles CL, Bonatto SL, Freitas LB, Salzano FM. 2005. Is haplogroup X present in extant South American Indians?. *Am J Phys Anthropol* 127: 439-448
- Duverger C. 2007. El primer mestizaje. La clave para entender el pasado mesoamericano. Consejo Nacional para la Cultura y las Artes, Instituto Nacional de Antropología e Historia, Taurus Historia, Universidad Nacional Autónoma de México, México
- Endicott P, Gilbert MT, Stringer C, Lalueza-Fox C, Willerslev E, Hansen AJ, Cooper A. 2003. The genetic origins of the Andaman Islanders. *Am J Hum Genet*. 72(1):178-184
- Engelman CD, Meyers KJ, Ziegler JT, Taylor KD, Palmer ND, Haffner SM, Fingerlin TE, Wagenknecht LE, Rotter JI, Bowden DW, Langefeld CD, Norris JM. 2010. Genome-wide association study of vitamin D concentrations in Hispanic Americans: The IRAS Family Study. *J Steroid Biochem Mol Biol*. Epub ahead of print
- Erlich HA. 1989. PCR technology. Principles and applications for DNA amplification. Stockton Press. Macmillan Publishers. England
- Escobar Ohmstede A. 1998. De la costa a la sierra. Las Huastecas, 1750-1900. CIESAS-INI. Colección Historia de los Pueblos Indígenas de México. México
- Eshelman JA, Malhi RS, Smith DG. 2003. Mitochondrial DNA studies of Native Americans: conceptions and misconceptions of the population prehistory of the Americas. *Evo Anth* 12: 7-18
- Excoffier L, Smouse P, Quattro J. 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131:479-491
- Excoffier L. 2004. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Mol Ecol* 13(4): 853-864

- Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 10:564-567
- Eyre-Walker A, Smith NH, Maynard Smith J. 1999. How clonal are human mitochondria?. *Proc R Soc Lond.* 266: 477-483
- Eyre-Walker A, Awadalla P. 2001. Does human mtDNA recombine? *Journal Molecular Evolution.* 53: 430-435
- Fagan BM. 2004. *The great journey: the peopling of ancient America.* Gainesville, FL. University Press of Florida
- Fagundes NJ, Bonatto SL, Callegari-Jaques SM, Salzano FM. 2002 Genetic, geographic, and linguistic variation among South American Indians: possible sex influence. *Am J Phys Anthropol* 117: 68-78
- Fagundes NJ, Ray N, Beaumont M, Neuenschwander S, Salzano FM, Bonatto SL, Excoffier LE. 2007. Statistical evaluation of alternative models of human evolution. *Proc Natl Acad Sci USA* 104: 17614-17619
- Fagundes NJ, Kanitz R, Eckert R, Valls AC, Bogó MR, Salzano FM, Smith DG, Silva WA Jr, Zago MA, Ribeiro-dos-Santos AK, Santos SE, Petzl-Erler ML, Bonatto SL. 2008a. Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. *Am J Hum Genet.* 82(3):583-592
- Fagundes NJ, Kanitz R, Bonatto SL. 2008b. A reevaluation of the Native American mtDNA genome diversity and its bearing on the models of early colonization of Beringia. *PLoS One.* 3(9):e3157
- Falk T. 2004. *Science* 305: 590
- Fenton M, Geiselhart S, Rohling EJ, Hemleben C. 2000. A planktonic zones in the Red Sea. *Marine Micropaleontology.* 40(3): 277-294
- Flores C, Maca-Meyer N, González AM, Oefner PJ, Shen P, Pérez JA, Rojas A, Larruga JM, Underhill PA. 2004. Reduced genetic structure of the Iberian peninsula revealed by Y-chromosome analysis: implications for population demography. *Eur J Hum Genet.* 12(10):855-863
- Ford EB. 1966. Genetic Polymorphism. *Proc R Soc Biol Sci* 164: 350-361
- Forster P, Harding R, Torroni A, Bandelt HJ. 1996. Origin and evolution of Native American mtDNA variation: A reappraisal. *Am J Hum Genet.* 59: 935-945
- Forster P, Torroni A, Renfrew C, Röhl A. 2001 Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution. *Molecular Biology Evolution.* 18: 1864–1881.
- Fuselli S, Tarazona-Santos E, Dupanloup I, Soto A, Luiselli D, Pettener D. 2003. Mitochondrial DNA diversity in South America and the genetic history of Andean highlanders. *Mol Biol Evol* 20:1682-1691
- Gamble C. 1999. *The Paleolithic Societies of Europe.* Cambridge University Press

- Garrigan D, Hammer MF. 2006. Reconstructing human origins in the genomic era. *Nat Rev Genet.* 7(9):669-680
- Genographic Project. nnhgtool.nationalgeographic.com/classify/.
- Gibbons A. 1996. The peopling of the Americas. *Science.* 274: 31-33
- Gilbert MT, Jenkins DL, Götherstrom A, Naveran N, Sanchez JJ, Hofreiter M, Thomsen PF, Binladen J, Higham TF, Yohe RM 2nd, Parr R, Cummings LS, Willerslev E. 2008. DNA from pre-Clovis human coprolites in Oregon, North America. *Science* 320:786-789
- Giles RE, Blanc H, Cann HM, Wallace DC. 1980. Maternal inheritance of human mitochondrial DNA. *Proc Natl Acad Sci USA.* 77: 6715-6719
- Gill, P., Ivanov, P.I., Kimpton, C., Piercy, R., Benson, N., Tully, G., Evett, I., Hagelberg, E., Sullivan, K. 1994. Identification of the remains of the Romanov family by DNA analysis. *Nat Genetics* 6:130-136
- Ginther C, Corach D, Penacino GA, Rey JA, Carnese FR, Hutz MH, Anderson A. 1993. Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial DNA sequence, nuclear variation and allele frequencies of several nuclear genes. In: Pena SDJ, Chakraborty R, Epplen J. T., Jeffrey A. J. (eds) *DNA fingerprinting: state of the science*, Birkhauser, Basel, pp 211-219
- Goebel T, Derevianko A, Petrin V. 1993. Dating the Middle-to-Upper Paleolithic transition at Kara Bom. *Curr Anthropol.* 34: 452-458
- Goebel T, Waters MR, Dikova M. 2003. The archaeology of Ushki Lake, Kamchatka, and the Pleistocene peopling of the Americas. *Science.* 301: 501-505
- Goebel T, Waters MR, O'Rourke DH. 2008 The late Pleistocene dispersal of modern humans in the Americas. *Science* 319: 1497-1502
- Goicoechea AS, Carnese FE, Dejean C, Avena SA, Weimer TA, Franco MHL, Callegari-Jacques SM, Estalote AC, Simões MLMS, Palatnik M, Salzano FM. 2001. Genetic relationships between Amerindian populations. *Am J Phys Anthropol* 115:133-143
- Gokey NG, Cao Z, Pak JW, Lee D, McKiernan SH, McKenzie D, Weindrich R, Aiken JM. 2004. Molecular analyses of mtDNA deletion mutations in microdissected skeletal muscle fibers from aged rhesus monkeys. *Aging Cell.* 3(5):319-326
- Goldberg P, Berna F, Macphail RI. 2009. Comment on "DNA from pre-Clovis human coprolites in Oregon, North America". *Science* 325:148
- Goldberg T. 1996. *Genetics and biogeography of East African chimpanzees, Pan troglodytes schweinfurthii*. Cambridge, USA: Harvard University
- Gonzalez S, Jiménez-López JC, Hedges R, Huddart D, Ohman JC, Turner A, Pompa y Padilla JA. 2001. Earliest humans in the Americas: new evidence from México. *J Hum Evol.* 44(3):379-487
- González-José R, González-Martín A, Hernández M, Pucciarelli HM, Sardi M, Rosales A, Van Der Molen S. 2003. Craniometric evidence for Palaeoamerican survival in Baja California. *Nature.* 425(6953):62-65

- González-José R, Neves W, Lahr MM, González S, Pucciarelli H, Hernández Martínez M, Correal G. 2005. Late Pleistocene/Holocene craniofacial morphology in Mesoamerican Paleoindians: Implications for the peopling of the New World. *Am J Phys Anthropol* 128: 772–780
- González-José R, Martínez-Abadías N, González-Martín A, Bautista-Martínez J, Gómez-Valdés J, Quinto M, Hernández M. 2007. Detection of a population replacement at the Classic-Postclassic transition in Mexico. *Proc Biol Sci*. 274(1610):681-688
- González-José R, Bortolini MC, Santos FR, Bonatto SL. 2008. The peopling of America: craniofacial shape variation on a continental scale and its interpretation from an interdisciplinary view. *Am J Phys Anthropol*. 137(2):175-187
- González-Martín A, Gorostiza A, Rangel-Villalobos H, Acunha V, Barrot C, Sánchez C, Ortega M, Gené M, Calderón R. 2008. Analyzing the genetic structure of the Tepehua in relation to other neighbouring Mesoamerican populations. A study based on allele frequencies of STR markers. *Am J Hum Biol*. 20(5):605-613
- González-Olivier A, Márquez-Morfín L, Jiménez JC, Torre-Blanco A. 2001. Founding amerindian mitochondrial DNA lineages in ancient Maya from Xcaret, Quintana Roo. *Am J Phys Anthropol*. 116: 230-235
- Goodman M, Bailey WJ, Hayasaka K, Stanhope MJ, Slightom J, Czelusniak J. 1994. Molecular evidence on primate phylogeny from DNA sequences. *Am J Phy Anthropol*. 94:3-24
- Gordon RG Jr. 2005. *Ethnologue: Languages of the World*. Fifteenth edition. Summer Institute of Linguistics International, Dallas, Texas. Online version www.ethnologue.com
- Gorostiza A, González-Martín A, Ramírez CL, Sánchez C, Barrot C, Ortega M, Huguet E, Corbella J, Gené M. 2007. Allele frequencies of the 15 AmpF/Str Identifier loci in the population of Metztlán (Estado de Hidalgo), México. *Forensic Sci Int*. 166(2-3):230-232
- Goudet J, Raymond M, de Meeüs T, Rousset F. 1996. Testing differentiation in diploid populations. *Genetics* 144: 1933-1940
- Green LD, Derr JN, Knight A. 2000. MtDNA affinities of the peoples of north-central Mexico. *Am J Hum Genet* . 66:989-998
- Green RE, Malaspina AS, Krause J, Briggs AW, Jonson PL, Uhler C, Meyer M, Good JM, Maricic T, Stenzel U, Prüfer K, Siebauer M, Burbano HA, Ronan M, Rothberg JM, Egholm M, Rudan P, Brajkovic D, Kucan Z, music I, Wikstrom M, Laakkonen L, celso J, Slatkin M, Paabo S. 2008. A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* 134 (3): 416-426
- Greenberg BD, Newbold JE, Sugino A. 1983. Intraspecific nucleotide sequence variability surrounding the origin of replication in human mitochondrial DNA. *Gene* 21(1-2): 33-49
- Greenberg JH, Turner CG, Zegura SL. 1986. The settlement of the Americas: A comparison of the linguistic, dental, and genetic evidence. *Current Anthropology*, 27(5): 477-497

- Greenberg JH. 1987. *Language in the Americas*. Stanford University Press, Stanford CA
- Grün R. 2006. Direct dating of human fossils. *Am J Phys Anthropol. Suppl* 43:2-48
- Guardado-Estrada M, Juarez-Torres E, Medina-Martinez I, Wegier A, Macías A, Gomez G, Cruz-Talonia F, Roman-Bassaure E, Piñero D, Kofman-Alfaro S, Berumen J. 2009. A great diversity of Amerindian mitochondrial DNA ancestry is present in the Mexican mestizo population. *J Hum Genet.* 54(12):695-705
- Guo X, Popadin KY, Markuzon N, Orlov YL, Kraytsberg Y, Krishnan KJ, Zsurka G, Turnbull DM, Kunz WS, Khrapko K. 2010. Repeats, longevity and the sources of mtDNA deletions: evidence from 'deletional spectra'. *Trends Genet.* 8:340-433
- Hagelberg E, Goldman N, Lio P, Whelan S, Schiefenhövel W, Clegg JB, Bowden DK. 1999. Evidence for mitochondrial DNA recombination in a human population of island Melanesia. *Proceedings Royal Society London B Biological Science.* 266:485–492
- Hamilton MJ, Buchanan B. 2007. Spatial gradients in Clovis-age radiocarbon dates across North America suggest rapid colonization from the north. *Proc Natl Acad Sci U S A.* 104: 15625–15630
- Harpending HC, Batzer MA, Gurven M, Jorde LB, Rogers AR, Sherry ST. 1998. Genetic traces of ancient demography. *Proc Natl Acad Sci U S A.* 95(4):1961-1967
- Haynes G, Anderson DG, Ferring CR, Fiedel SJ, Grayson DK, Haynes CV Jr, Holliday VT, Huckell BB, Kornfeld M, Meltzer DJ, Morrow J, Surovell T, Waguespack NM, Wigand P, Yohe RM. 2007. 2nd. Comment on "Redefining the age of Clovis: implications for the peopling of the Americas". *Science.* 317(5836):320
- He M, Gitschier J, Zerjal T, de Knijff P, Tyler-Smith C, Xue Y. 2009. Geographical affinities of the HapMap samples. *PLoS One.* 4(3):e4684
- Hekking E. 1995. *El otomí de Santiago Mexquititlán: desplazamiento lingüístico, préstamos y cambios gramaticales*. Institute for Functional Research into Language and Language Use, Instituut voor Functioneel Onderzoek van Taal en Taalgebruik, Amsterdam
- Henshilwood CS, D'Errido F, Yates R, Jacobs Z, Tribolo C, Duller GA, Mercier N, Sealy JC, Valladas H, Watts I, Wintle AG. 2002. Emergence of modern human behaviour: Middle Stone Age engravings from South Africa. *Science.* 295: 1278-1280
- Hernández Montes M, Heiras Rodriguez CG. 2006. *Tepehuas*. Colección Pueblos del México contemporáneo. Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CoNaDePI). Programa de las Naciones Unidas para el desarrollo. México
- Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, Howell N. 2002a. Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *American Journal Human Genetics.* 70(5):1152-71.

- Herrnstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE, Howell N. 2002b. Erratum. *American Journal Human Genetics*. 71(2):448-9.
- Hervás y Panduro L. 1800-1805. Catálogo de las lenguas de las naciones conocidas y numeración, división y clases de estas según la diversidad de sus idiomas y dialectos. 6 volúmenes. Imprenta de la Administración del Real Arbitrio de Beneficencia. Madrid
- Hey J. 2005. On the number of new world founders: a population genetic portrait of the peopling of the Americas. *PLoS Biol* 3:e193
- Heyer E, Balaesque P, Jobling MA, Quintana-Murci L, Chaix R, Segurel L, Aldashev A, Hegay T. 2009. Genetic diversity and the emergence of ethnic groups in Central Asia. *BMC Genet*. 1;10:49
- Higuchi R, Von Beroldingen C, Sensabough G, Erlich G. 1988. DNA typing from single hairs. *Nature* 332: 543-546
- Hill C, Soares P, Mormina M, Macaulay V, Meehan W, Blackburn J, Clarke D, Raja JM, Ismail P, Bulbeck D, Oppenheimer S, Richards M. 2006. Phylogeography and ethnogenesis of aboriginal Southeast Asians. *Mol Biol Evol*. 23(12):2480-2491
- Hoffecker JF, Powers WR, Goebel T. 1993. The colonization of Beringia and the peopling of the new world. *Science*. 259 :46-53
- Holland MM, Parsons TJ. 1999. Mitochondrial DNA sequence analysis, validation and use in forensic casework. *Forensic Sci Rev* 11: 21-50
- Holt BM, Formicola V. 2008. Hunters of the Ice Age: The biology of Upper Paleolithic people. *Am J Phys Anthropol*. Suppl 47:70-99
- Hope M. 2006. Pimas. Colección Pueblos del México contemporáneo. Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CoNaDePI). Programa de las Naciones Unidas para el desarrollo. México
- Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima K. 1993. Peopling of the Americas, Founded by Four Mayor Lineages of Mitochondrial DNA. *Mol Biol Evol*. 10(1): 23-47
- Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N. 1995. Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. *Proc Natl Acad Sci U.S.A.* 92(2): 532-536
- Howell N, Kubacka I, Mackey DA. 1996. How rapidly does the human mitochondrial genome evolve? *Am J Hum Genet*. 59(3):501-509
- Howell N, Smejkal CB, Mackey DA, Chinnery PF, Turnbull DM, Herrnstadt C. 2003. The pedigree rate of sequence divergence in the human mitochondrial genome: there is a differences between phylogenetic and pedigree rates. *Am J Hum Genet*. 72: 659-670
- Hrdlicka A. 1925. The origin and antiquity of the American Indian. En: *Ann. Rep. Smith. Inst. for 1923*. Washington. pp. 481-494

- Hubbe M, Neves WA, Harvati K. 2010. Testing evolutionary and dispersion scenarios for the settlement of the New World. *PLoS ONE*. 5(6): e11105
- Humboldt A. 1810. *Vues des cordillères et monuments des peuples indigènes de l'Amérique*. París.
- Humboldt WV. 1960. *Ueber die Verschiedenheit des menschlichen Sprachbaues und ihre Einfluss auf die geistige Entwicklung des Menschengeschlechtes*. Ferd. Dümmlers. Verlag, Bonn
- Hunley K, Long JC. 2005. Gene flow across linguistic boundaries in native North American populations. *Proc Natl Acad Sci USA*. 102:1312-1317
- Hutchison CA, Newbold JE, Potter SS, Edgell MH. 1974. Maternal inheritance of mammalian mitochondrial DNA. *Nature* 251: 536-538
- Ibarra-Rivera L, Mirabal S, Regueiro MM, Herrera RJ. 2008. Delineating genetic relationships among the Maya. *Am J Phys Anthropol*. 135(3):329-347
- Ingman M, Kaessmann H, Paabo S, Gyllensten U. 2000. Mitochondrial genome variation and the origin of modern humans. *Nature* 408 (6813): 708-713
- Ingman M, Gyllensten U. 2001. Analysis of complete human mtDNA genome: methodology and inferences for human evolution. *J Hered* 92:454-461
- Ingman M, Gyllensten U. 2003. Mitochondrial genome variation and the origin of modern humans. *Nature*. 408: 708-713
- Instituto Nacional de Estadística, Geografía e Informática (INEGI). 1990. X Censo General de población y vivienda. México
- Instituto Nacional de Estadística, Geografía e Informática (INEGI). 2000. XII Censo General de población y vivienda. México
- Instituto Nacional de Estadística, Geografía e Informática (INEGI). 2005. XIII Censo General de población y vivienda. México
- Instituto Nacional de Lenguas Indígenas (INALI). 2007. *Catálogo de las lenguas indígenas nacionales: Variantes lingüísticas de México con sus autodenominaciones y referencias geoestadísticas*. México
- Instituto Nacional de Lenguas Indígenas (INALI). 2008. *Catálogo de las Lenguas Indígenas Nacionales: Variantes Lingüísticas de México con sus autodenominaciones y referencias geoestadísticas*. México
- Iriondo M, Barbero MC, Manzano C. 2003. DNA polymorphism detect ancient barriers to gen flow in Basques. *Am J Phys Anthropol*. 122: 73- 84
- Jackson BA, Wilson JL, Kirbah S, Sidney SS, Rosenberger J, Bassie L, Alie JAD, McLean DC, Garvey WT, Ely B. 2005. Mitochondrial DNA genetic diversity among four ethnic groups in Sierra Leone. *Am J Phys Anthropol* 128: 156-163
- Jantz RL, Owsley DW. 1997. Pathology, taphonomy and cranial morphometrics of the Spirit Cave mummy. *Nev Hist Soc Q* 40:62-84

- Jantz RL, Owsley DW. 2001. Variation Among Early North American Crania. *Am J Phy Anthropol* 114:146–155
- Jenks AE. 1937. Minnesota's Browns Valley Man and associated burial artifacts. Menasha, WI: *Memoirs of the American Anthropological Association*, no 49
- Jobling MA, Hurles ME, Tyler-Smith C. 2004. Into new found lands. Human evolutionary genetics. *Origins, Peoples & Diseases*. New York: Garland Science. pp 339-372
- Jorde LB, Bamshad M. 2000. Questioning evidence for recombination in human mitochondrial DNA. *Science* 288:1931a
- Jorde LB, Rogers AR, Bamshad M, Watkins WS, Kralowiak P, Sung S, Kere J, Haspending HC. 1997. Microsatellite diversity and the demographic history of modern humans. *Proc Natl Acad Sci USA*. 94:3100-3103
- Jorde LB, Watkins WS, Bamshad MJ, Dixon ME, Ricker CE, Seielstad MT, Batzer MA. 2000. The Distribution of Human Genetic Diversity: A Comparison of Mitochondrial, Autosomal and Y-Chromosome Data. *Am J Hum Genet* 66: 979-988
- Jorde LB, Watkins WS, Bamshad MJ. 2001. Population genomics: a bridge from evolutionary history to genetic medicine. *Hum Mol Genet* 10: 2199-2207
- Juárez-Cedillo T, Zuñiga J, Acuña-Alonzo V, Pérez-Hernández N, Rodríguez-Pérez JM, Barquera R, Gallardo GJ, Sánchez-Arenas R, García-Peña Mdel C, Granados J, Vargas-Alarcón G. 2008. Genetic admixture and diversity estimations in the Mexican Mestizo population from Mexico City using 15 STR polymorphic markers. *Forensic Sci Int Genet*. 2(3):e37-39
- Karafet TM, Zegura SL, Posukh O, Osipova L, Bergen A, Long J, Goldman D, Klitz W, Harihara S, de Knijff P, Wiebe V, Griffiths RC, Templeton AR, Hammer MF. 1999. Ancestral Asian source(s) of new world Y-chromosome founder haplotypes. *Am J Hum Genet*. 64(3):817-831
- Kaufman T, Golla V. 2000. Language groupings in the New World: their reliability and usability in cross-disciplinary studies. In *America past, America present: genes and language in the Americas and beyond*, C. Renfrew, ed. (Cambridge: McDonald Institute for Archeological Research), pp. 47-57
- Kellogg S. 2005. *Weaving the Past: A History of Latin America's Indigenous Women from the Prehispanic Period to the Present*. Oxford University Press
- Kemp BM, Malhi RS, McDonough J, Bolnick DA, Eshleman JA, Rickards O, Martinez-Labarga C, Johnson JR, Lorenz JG, Dixon EJ, Fifield TE, Heaton TH, Worl R, Smith DG. 2007. Genetic analysis of early holocene skeletal remains from Alaska and its implications for the settlement of the Americas. *Am J Phys Anthropol*. 132(4):605-621.
- Kemp BM, González-Oliver A, Malhi RS, Monroe C, Schroeder KB, McDonough J, Rhatt G, Resendéz A, Peñaloza-Espinosa RI, Buentello-Malo L, Gorodesky C, Smith DG. 2010. Evaluating the Farming/Language Dispersal Hypothesis with genetic variation exhibited by populations in the Southwest and Mesoamerica. *Proc Natl Acad Sci USA*. 107(15):6759-6764

- Kim HS, Takenaka O. 1996. A comparison of TSPY genes from Y-chromosomal DNA of the great apes and humans: sequence, evolution, and phylogeny. *Am J Phys Anthropol.* 100:301-309
- Kimura M. 1968. Evolutionary Rate at the Molecular Level. *Nature* 217: 624-626
- King RJ, Ozcan SS, Carter T, Kalfoğlu E, Atasoy S, Triantaphyllidis C, Kouvatsi A, Lin AA, Chow CE, Zhivotovsky LA, Michalodimitrakis M, Underhill PA. 2008. Differential Y-chromosome Anatolian influences on the Greek and Cretan Neolithic. *Ann Hum Genet.*72(Pt 2):205-214
- Kirchhoff P. 1943. Mesoamérica: sus límites geográficos, composición étnica y caracteres culturales. *Acta Americana* 1: 92-107
- Kirchhoff P. 1952. Mesoamerica: Its geographic limits, ethnic composition and cultural characteristics. *Heritage of Conquest*, ed Tax S (The Free Press Publishers. Traslated by Norman McQuown from *Acta Americana*, Vol I, No I, 1943, Gleonce Illinois), pp 17-31
- Kitchen A, Miyamoto MM, Mulligan CJ. 2008. A three stage colonization model for the peopling of the Americas. *PLoS One* 3:e1596
- Kivisild T, Kaldma K, Metspalu M, Parik J, Papiha S, Villems R. 1999. The Place of the Indian mtDNA Variants in the Global Network of Maternal Lineages and the Peopling of the Old World. In: *Genomic Diversity*. Edited by Deka R, Papih SS Kluwer/Academic/Plenum Publishers. 135-152
- Kivisild T, Rootsi S, Metspalu M, Mastana S, Kaldma K, Parik J, Metspalu E, Adojaan M, Tolk HV, Stepanov V, Gölge M, Usanga E, Papiha SS, Cinniöglu C, King R, Cavalli-Sforza L, Underhill PA, Villems R. 2003. The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *Am J Hum Genet.* 72(2):313-332
- Kivisild T, Reidla M, Metspalu E, Rosa A, Brehm A, Pennarun E, Parik J, Geberhiwot T, Usanga E, Villems R. 2004. Ethiopian mitochondrial DNA heritage: tracking gene flow across and around the gate of tears. *Am J Hum Genet.* 75(5):752-770
- Kivisild T, Shen P, Wall DP, Do B, Sung R, et al., 2006. The role of selection in the evolution i;of human mitocohondrial genomes. *Genetics* 172:373-387
- Kolman CJ, Bermingham E, Cooke R, Ward RH, Arias TD, Guionneau-Sinclair F. 1995 Reduced mtDNA diversity in the Ngöbe amerinds of Panama. *Genetics.* 140: 275-283
- Kolman CJ, Bermingham E. 1997. Mitochondrial and nuclear diversity in the Chocó and Chibcha amerinds of Panama. *Genetics.* 147: 1289-1302
- Lahr MM. 1995. Patterns of modern human diversification: implications for Amerindian origins. *Yearb. Phys. Anthropol.* 38:163-198
- Lahr MM, Foley R. 1998. Towards a theory of modern human origins: geography, demography, and diversity in recent human evolution. *Yearbook of Physical Anthropology.* 41: 137-176

- Lalueza C, Pérez-Pérez A, Prats E, Cornudella L, Turbón D. 1997. Lack of founding amerindian mitochondrial DNA lineages in extinct aborigines from Tierra del Fuego-Patagonia. *Hum Molec Genet* 6: 41-46
- Lewis CM, Tito RY, Lizarraga B and Stone A. 2004. Land, language, and loci: mtDNA in native Americans and the genetic history of Peru. *Am J Phys Anthropol* 127: 351-360
- Lewis MP. 2009. *Ethnologue: Languages of the World*. Sixteenth edition. Summer Institute of Linguistics International, Dallas, Texas. Online version www.ethnologue.com
- Lewontin RC. 1972. Uses of knowledge. *Science*. 177:386
- Librado P, Rozas, J. 2009. A software for comprehensive analysis of DNA polymorphism data. *Bioinformatica*. 25: 1451-1452
- Lisker R. 1981. La estructura genética de la población mexicana. Aspectos médicos y antropológicos. Ed. Salvat Mexicana, México
- Lisker R, Ramírez E, Babinsky V. 1996. Genetic Structure of Autochthonous Populations of Meso-America: Mexico. *Hum Biol.* 6 (3), 395-404
- Livshits G, Nei M. 1990. Relationships between Intrapopulation and Interpopulation Genetic Diversity in Man. *Ann Hum Biol.* 17:501-513
- Loehr J, Worley K, Grapputo A, Carey J, Veitch A, Coltman DW. 2006. Evidence for cryptical glacial refugia from North American mountain sheep mitochondrial DNA. *J Evol Biol* 19, 419-430
- López Austin A, López Luján L. 1996. El pasado indígena. Fondo de Cultura Económica. México
- López Luján L, Manzanilla L. 1989. Atlas histórico de Mesoamérica. Ediciones Larousse, Colección Referencias. México
- Lorenz JG, Smith DG. 1996. Distribution of four founding mtDNA haplogroups among Native North Americans. *Am J Phys Anthropol.* 101(3):307-323
- Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, Bonnè-Tamir B, Sykes B, Torroni A. 1999. The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am J Hum Genet* 64:232-249
- Macaulay V, Hill C, Achilli A, Rengo C, Clarke D, Meehan W, Blackburn J, Semino O, Scozzari R, Cruciani F et al., 2005. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 308: 1034-1036
- Majid Z. 1998. Radiocarbon dates and culture sequence in the Lenggong Valley and beyond. *Malaysia Museums Journal.* 34: 241-249
- Malhi RS, Smith DG. 2002. Haplogroup X confirmed in prehistoric North America. *Am J Phys Anthropol* 119:84-86
- Malhi RS, Mortensen HM, Eshleman JA, Kemp BM, Lorenz JG, Kaestle FA, Johnson JR, Gorodezky C, Smith DG. 2003. Native American mtDNA prehistory in the American Southwest. *Am J Phys Anthropol.* 120(2):108-124

- Malhi RS, Cybulski JS, Tito RY, Johnson J, Harry H, Dan C. 2010. Brief communication: mitochondrial haplotype C4c confirmed as a founding genome in the Americas. *Am J Phys Anthropol.* 141(3):494-497
- Manfredi G, Thyagarajan D, Papadopoulou LC, Pallotti F, Schon EA. 1997. The fate of human sperm-derived mtDNA in somatic cells. *American Journal Human Genetics.* 61:953–960
- Mann CC. 2007. Archaeology. Clovis technology flowered briefly and late, dates suggest. *Science.* 315(5815):1067
- Manni F, Barrai I. 2001. Genetic structures and linguistic boundaries in Italy: a microregional approach. *Hum Biol.* 73(3):335-347
- Manni F, Guérard E, Heyer E. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by “Monmonier’s algorithm”. *Human Biology,* 76(2): 173-190
- Manrique Castañeda L 1988a. Atlas cultural de México: lingüística. Secretaría de Educación Pública-Instituto Nacional de Antropología e Historia (INAH) Editorial Planeta México. Colección Atlas Cultural de México. México DF, México
- Manrique Castañeda L. 1998b. Lenguas indígenas. Enciclopedia de México. 2a ed. Secretaría de Educación Pública. México DF, México
- Manrique Castañeda L, Lastra Y, Bartholomew D. 1994-95. Panorama de los estudios de las lenguas indígenas de México. Quito, Abya-Yala
- Manrique Castañeda L. 1995. La población indígena mexicana. Instituto Nacional de Estadística, Geografía e Informática (INEGI), Instituto Nacional de Antropología e Historia (INAH), Universidad Nacional Autónoma de México (UNAM). México DF, México
- Mao X, Bigham AW, Mei R, Gutierrez G, Weiss KM, Brutsaert TD, Leon-Velarde F, Moore LG, Vargas E, McKeigue PM, Shriver MD, Parra EJ. 2007. Genome-wide admixture mapping panel for Hispanic/Latino Populations. *Am J Hum Genet* 80:1171- 1178
- Margulis L. 1981. Symbiosis in Cell evolution: Life and its environment on the Early Earth. W.H Freeman & Co., New York.
- Margulis L, Sagan D. 1987. Microcosmos: the universe within us reveals evolution’s secrets. *Bostonia* 61: 55-58
- Martínez-Cortés G, Nuño-Arana I, Rubi-Castellanos R, Vilchis-Dorantes G, Luna-Vázquez A, Coral-Vázquez RM, Canto-Cetina T, Salazar-Flores J, Muñoz-Valle JF, Sandoval-Mendoza K, López Z, Gamero-Lucas JJ, Rangel-Villalobos H. 2010. Origin and genetic differentiation of three Native Mexican groups (Purépechas, Triquis and Mayas): Contribution of CODIS-STRs to the history of human populations of Mesoamerica. *Ann Hum Biol.* Epub ahead of print
- Martinez-Fierro ML, Beuten J, Leach RJ, Parra EJ, Cruz-Lopez M, Rangel-Villalobos H, Riego-Ruiz LR, Ortiz-Lopez R, Martinez-Rodriguez HG, Rojas-Martinez A. 2009. Ancestry informative markers and admixture proportions in northeastern Mexico. *J Hum Genet.* 54(9):504-509

- McKenney K, Spriggs T, Reinhard J, Chavez JA, Lorenzi HA, Caler E, Shumway M, Sapiro V, Fraser C. 2007. Mitochondrial genome sequencing of an Incan pre-Columbian mummy recovered from a Peruvian glacier ('Juanita, the Ice Maiden of Ampato'). Unpublished. GenBank Accession Numbers EF660742, EF660743
- Meltzer DJ, Michael BC. 1987. Prehistoric water wells on the Southern High Plains: Clues to Altithermal climate. *Journal of Field Archaeology*. 14: 9-28
- Merriwether DA, Rothhammer F, Ferrell RE. 1994. Genetic variation in the New World: ancient teeth, bone, and tissue as sources of DNA. *Experientia* 50: 592-601
- Merriwether DA, Rothhammer F, Ferrell RE. 1995. Distribution of the four founding lineages haplotypes in Natives Americans suggests a single wave of migration for the New World. *Am J Phys Anthropol* 98: 411-430
- Merriwether DA, Hall WW, Vahlne A, Ferrell RE. 1996. MtDNA variation indicates Mongolia may have been the source for the founding population for the New World. *Am J Hum Genet* 59: 204-212
- Merriwether DA, Reed DM, Ferrell RE. 1997. Ancient and contemporary mitochondrial DNA variation in the Maya. In: Whittington SL, Reed DM, editors. *Bones of the Maya: studies of ancient skeletons*. Washington, DC: Smithsonian Institute Press. p208-217
- Merriwether DA, Hodgson JA, Friedlaender FR, Allaby R, Cerchio S, Koki G, Friedlaender JS. 2005. Ancient mitochondrial M haplogroups identified in the Southwest Pacific. *Proc Natl Acad Sci USA*. 102(37):13034-13039
- Mesa NR, Mondragón MC, Soto ID, Parra MV, Duque C, Ortiz-Barrientos D, García LF, Velez ID, Bravo ML, Múnera JG, Bedoya G, Bortoloni MC, Ruíz-Linares A. 2000. Autosomal mtDNA, and Y-Chromosome Diversity in Amerinds: Pre- and Post-Columbian Patterns of Gene Flow in South America. *Am J Hum Genet* 67: 1277-1286
- Metspalu E, Kivisild T, Kaldma K, Parik Y, Reidla M, Tambets K, Villems R. 1999. The trans-caucasus and the expansion of the caucasoid-specific human mitochondrial DNA. In: *Genomic Diversity: Applications in Human Population Genetics*. Ed Papiha, D. & Chakraborty, K. Academic / Plenum Publishers, New York. 121-133
- Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, Brandon M, Easley K, Chen E, Brown MD, Sukernik RI, Olckers A, Wallace DC. 2003. Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci U S A*. 100(1):171-176
- MITOMAP: A human mitochondrial genome database. <http://www.mitomap.org/>
- MITOWEB: www.mitomap.org/Main/WebHome
- Moreno Alcántara B, Garret Ríos MG, Fierro Alonso UJ. 2006. Otomíes del valle del Mezquital. Colección Pueblos del México contemporáneo. Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CoNaDePI). Programa de las Naciones Unidas para el desarrollo. México
- Monmonier M. 1973. Maximum-difference barriers: an alternative numerical regionalization method. *Geographical Analysis*. 3: 245- 261

- Morral N, Bertranpetit J, Estivill X, Nunes V, Casals T, Giménez J, Reis A, Varon-Mateeva R, Macek M Jr, Kalaydjieva L, Angelicheva D, Dancheva R, Romeo G, Russo MP, Garnerone S, Restagno G, Ferrari M, Magnani C, Claustres M, Desgeorges M, Schwartz M, Schwarz M, Dallapiccola B, Novelli G, Ferec C, Arce M, Nemeti M, Kere J, Anvret M, Dahl N, Kadasi L. 1994. The origin of the major cystic fibrosis mutation (delta F508) in European populations. *Nat Genet.* 7(2):169-175
- Muhammad JS. 1995. No longer invisible: Afro-Latin Americans today. Mexico and Central America. In: *Minority Rights Group (ed) Minority Rights Publications, London, pp 163-180*
- Mukherjee N, Nebel A, Oppenheim A, Majumder PP. 2001. High-resolution analysis of Y-chromosomal polymorphisms reveals signatures of population movements from Central Asia and West Asia into India. *J Genet.* 80(3):125-135
- Mulligan CJ, Kitchen A, Miyamoto MM. 2008. Updated three-stage model for the peopling of the Americas. *PLoS One.* 17;3(9):e3199
- Mullis KB, Faloona FA. 1987. Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. *Meth Enzymol* 155: 335-350
- Muñoz Orozco M. Sin fecha. Conflicto Pima: formas de organización para la producción y la cultura
- Myster SMT, O'Connell B. 1997. Bioarcheology of Iowa, Wisconsin, and Minnesota. In: Owsley DW, Rose JC, editors. *Bioarcheology of the North Central United States.* Fayetteville, AR: Arkansas Archeological Survey Research Series. 147- 239
- Neckelmann N, Li K, Wade RP, Shuster R, Wallace DC. 1987. cDNA sequence of a human skeletal muscle ADP/ATP translocator: lack of a leader peptide, divergence from a fibroblast translocator cDNA, and coevolution with mitochondrial DNA genes. *Proc Natl Acad Sci U S A.* 84(21):7580-1584
- Nei, M. 1987 *Molecular Evolutionary Genetics.* Columbia University Press, New York, NY, USA
- Nei M. 1977. F-Statistics and analysis of gene diversity in subdivided populations. *Ann Hum Genet.* 41:225-233
- Nettle D. 1999. Linguistic diversity of the Americas can be reconciled with a recent colonization. *Proceedings of the National Academy of Sciences, USA.* 96, 33325-33329
- Neurath J. 2003. Huicholes. Colección Pueblos del México contemporáneo. Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CoNaDePI). Programa de las Naciones Unidas para el desarrollo. México
- Neves WA, Pucciarelli HM. 1989. Extra continental biological relationships of early South American human remains: a multivariate analysis. *Ciencia e Cultura, J Braz Assoc Adv Sc.* 41:566-575
- Neves WA, Pucciarelli HM. 1990. The origin of the first Americans: an analysis based on the cranial morphology of early South American human remains. *Am J Phys Anthropol.* 81: 274-

- Neves WA, Pucciarelli HM. 1991. Morphological affinities of the first Americans: an exploratory analysis based on early South American human remains. *J Hum Evol* 21:261-273
- Neves W, González-José R, Hubbe M, Kipnis R, Araujo A, Blasi O. 2004. Early Holocene Human Skeletal Remains from Cerca Grande, Lagoa Santa, Central Brazil, and the Origins of the First Americans. *World Archaeology*. 36: 479-502
- Neves WA, Hubbe M. 2005. Cranial morphology of early Americans from Lagoa Santa, Brazil: implications from the settlement of the New World. *Proc Natl Acad Sci USA*. 102:18309-18314
- Neves WA, Prous A, González-José R, Kipnis R, Powell J. 2003. Early Holocene Human Skeletal from Santana do Riacho, Brazil: Implications for the Settlement of the New World. *J Hum Evol*. 45: 759-782
- Neves WA, Hubbe M, Okumura MM, González-José R, Figuti L, Eggers S, De Blasis PA. 2005. A new early Holocene human skeleton from Brazil: implications for the settlement of the New World. *J Hum Evol*. 48(4):403-414
- Nichols J. 2002. The first American languages. In Jablonski NG (ed.). *The first Americans*. San Francisco CA: California Academy of Sciences. Pp 273-293
- Noonan JP, Coop G, Kudaravalli S, Smith D, Krause J, Alessi J, Chen F, Platt D, Paabo S, Pritchard JK, Rubin EM. 2006. Sequencing and Analysis of Neanderthal Genomic DNA. *Science*. 314: 1113-1118
- Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, Stephens M, Bustamante CD. 2008. Genes mirror geography within Europe. *Nature*. 456(7218): 98-101
- Ochoa L. 1989. *Huastecos y Totonacos*. COMisión NAcional para la CULTura y las Artes (CONCULTA), México
- Olszewsky DI, Dibble HL. 1994. The Zagros aurignacian. *Curr Anthropol*. 35(1): 68-75
- Oota H, Settheetham-Ishida W, Tiwawech D, Ishida T, Stoneking M. 2001. Human mtDNA and Y-chromosome variation is correlated with matrilocality versus patrilocality residence. *Nature Genetics*. 29: 20-21
- OPPENHEIMER, S. 2004. *Out of Eden. The peopling of the world*. Constable & Robinson Ltd. London.
- O'Rourke DH. 2006. Blood groups, immunoglobulins, and genetic variation. In *Handbook of North American Indians*. Vol. 3. Environment, origins, and population. D Ubelaker, ed. Washington, DC; Smithsonian Institution. Pp. 762-776
- O'Rourke DH, Raff JA. 2009. The Human Genetic History of the Americas: the final frontier. *Curr Biol*. 20(4):R202-207
- Orozco y Berra M. 1864. *Geografía de las lenguas y carta etnográfica de México*. Precedidas de un ensayo de clasificación de las mismas lenguas y de apuntes para las inmigraciones. México

- Otte M. 2004. The aurignacian in asia. In: *The Early Upper Paleolithic beyond Western Europe*. Edited by P. Jeffrey Brantingham, Steven L. Kuhn, and Kristopher W. Kerry. University of California Press.
- Owsley DW, Jantz RL. 1999. Databases for Paleoamerican skeletal biology research. In: *Bonnischsen R, editor. Who were the first Americans?* Corvallis, OR: Center for the Study of the First Americans, Oregon State University. 79–96
- Pailles RA, Whitecotton JE. 1995. The frontiers of Mesoamerica: Northern and Southern. In: *Reyman JE (ed) The Gran Chichimeca: essays of the archaeology and ethnohistory of Northern Mesoamerica*. Worldwide Archaeology series, Avebury
- Parra E, Saha N, Soemantri AG, McGarvey ST, Hundrieser J, Shriver MD, Deka R. 1999. Genetic variation at 9 autosomal microsatellite loci in Asian and Pacific populations. *Hum Biol.* 71(5):757-779
- Parsons TJ, Muniec DS, Sullivan K, Allison-Greiner R, Wilson MR, Berry DL, Holland KA, Weedn VW, Gill P, Holland MM. 1997. A high observed substitution rate in the human mitochondrial DNA control region. *Nature Genet.* 15: 363-368
- Peñaloza-Espinosa RI, Arenas-Aranda D, Cerda-Flores RM, Buentello-Malo L, González-Valencia G, Torres J, Alvarez B, Mendoza I, Flores M, Sandoval L, Loeza F, Ramos I, Muñoz L, Salamanca F. 2007. Characterization of mtDNA haplogroups in 14 Mexican indigenous populations. *Hum Biol.*79(3):313-320
- Perego UA, Achilli A, Angerhofer N, Accetturo M, Pala M, Olivieri A, Kashani BH, Ritchie KH, Scozzari R, Kong QP, Myres NM, Salas A, Semino O, Bandelt HJ, Woodward SR, Torroni A. 2009. Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. *Curr Biol.* 19(1):1-8
- Perego UA, Angerhofer N, Pala M, Olivieri A, Lancioni H, Kashani BH, Carossa V, Ekins JE, Gómez-Carballa A, Huber G, Zimmermann B, Corach D, Babudri N, Panara F, Myres NM, Parson W, Semino O, Salas A, Woodward SR, Achilli A, Torroni A. 2010. The initial peopling of the Americas: a growing number of founding mitochondrial genomes from Beringia. *Genome Res.* 20(9):1174-1179
- Pereira L, Richards M, Goios A, Alonso A, Albarrán C, Garcia O, Behar DM, Gölge M, Hatina J, Al-Gazali L, Bradley DG, Macaulay V, Amorim A. 2005. High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. *Genome Res.* 15(1):19-24
- Pereira L, Cerný V, Cerezo M, Silva NM, Hájek M, Vasíková A, Kujanová M, Brdicka R, Salas A. 2010. Linking the sub-Saharan and West Eurasian gene pools: maternal and paternal heritage of the Tuareg nomads from the African Sahel. *Eur J Hum Genet.* 18(8):915-923
- Petraglia MD, Haslam M, Fuller DQ, Boivin N, Clarkson C. 2010. Out of Africa: new hypotheses and evidence for the dispersal of *Homo sapiens* along the Indian Ocean rim. *Ann Hum Biol.* 37(3):288-311
- PHYLOTREE: www.phylotree.org/
- Pimentel F. 1874-1875. Cuadro descriptivo y comparativo de las lenguas indígenas de México, o tratado de filología mexicana. (II): Págs. 303-305

- Piña Chan R. 1971. Historia, arqueología y arte prehispánico. Fondo de Cultura Económica. México
- Piña Chan R. 1997. Un modelo de evolución social y cultural del México precolombino. Instituto Nacional de Antropología e Historia, México
- Poinar H, Fiedel S, King CE, Devault AM, Bos K, Kuch M, Debruyne R. 2009. Comment on "DNA from pre-Clovis human coprolites in Oregon, North America". *Science* 325:148
- Pompa y Padilla JA, Serrano Carreto E. 2001. Los más antiguos americanos. *Arqueología Mexicana*. 9: 36-41
- Powell JF, Neves WA. 1999. Craniofacial morphology of the first americans: pattern and process in the peopling of the New World. *Yearb Phys Anthropol*. 42:153-188
- Prager EM, Orrego C, Sage RD. 1998. Genetic variation and phylogeography of central Asian and other house mice, including a major new mitochondrial lineage in Yemen. *Genetics*. 150, 835- 861
- Przeworski M, Hudson RR, Di Rienzo A. 2000. Adjusting the focus o human variation. *Trends Genet*, 16: 296-302
- Pucciarelli HM, Neves WA, González-José R, Sardi ML, Rozzi FR, Struck A, Bonilla MY. 2006. East-West cranial differentiation in pre-Columbian human populations of South America. *Homo*. 57(2):133-150
- Pucciarelli HM, González-José R, Neves WA, Sardi ML, Rozzi FR. 2008. East-West cranial differentiation in pre-Columbian populations from Central and North America. *J Hum Evol*. 54(3):296-308
- Qin Z, Yang Y, Kang L, Yan S, Cho K, Cai X, Lu Y, Zheng H, Zhu D, Fei D, Li S, Jin L, Li H; and The Genographic Consortium. 2010. A mitochondrial revelation of early human migrations to the Tibetan Plateau before and after the last glacial maximum. *Am J Phys Anthropol*. [Epub ahead of print]
- Questa Rebolledo A, Utrilla Sarmiento B. 2006. Otomíes del norte del estado de México y sur de Querétaro. Colección Pueblos del México contemporáneo. Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CoNaDePI). Programa de las Naciones Unidas para el desarrollo. México
- Quintana-Murci L, Semino O, Bandelt HJ, Passarino G, McElreavey K, Santachiara-Benerecetti AS. 1999. Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa. *Nat Genet*. 23(4):437-441
- Quintana-Murci L, Chaix R, Wells RS, Behar DM, Sayar H, Scozzari R, Rengo C, Al-Zahery N, Semino O, Santachiara-Benerecetti AS, Coppa A, Ayub Q, Mohyuddin A, Tyler-Smith C, Qasim Mehdi S, Torroni A, McElreavey K. 2004. Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. *Am J Hum Genet*. 74(5):827-845
- Ramsay RM. 2003. "Evolución y diversidad en el trabajo migratorio: Gundho, un pueblo Hñahñu del Mezquital". México
- Rando JC, Pinto F, González AM, Hernández M, Larruga JM, Cabrera VM, Bandelt HJ. 1998. Mitochondrial DNA analysis of northwest African populations reveals genetic

- exchanges with European, near-eastern, and sub-Saharan populations. *Ann Hum Genet.* 6:531-550
- Rangel-Villalobos H, Rubi-Castellanos R, Morales-Vallejo ME, Molina-Araujo V, Licea-Cadena RA, Rizzo-Juárez S, Muñoz-Lozano FE, Nuño-Arana I, Muñoz-Valle JF. 2010. Admixture estimates and statistical parameters of forensic importance based on PowerPlex 16 system in Mexican-Mestizos from the States of Guanajuato (Center) and Veracruz (East). *Forensic Sci Int Genet.* 2010 4(4):271-272
- Ray N, Currat M, Excoffier L. 2003. Intra-Deme Molecular Diversity in Spatially Expanding Populations. *Mol Biol Evol* 20(1): 76-86
- Raymond M, Rousset F. 1995 An exact test for population differentiation. *Evolution* 49:1280-1283
- Redd AJ, Stoneking M. 1999. Peopling of Sahul: mtDNA variation in aboriginal Australian and Papua New Guinean populations. *Am J Hum Genet.* 65(3):808-828
- Refrew C, Boyle K (Eds). 2000. *Archaeogenetics: DNA and the populations prehistory of Europe.* MacDonald Institute for Archaeological Research. Cambridge
- Reich D, Thangaraj K, Patterson N, Price AL, Singh L. 2009. Reconstructing Indian population history. *Nature.* 24;461:489-494
- Reidla M, Kivisild T, Metspalu E, Kaldma K, Tambets K, Tolk HV, Parik J, Loogväli EL, Derenko M, Malyarchuk B, Bermisheva M, Zhadanov S, Pennarun E, Gubina M, Golubenko M, Damba L, Fedorova S, Gusar V, Grechanina E, Mikerezi I, Moisan JP, Chaventré A, Khusnutdinova E, Osipova L, Stepanov V, Voevoda M, Achilli A, Rengo C, Rickards O, De Stefano GF, Papiha S, Beckman L, Janicijevic B, Rudan P, Anagnou N, Michalodimitrakis E, Koziel S, Usanga E, Geberhiwot T, Herrnstadt C, Howell N, Torroni A, Villems R. 2003. Origin and diffusion of mtDNA haplogroup X. *Am J Hum Genet.* 73(5):1178-1190
- Relethford JH, Harpending HC. 1994. Craniometric variation, genetic theory, and modern human origins. *Am J Phys Anthropol.* 95:249-270
- Richards MB, Macaulay VA, Bandelt HJ, Sykes BC. 1998. Phylogeography of mitochondrial DNA in western Europe. *Ann Hum Genet.* 62(Pt 3):241-260
- Richards M, Macaulay V, Hickey E, Vega E, Sykes B, Guida V, Rengo C, Sellitto D, Cruciani F, Kivisild T, Villems R, Thomas M, Rychkov S, Rychkov O, Rychkov Y, Gölge M, Dimitrov D, Hill E, Bradley D, Romano V, Cali F, Vona G, Damaine A, Papiha S, Triantaphyllidis C, Stefanescu G, Hatina J, Belledi M, Rienzo AD, Novelletto A, Oppenheim A, Norby S, Al-Zaheri N, Santachiara-Benerencetti S, Scozzari R, Torroni A, Bandelt HJ. 2000. Tracing European Founder Lineages in the Near mtDNA Pool. *Am J Hum Genet* 67: 1251-1276
- Richards M, Macaulay V. 2001. The Mitochondrial Gene Tree Comes of Age. *Am J Hum Genet* 68:1315-1320
- Rivet P. 1943. *Les Origines de l'homme américain.* Les Éditions de l'Arbre. Montréal. Reeditado en 1957 por Les éditions Gallimard
- Roberts R, Jones G, Smith MA. 1990. Thermoluminescence dating of a 50.000 year-old human occupation in northern Australia. *Nature.* 345: 153-156

- Rogaev EI, Grigorenko AP, Moliaka YK, Faskhutdinova G, Goltsov A, Lahti A, Hildebrandt C, Kittler EL, Morozova I. 2009. Genomic identification in the historical case of the Nicholas II royal family. *Proc Natl Acad Sci U S A*. 106(13):5258-5263
- Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Mol Biol Evol*. 9(3):552-569
- Rohling EJ, Fenton M, Jorissen FJ, Bertrand P, Ganssen G, Caulet JP. 1998. Magnitudes of sea-level lowstands of the past 500,000 years. *Nature* 394, 162-165
- Roostalu U, Kutuev I, Loogväli EL, Metspalu E, Tambets K, Reidla M, Khusnutdinova EK, Usanga E, Kivisild T, Villems R. 2007. Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: the Near Eastern and Caucasian perspective. *Mol Biol Evol*. 24(2):436-448
- Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, Armenteros M, Arroyo E, Barbujani G, Beckman G, Beckman L, Bertranpetit J, Bosch E, Bradley DG, Brede G, Cooper G, Côté-Real HB, de Knijff P, Decorte R, Dubrova YE, Evgrafov O, Gilissen A, Glisic S, Gölge M, Hill EW, Jeziorowska A, Kalaydjieva L, Kayser M, Kivisild T, Kravchenko SA, Krumina A, Kucinskas V, Lavinha J, Livshits LA, Malaspina P, Maria S, McElreavey K, Meitinger TA, Mikelsaar AV, Mitchell RJ, Nafa K, Nicholson J, Nørby S, Pandya A, Parik J, Patsalis PC, Pereira L, Peterlin B, Pielberg G, Prata MJ, Previderé C, Roewer L, Rootsi S, Rubinsztein DC, Saillard J, Santos FR, Stefanescu G, Sykes BC, Tolun A, Villems R, Tyler-Smith C, Jobling MA. 2000. Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet*. 67(6):1526-1543
- Ruiz-Linares A, Ortíz-Barrientos D, Figueroa M, Mesa N, Múnera JG, Bedoya G, Vélez ID, García LF, Pérez-Lezaun A, Bertranpetit J, Feldman MW, Goldstein DB. 1999. Microsatellites provide evidence for Y chromosome diversity among the founders of the New World. *Proc Natl Acad Sci U S A*. 96(11):6312-6317
- Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC. 2004. Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* 303:223-226
- Ruz MH. 2006. Mayas. Primera y segunda parte. Colección Pueblos del México contemporáneo. Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CoNaDePI). Programa de las Naciones Unidas para el desarrollo. México
- Sagan L. 1967. On the origin of mitosing cells. *Journal of Theoretical Biology* 14(3): 255-274
- Saillard J, Forster P, Lynnerup N, Bandelt HJ and Nørby S. 2000. mtDNA variation among Greenland Eskimos: the edge of the beringian expansion. *Am J Hum Genet* 67: 718-726
- Salas A, Richards M, De la Fe T, Lareu MV, Sobrino B, Sánchez-Díz P, Macaulay V, Carracedo A. 2002. The making of the African mtDNA Landscape. *Am J Hum Genet*. 71(5): 1082-1111
- Salas A, Richards M, Lareu MV, Scorazzi R, Coppa A, Torroni A, Macaulay V, Carracedo A. 2004. The African diaspora: mitochondrial DNA and the Atlantic slave trade. *Am J Hum Genet* 74(3): 454-465

- Salas A, Carracedo A, Macaulay V, Richards M, Bandelt HJ. 2005. A practical guide to mitochondrial DNA error prevention in clinical, forensic, and population genetics. *Biochem Biophys Res Commun.* 335(3):891-899
- Salas A, Bandelt HJ, Macaulay V, Richards MB. 2007. Phylogeographic investigations: the role of trees in forensic genetics. *Forensic Sci Int.* 168(1):1-13
- Salas A, Fachal L, Marcos-Alonso S, Vega A, Martín-Torres F; Grupo de investigación ESIGEM (Estudio Sobre la Influencia Genética en la Enfermedad Meningocócica). 2009. Investigating the role of mitochondrial haplogroups in genetic predisposition to meningococcal disease. *PLoS One.* 4(12):e8347
- Saltonstal K, Amato G, Powell J. 1998. Mitochondrial DNA variability in Grauer's gorillas of Kahuzi-Biega National Park. *J. Hered.* 89, 129- 135
- Salzano FM. 2002. Molecular variability in Amerindians: widespread but uneven information. *An Acad Bras Cienc.* 74:223-263
- Salzano FM, Callegari Jacques SM. 1988. South American Indians: A Case Study in Evolution. Oxford, England. Clarendon Press
- Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*, second ed., Cold Spring Harbor Laboratory Press, Cold Spring, New York
- Sandoval-Mendoza K, Buentello-Malo L, Peñaloza-Espinosa R, Comas D. 2006. Elementos clave sobre la diversidad genética de poblaciones indígenas mexicanas. *Actas del II Congreso Nacional de Medicina Genómica.* México DF, México
- Sandoval K, Buentello-Malo L, Peñaloza-Espinosa R, Avelino H, Salas A, Calafell F, Comas D. 2009. Linguistic and maternal genetic diversity are not correlated in Native Mexicans. *Hum Genet.* 126(4): 521-531
- Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463-5467
- Santos M, Ward RH, Barrantes R. 1994. mtDNA variation in the Chibcha Amerindian Huetar from Costa Rica. *Hum Biol* 66:963-977
- Santos C, Montiel R, Sierra B, Bettencourt C, Fernandez E, Alvarez L, Lima M, Abade A, Aluja MP. 2005. Understanding differences between phylogenetic and pedigree-derived mtDNA mutation rate: a model using families from the Azores Islands (Portugal). *Mol Biol Evol.* 22(6):1490-1505
- Sarich VM, Wilson AC. 1967. Immunological time scale for hominid evolution. *Science* 158: 1200-1203
- Schaefer SB, Furst PT. 1996. *People of the Peyote.* Albuquerque, Nuevo México
- Schneider S, Excoffier L. 1999. Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics.* 152(3):1079-1089
- Schroeder KB, Schurr TG, Long JC, Rosenberg NA, Crawford MH, Tarskaia LA, Osipova LP, Zhadanov SI, Smith DG. 2007. A private allele ubiquitous in the Americas. *Biol Lett.* 3(2):218-223

- Schurr, T.G., Ballinger, S.W., Gan Y-Y., Hodge J.A., Merriwether D.A., Lawrence, D. N., Knowler, W.C., Weiss, K.M. and Wallace, D.C. 1990. Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. *Am J Hum Genet.* 46(3): 613-623
- Schurr TG, Wallace DC. Mitochondrial DNA diversity in Southeast Asian populations. 2002. *Hum Biol.* 74(3):431-452
- Schurr TG, Sherry ST. 2004. Mitochondrial DNA and Y chromosome diversity and the peopling of the Americas: evolutionary and demographic evidence. *Am J Hum Biol* 16: 420-439
- Schurr TG. 2004. The peopling of the New World: Perspectives from molecular anthropology. *Annu Rev Anthropol* 33: 551-583
- Schwartz M, Vissing J. 2002. Paternal inheritance of mitochondrial DNA. *N. England Journal Medicine.* 347:576–580
- Scozzari R, Torroni A, Semino A, Sirugo G, Brega A, Santa-chiara-Benerecetti AS. 1988. Genetic studies on the Senegal population. Mitochondrial DNA polymorphisms. *Am J Hum Genet.* 43(4): 534-544
- Seebach JD. 2002. Stratigraphy and Bonebed Taphonomy at Blackwater Draw Locality No. 1 during the Middle Holocene (Altithermal). *Plains Anthropologist.* 47(183):339-358
- Serrano Carreto E, Gámez Montes V, Maldonado Salazar I, Bello Jiménez E, Velázquez Rosendo B, de Lourdes Ayala M, Virginia García Vidales L, Ramírez Martínez B, Salas Serrato JA. 2006. Regiones indígenas de México. Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CoNaDePI). Programa de las Naciones Unidas para el desarrollo. México
- Serrano C. 1995. Genes, evolución y diversidad humana. Temas de antropología molecular. Universidad Autónoma de México. Instituto de Investigaciones Antropológicas. México D.F., México
- Shi H, Zhong H, Peng Y, Dong YL, Qi XB, Zhang F, Liu LF, Tan SJ, Ma RZ, Xiao CJ, Wells RS, Jin L, Su B. 2008. Y chromosome evidence of earliest modern human settlement in East Asia and multiple origins of Tibetan and Japanese populations. *BMC Biol.* 6:45
- Siddall M, Rohling EJ, Almogi-Labin A, Hemleben Ch, Meischner D, Schmelzer I, Smeed DA. 2003. Sea-level fluctuations during the last glacial cycle. *Nature.* 423:853-858
- Silva-Júnior WA, Bonatto SL, Holanda AJ, Ribeiro-dos-Santos AK, Paixão BM, Goldman GH, Abe-Sandes K, Rodriguez-Delfin L, Barbosa M, Paçó-Larson ML, Petzl-Erler ML, Valente V, Santos SEB, Zago MA. 2002. Mitochondrial genome diversity of native Americans supports a single early entry of founder populations into America. *Am J Hum Genet* 71: 187-192
- Simoni L, Calafell F, Pettener D, Bertranpetit J, Barbujani G. 2000a. Reconstruction of prehistory on the basis of genetic data. *Am J Hum Genet.* 66: 1177-1179
- Simoni L, Calafell F, Pettener D, Bertranpetit J, Barbujani G. 2000b. Geographic Patterns of mtDNA Diversity in Europe. *Am J Hum Genet* 66: 262-278

- Smith DG, Malhi RS, Eshleman J, Lorenz JG, Kaestle FA. 1999. Distribution of mtDNA haplogroup X among Native North Americans. *Am J Phys Anthropol.* 110(3):271-284
- Smith CI, Chamberlain AT, Riley MS, Stringer C, Collins MJ. 2003. The thermal history of human fossils and the likelihood of successful DNA amplification. *J Hum Evol.* 45(3):203-317
- Soares P, Achilli A, Semino O, Davies W, Macaulay V, Bandelt HJ, Torroni A, Richards MB. 2010. The archaeogenetics of Europe. *Curr Biol.* 20(4):174-183
- Solanes Carraro MC, Vela Ramírez E. 2005. Atlas del México Prehispánico. Mapas, regiones y culturas. Arqueología Mexicana. Consejo Nacional para la Cultura y las Artes, Instituto Nacional de Antropología e Historia. México
- Starikovskaya YB, Sukernik RI, Schurr TG, Kogelnik AM, Wallace DC. 1998. MtDNA diversity in Chukchi and Siberian Eskimos: implications for the genetic history of ancient Beringia and peopling of the New World. *Am J Hum Genet* 63: 1473-1491
- Steele G, Powell JF. 1999. Peopling of the Americas: a historical and comparative perspective. En: Bonnichsen R (ed.) *Who were the first americans?* Center for the study of the first americans, Oregon State University, Corvallis. Pags. 97-126
- Steffens DL, Roy R. 1998. Sequence analysis of mitochondrial DNA hypervariable regions using infrared fluorescence detection. *Biotechniques* 24, 1044- 1046
- Stone AC, Stoneking M. 1998. mtDNA analysis of a prehistoric Oneota population: Implications for the peopling of the New World. *Am J Hum Genet* 62:1153-1170
- Stoneking M. 1997. The human genome project and molecular anthropology. *Genome Res.* 7(2):87-91
- Stresser-Péan G. 2008. Viaje a la Huasteca con Guy Stresser-Péan. Fondo de Cultura Económica. México D.F., México
- Stringer C, McKie R. 1996. *The African exodus.* Jonathan Cape. Londres.
- Sturrock K, Rocha J. 2000. A Multidimensional Scaling Stress Evaluation Table. *Fields Methods.* 12(1):49-60
- Suárez JA. 1995. Las lenguas indígenas mesoamericanas. Instituto Nacional Indigenista (INI). Centro de Investigaciones y Estudios Superiores e Antropología Social (CIESAS). México DF, México
- Swadesh M. 1959. Indian linguistic groups of Mexico. Escuela Nacional de Antropología e Historia, Instituto Nacional de Antropología e Historia. México DF, México
- Swadesh M. 1960. The Oto-manguanean hypothesis and Macro-mixtecan. *International Journal of American Linguistics* 26: 79-111
- Swadesh M. 1963. Nuevo ensayo de glotocronología yutonahua. *Anales del Instituto Nacional de Antropología e Historia* XV, 44: 263-302
- Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105: 437-460

- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis. *Genetics* 123: 585-595
- Tajima F. 1993. Measurement of DNA polymorphism. In: *Mechanisms of Molecular Evolution. Introduction to Molecular Paleopopulation Biology*, edited by Takahata, N. and Clark, A.G., Tokyo, Sunderland, MA:Japan Scientific Societies Press, Sinauer Associates, Inc., p. 37-59.
- Tajima F. 1996. The amount of DNA polymorphism maintained in a finite population when the neutral mutation rate varies among sites. *Genetics* 143: 1457-1465
- Tamm E, Kivisild T, Reidla M, Metspalu M, Smith DG, Mulligan CJ, Bravi CM, Rickards O, Martinez-Labarga C, Khusnutdinova EK, Fedorova SA, Golubenko MV, Stepanov VA, Gubina MA, Zhadanov SI, Ossipova LP, Damba L, Voevoda MI, Dipierri JE, Villems R, Malhi RS. 2007. Beringian standstill and spread of Native American founders. *PLoS One*. 2(9):e829
- Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology Evolution*. 10: 512-526
- Tanaka M, Ozawa T. 1994. Strand asymmetry in human mitochondrial DNA mutations. *Genomics* 22(2):327-335
- Templeton A. 2002. Out of Africa again and again. *Nature*. 416: 45-51
- Thangaraj K, Chaubey G, Kivisild T, Reddy AG, Singh VK, Rasalkar AA, Singh L. 2005. Reconstructing the origin of Andaman Islanders. *Science*. 13;308(5724):996
- Thomas C. 1911. *Indian Languages of Mexico and Central America*. Washington, DC
- Thorne A, Curnoe D. 2000. Sex and significance of Lake Mungo 3: reply to Brown "Australian pleistocene variation and the sex of Lake Mungo 3". *J Hum Evol*. 39(6):587-600
- Thorne A, Grün R, Mortimer G, Spooner NA, Simpson JJ, McCulloch M, Taylor L, Curnoe D. 1999. Australia's oldest human remains: age of the Lake Mungo 3 skeleton. *J Hum Evol*. 36(6):591-612
- Tian C, Plenge RM, Ransom M, Lee A, Villoslada P, Selmi C, Klareskog L, Pulver AE, Qi L, Gregersen PK, Seldin MF. 2008. Analysis and application of European genetic substructure using 300 KSNP information. *PLoS Genet*. 4(1):e4
- Torroni A., Schurr, T.G., Yang, C-C., Szathmary, E.J.E., Williams, R.C., Schanfield, M.S., Troup, G.A., Knowler, W.C., Lawrence, D.N., Weiss, K.M., Wallace, D.C. 1992. Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. *Genetics* 130: 153-162
- Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC 1993a. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am J Hum Genet* 53: 563-590
- Torroni A, Sukernik RI, Schurr TG, Starikorskaya YB, Cabell MF, Crawford MH, Comuzzie AG, Wallace DC. 1993b. mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. *Am J Hum Genet*. 53(3): 591-608

- Torrioni A, Miller JA, Moore LG, Zamudio S, Zhuang J, Droma T, Wallace DC. 1994a. Mitochondrial DNA analysis in Tibet: implications for the origin of the Tibetan population and its adaptation to high altitude. *Am J Phys Anthropol* 93:189-199
- Torrioni A, Lott MT, Cabell MF, Chen YS, Lavergne L, Wallace DC. 1994b. Mitochondrial DNA and the origin of the Caucasians: identification of Ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. *Am J Hum Genet* 55:760-776
- Torrioni A, Chen YS, Semino O, Santachiara-Beneceretti AS, Scott CR, Lott MT, Winter M, Wallace DC. 1994c. mtDNA and Y-Chromosome Polymorphisms in Four Native American Populations from Southern Mexico. *Am J Hum Genet* 54: 303-318
- Torrioni A, Neel JV, Barrantes R, Schurr TG, Wallace DC. 1994d. Mitochondrial DNA "clock" for the Amerinds and its implications for timing their entry into North America. *Proc Natl Acad Sci U S A*. 91(3):1158-1162
- Torrioni A, Bandelt HJ, Macaulay V, Richards M, Cruciani F, Rengo C, Martinez-Cabrera V, Villems R, Kivisild T, Metspalu E, Parik J, Tolk HV, Tambets K, Forster P, Karger B, Francalacci P, Rudan P, Janicijevic B, Rickards O, Savontaus ML, Huoponen K, Laitinen V, Koivumäki S, Sykes B, Hickey E, Novelletto A, Moral P, Sellitto D, Coppa A, Al-Zaheri N, Santachiara-Benerecetti AS, Semino O, Scozzari R. 2001. A signal, from human mtDNA, of postglacial recolonization in Europe. *Am J Hum Genet*. 69(4):844-852
- Torrioni A, Achilli A, Macaulay V, Richards M, Bandelt HJ. 2006. Harvesting the fruit of the human mtDNA tree. *Trends in Genetics* 22(6): 339-345
- Tully LA, Parsons TJ, Steighner RJ, Holland MM, Marino MA, Prenger VL. 2000. A sensitive denaturing gradient-gel electrophoresis assay reveals a high frequency of heteroplasmy in hypervariable region 1 of the human mtDNA control region. *Am J of Hum Genet*. 67:432-443
- Underhill PA, Passarino G, Lin AA, Shen P, Mirazón Lahr M, Foley RA, Oefner PJ, Cavalli-Sforza LL. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet*. 65(Pt 1):43-62
- Underhill PA, Kivisild T. 2007. Use of y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu Rev Genet*. 2007;41:539-564
- Valle Esquivel J. 2006. Nahuas de la Huasteca. Colección Pueblos del México contemporáneo. Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CoNaDePI). Programa de las Naciones Unidas para el desarrollo. México
- Valle Esquivel J, Hernández Alvarado JB. 2006. Huastecos de Veracruz. Colección Pueblos del México contemporáneo. Comisión Nacional para el Desarrollo de los Pueblos Indígenas (CoNaDePI). Programa de las Naciones Unidas para el desarrollo. México
- Van Oven M, Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 30(2):E386-E394
- Via García M. 2006. Variació molecular en les sintases d'oxid nítric en poblacions humanes i susceptibilitat cardiovascular. Tesi Doctoral. Universitat de Barcelona

- Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC. 1989. Mitochondria DNA sequences in single hairs from a southern African population. *Proc Natl Acad Sci USA* 86: 9350- 9354
- Wallace DC, Garrison K, Knowler, W.C. 1985. Dramatic founder effects in Amerindian mitochondrial DNAs. *Am J Phys Anthropol* 68(2): 149-155
- Wallace DC, Ye JH, Neckelmann SN, Singh G, Webster KA, Greenberg BD. 1987. Sequence analysis of human and bovine ATP synthase beta subunit: mitochondrial DNA genes sustain seventeen times more mutations. *Curr Genet.* 12: 81-90
- Wallace DC, Lott MT, Torroni A, Shoffner JM. 1991. Report of the committee on human mitochondrial DNA. *Cytogene Cell Genet* 58: 1103-1123
- Wallace DC, Stugard C, Murdock D, Schurr T, Brown MD. 1997. Ancient mtDNA sequences in the human nuclear genome: a potential source of errors in identifying pathogenic mutations. *Proc Natl Acad Sci U S A.* 23;94(26):14900-14905
- Walsh PS, Metzger DA, Higuchi R. 1991. Chelex® 100 as a medium for a simple extraction of DNA for PCR- based typing for forensic material. *Biotechniques.* 10: 506-513
- Walter RC, Buffler RT, Bruggemann JH, Guillaume MM, Berhe SM, Negassi B, Libsekal Y, Cheng H, Edwards RL, von Cosel R, Néraudeau D, Gagnon M. 2000. Early human occupation of the Red Sea coast of Eritrea during the last interglacial. *Nature.* 405:65-69
- Wang S, Lewis CM, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W, Parra MV, Molina JA, Gallo C, Mazzotti G, Poletti G, Hill K, Hurtado AM, Labuda D, Klitz W, Barrantes R, Bortolini MC, Salzano FM, Petzl-Erler ML, Tsuneto LT, Llop E, Rothhammer F, Excoffier L, Feldman MW, Rosenberg NA, Ruiz-Linares A. 2007. Genetic variation and population structure in Native Americans. *PLoS Genet* 3:e185
- Wang S, Ray N, Rojas W, Parra MV, Bedoya G, Gallo C, Poletti G, Mazzotti G, Hill K, Hurtado AM, Camrena B, Nicolini H, Klitz W, Barrantes R, Molina JA, Freimer NB, Bortolini MC, Salzano FM, Petzl-Erler ML, Tsuneto LT, Dipierri JE, Alfaro EL, Bailliet G, Bianchi NO, Llop E, Rothhammer F, Excoffier L, Ruiz-Linares A. 2008. Geographic patterns of genome admixture in Latin American Mestizos. *PLoS Genet* 4:e1000037
- Ward RH, Frazier BL, Dew-Jager K, Pääbo S. 1991. Extensive mitochondrial diversity within a single Amerindian tribe. *Proc Natl Acad Sci USA* 88: 8720-8724
- Waters MR, Stafford TW Jr. 2007. Redefining the age of Clovis: implications for the peopling of the Americas. *Science.* 315(5815):1122-1126
- Watson E, Forster P, Richards M, Bandelt HJ. 1997. Mitochondrial footprints of human expansions in Africa. *Am J Hum Genet.* 61, 691–704
- Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358-1370
- Weir BS. 1996. *Genetic Data Analysis II: Methods for Discrete Population Genetic Data.* Sinauer Assoc., Inc., Sunderland, MA, USA

- Weitlaner R. 1961. Cantos otomíes de la sierra de Puebla. Homenaje a Pablo Martínez del Río, México, Universidad Nacional Autónoma de México UNAM), México D.F., México
- Williams García R. 1953. Etnografía prehispánica de la zona central de Veracruz. *Revista Mexicana de Estudios Antropológicos* 13(2-3):157-162
- Williams García R. 1963. Los Tepehuas. Instituto de Antropología de la Universidad Veracruzana. Xalapa, México
- Wright S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution*. 19:395-420
- Xu X, Arnason U. 1996. A complete sequence of the mitochondrial genome of the western lowland gorilla. *Mol. Biol. Evol.* 13 (5): 691-698
- Xue Y, Zerjal T, Bao W, Zhu S, Shu Q, Xu J, Du R, Fu S, Li P, Hurles ME, Yang H, Tyler-Smith C. 2006. Male demography in East Asia: a north-south contrast in human population expansion times. *Genetics*. 172(4):2431-2439
- Yao YG, Kong QP, Bandelt HJ, Kivisild T, Zhang YP. 2002. Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am J Hum Genet.* 70(3):635-651
- Zakharov IA. 2000. Mitochondrial DNA variation in two South Siberian Aboriginal populations: implications for the genetic history of North Asia. *Hum Biol.* 72(6):945-973
- Zalloua PA, Platt DE, El Sibai M, Khalife J, Makhoul N, Haber M, Xue Y, Izaabel H, Bosch E, Adams SM, Arroyo E, López-Parra AM, Aler M, Picornell A, Ramon M, Jobling MA, Comas D, Bertranpetit J, Wells RS, Tyler-Smith C; Genographic Consortium. 2008. Identifying genetic traces of historical expansions: Phoenician footprints in the Mediterranean. *Am J Hum Genet.* 83(5):633-642
- Zaragoza MV, Fass J, Diegoli M, Lin D, Arbustini E. 2010. Mitochondrial DNA Variant Discovery and Evaluation in Human Cardiomyopathies through Next-Generation Sequencing. *PLoS One.* 5(8). e12295
- Zavala C, Alatorre S, Lisker R. 1982. Distancias génicas entre algunos grupos indígenas mexicanos. *Est Antropol Biol*, 1: 141-154
- Zegura SL, Karafet TM, Zhivotovsky LA, Hammer MF. 2004. High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Mol Biol Evol.* 21(1):164-175
- Zerjal T, Wells RS, Yuldasheva N, Ruzibakiev R, Tyler-Smith C. 2002. A genetic landscape reshaped by recent events: Y-chromosomal insights into central Asia. *Am J Hum Genet.* 71(3):466-482
- Zhong H, Shi H, Qi XB, Duan ZY, Tan PP, Jin L, Su B, Ma RZ. 2010. Extended Y-chromosome investigation suggests post-Glacial migrations of modern humans into East Asia via the northern route. *Mol Biol Evol.* [Epub ahead of print]
- Zlojutro M, Rubicz R, Devor EJ, Spitsyn VA, Makarov AV, Wilson K and Crawford MH. 2006. Genetic structure of the Aleuts and circumpolar populations based on mitochondrial DNA sequences: a synthesis. *Am J Phys Anthropol* 129: 446-464

- Zouros E, Freeman KR, Ball AO, Pogson GH. 1992. Direct evidence for extensive paternal mitochondrial DNA inheritance in the marine mussel *Mytilus*. *Nature* 359:412–414
- Zuckerandl E, Pauling L. 1962. Molecular disease, evolution, and genetic heterogeneity. In: *Horizons in Biochemistry*. edited by M. Kasha and B. Pullman. Academic Press, New York

VII. ANEXO

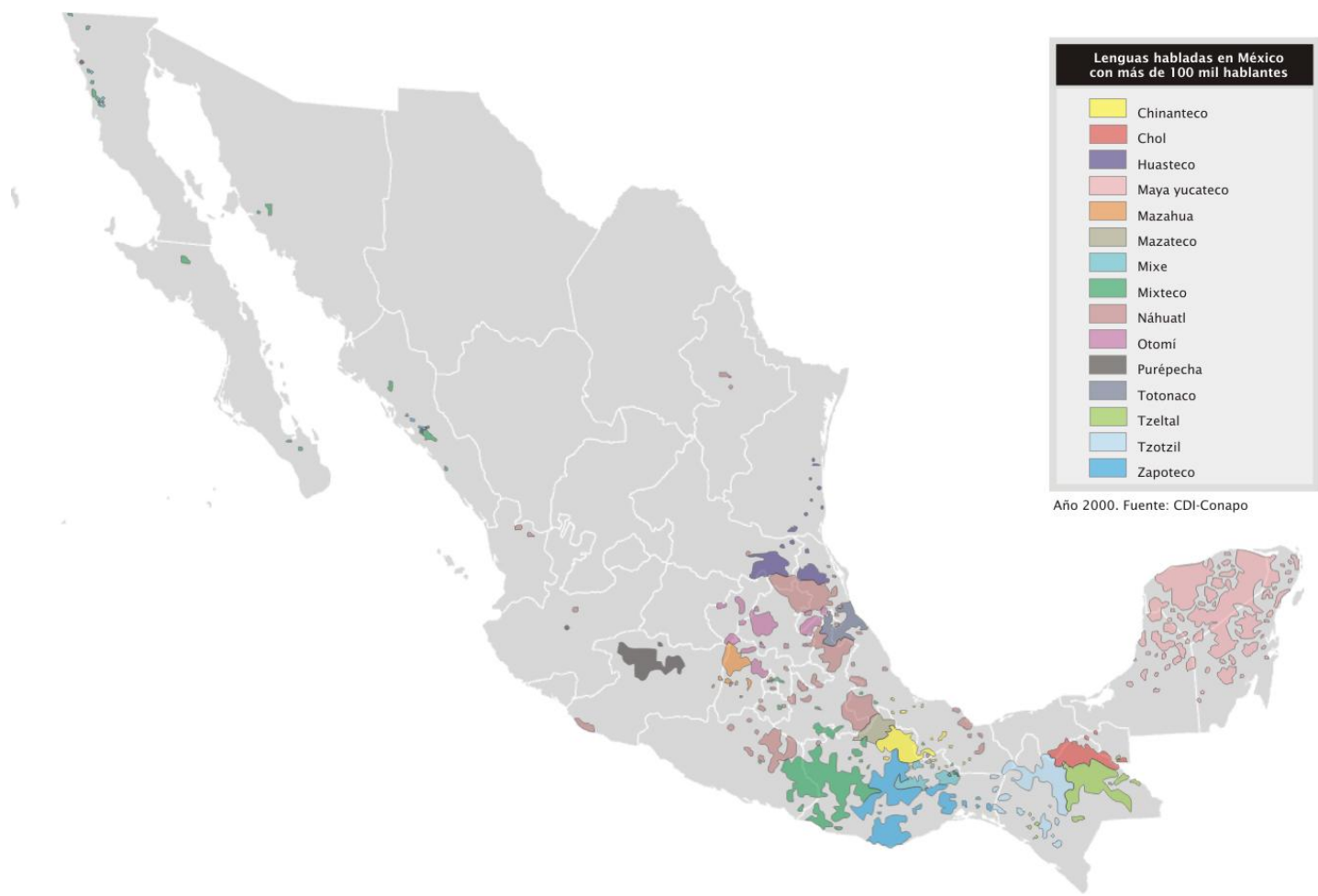


Figura VII.1. Distribución de lenguas indígenas habladas en México con más de 100.000 hablantes. Dentro de esta distribución quedan integradas las poblaciones Nahuas de la Huasteca, Otomíes y Mayas (INALI 2007, 2008)

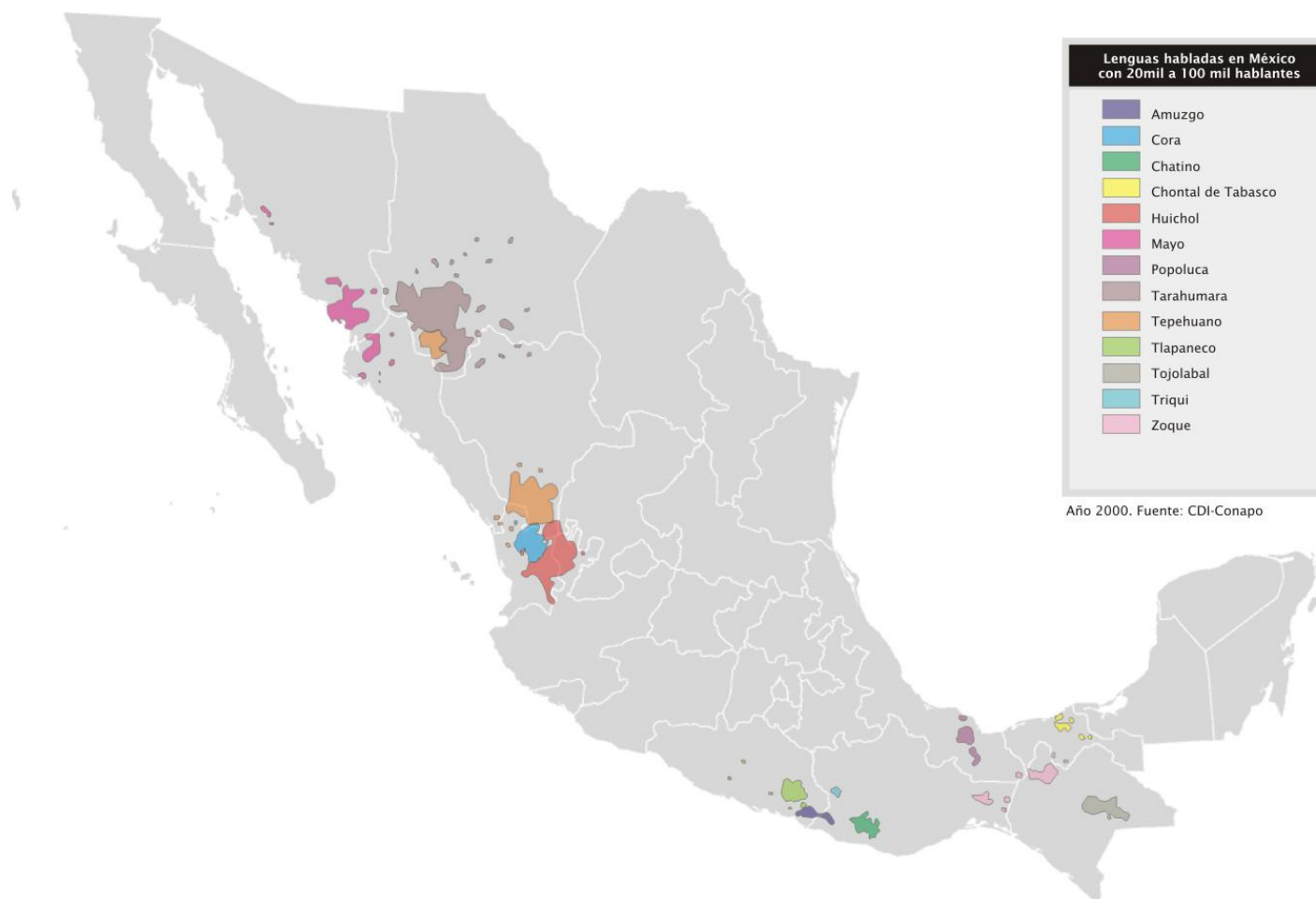


Figura VII.2. Distribución de lenguas indígenas habladas en México entre 20.000 y 100.000 hablantes. Dentro de esta distribución quedan integradas las poblaciones Huichol y Mayo (INALI 2007, 2008)



Figura VII.3. Distribución de lenguas indígenas habladas en México con menos de 20.000 hablantes. Dentro de esta distribución quedan integradas las poblaciones Pima y Tepehua (INALI 2007, 2008)



Cédula de Trabajo de Campo
Universidad Complutense de Madrid

ID individuo:

Fecha: Lugar: Lengua (ego y/o familia):
 Nombre: Edad: Sexo:
 ¿Participa algún familiar en este estudio?
 Lugar de nacimiento: Ego:
 Padre: Madre:
 Abuelo paterno: Abuelo materno
 Abuela paterna: Abuela materna:
 Enfermedades (ego y/o familia)
 Observaciones:

Autorizo a la Universidad Complutense de Madrid para analizar el material genético de mi muestra con el objetivo de realizar estudios antropológicos poblacionales. Asimismo, se me ha explicado que la muestra tomada será empleada únicamente para fines de investigación, y que toda la información, tanto la generada por la muestra como la que he proporcionado, son confidenciales. Soy consciente de que tengo acceso a los resultados que de mi muestra deriven y que sean de relevancia en mi historial médico

Nombre y firma

Figura VII.4. Encuesta modelo utilizada

Tabla VII.1. Consideraciones con respecto a la asignación de los haplogrupos y las mutaciones diagnóstico para cada uno de ellos

Individuo	Observaciones
PI78	A pesar de tener el polimorfismo 16189C, que diagnostica el HgB, la página del Genographic lo clasifica como HgD. Se trata de HgD1 ya que tiene los tres polimorfismos diagnósticos: 16223T, 16325C, 16362C
MO28, MO44 MO49, MO59 MO64, MO65 MO70, MO85 MO86, MO88	Presentan 16189C, podría ser una mutación fantasma (Brandstatter et al., 2005) aunque debemos considerar que no es así y realmente se trata de una mutación real. Todos los individuos con 16189C tienen además la 16217C, el otro punto diagnóstico que los define como HgB2, ahora bien, los individuos con 16189d no presentan la 16217C y además presentan la 16223T, 16325C y 16362C, los tres puntos diagnósticos para el HgD1. Consideramos todos estos individuos como HgD1
MO05, MO16 MO76, MO89	Presentan los cuatro puntos diagnósticos en HVI de HgC1 pero también el 16362C que diagnostica a HgD1. De todas formas todos ellos tienen la249d, 290-291d en la HVII por lo que los clasificamos como HgC1
MO87	Presenta en HVI los marcadores propios de HgD1, además del 1632T
MO15	No presenta las variantes 249d, 290-291d en la HVII por lo que no es HgC1. Al cargar su haplotipo en Mitosearch coincide en gran medida con HgD aunque no se ha descrito ningún haplotipo idéntico, todos los descritos en esta página tienen una o dos diferencias
MO77	El haplotipo en la página del Genographic da HgD
MO08	Es un haplotipo incompleto aunque presenta la mutación 16189C, lo que permite asignarlo como HgB, además, las variantes de la HVII son típicas de este haplogrupo
MO91	Presenta en HVI las variantes propias del HgB2 además la 249d, 290-291d en la HVII, estas últimas diagnósticas del HgC1. Lo clasificamos como HgB2
OS474	Junto con OS482, 490, 575, 589, 592, 593, 596, 679, 690, 691 y 692 tienen el mismo haplotipo. Al cargarlo en la página de Genographic se clasifica como HgA, lógico ya que tiene todas las cinco variantes de A2 más la 16176T
OS498, OS498 OS606, OS675 OS704	Presentan los siguientes marcadores: 16153A, 16223T, 16240G, 16290T, 16319A, 16362C, 16390A y 16362C. Cuatro de ellos son puntos diagnósticos para HgA2, solo les falta el 16111T. Genographic los clasificamos como HgA2
OS491, OS611 OS694, OS695 OS696	Todos ellos tienen los marcadores que definen al HgA2 excepto el 16111T. En un principio podría interpretarse como que podrían ser HgA, ya que carecen de uno de los marcadores que diferencia a estos dos Hg. Ahora bien, la presencia del 16362C permite clasificarlos como HgA2
OS604	Cargando solo la HVI en Mitosearch no da mach aunque los haplotipos más parecidos son HgH (-1). Carecen de la 16126C o de la 16278T. Genographic lo identifica como HgH
OS688	En Mitosearch no aparece ningún haplotipo parecido y Genographic lo clasifica como HgV. Desgraciadamente no ha salido la secuencia de la HVII lo que impide intentar otros sistemas para clasificarlo
OS503	Se le asigna al HgC1 ya que presenta la 249d y las 290-291d

OS476, OS493 OS499, OS500 OS591, OS597	Todos estos individuos presentan una combinación de marcadores de HgD1 y HgC1: del primero presentan 16223T, 16325C y 16362C y del HgC1 tienen 16223T, 16325C y 16327T, careciendo del 16298C. Ahora bien, presentan en el HVII las deleciones 249d, 290-291d propias de HgC1. Para confirmar el diagnóstico cargamos el haplotipo en Genographic y se clasifican como HgC1
OS481, OS483 OS579, OS673	Presentan algunos marcadores comunes infrecuentes (16209C y 16519C) pero tienen cuatro marcadores propios de HgA2: 16111T, 16223T, 16319A y 16362C
OS686	Individuo muy complicado de catalogar ya que en la HVII no presenta los marcadores 249d ni 290-291d propios del HgC1. Ahora bien, en la HVI tiene las variantes 16223, 16298C, 16325C y 16327T propias de este haplogrupo aunque la 16223T y 16325C son también diagnósticas para el HgD1. Tanto en Genographic, como en Mitosearch, utilizando en este último solo la HVI se clasifica como HgC1. Otra singularidad es que en Mitosearch aparecen bastantes individuos HgC1 iguales en la HVI que OS686 pero ninguno de ellos tiene una HVII parecida. Utilizando la premisa de que la mayor parte de los individuos se clasifican atendiendo a la HVI lo clasificamos a pesar de todo como HgC1
OS598	Presenta un haplogrupo bastante atípico, se le clasifica como HgC1 ya que tiene las variantes 16223T, 16298C, 16325C y 16327T
TE453, TE454 TE465, TE564 TE572, TE651 TE652, TE654 TE656, TE667	Todos estos individuos de la población Tepehua presentan los marcadores de HgA, carecen del 16111T que define a HgA2. En este sentido deberían de clasificarse como HgA ahora bien, en región HVII todos presentan, además de otras, las mutaciones 64T, 146C y 153G que caracterizan a HgA2. Los clasificamos, por lo tanto, como HgA2
TE565 TE655	Estos dos haplotipos no presentan la mutación 16111T aunque si el resto de marcadores propios de HgA. TE563 tiene en HVII 64T pero no la 153G, al 655 le pasa al contrario, tiene la 153G pero no la 64T. De todas formas, es muy difícil que se trate del HgA, los clasificamos como HgA2.
MA125, MA131	Ambos tienen el mismo haplotipo. Respecto a su clasificación como HgA2 carecen de la 16111T en la HVI y de la 153G en la HVII. De todas formas las mutaciones en la HVI hacen que la página Genographic los clasifique como HgA2. Mitosearch también los clasifica como A aunque no hay mach completo, el resultado más aproximado es (-1)
MA147, MA153	Todas las pruebas lo clasifican como HgM. Genographic confirma este resultado y Mitosearch da dos match de coincidencia utilizando la HVI, un resultado es HgM y otro HgD aunque en este último se puede consultar el HVII y no coincide con el de las muestras. Al consultar los polimorfismos en la base de datos de xls de Genographic, los puntos que definen al HgM no coinciden con los de la muestra. De todas formas estos individuos tienen dos de tres puntos diagnósticos de HVI (16223T y 16325C) y les falta tan solo el 16327T. Respecto a HVII presentan las dos mutaciones que definen a HgD, la 73G y la 263G. Consideramos que probabilísticamente deben de tratarse de individuos con HgD1.
MA159	Haplotipo bastante raro aunque claramente se clasifica como HgA2
NA176, NA188 NA257	Presentan haplotipos muy parecidos, con variantes semejantes a HgA2 pero sin el 16111T que podría interpretarse como que son HgA. De todas formas en la HVII presentan los polimorfismos que definen a HgA2
H253	No presenta la 16111T pero puede clasificarse con facilidad como HgA2
H321	Haplotipo raro que clasificamos como HgD1. Descartamos que sea HgA2 pues carece de la

	16111T y de la 16290T. Por otra parte presenta las variantes 16223Ty 16325T propias deHgD1
H408	Haplotipo bastante raro aunque claramente se clasifica como HgC1
H326	Genographic lo clasifica como HgC1 aunque no presenta 249d, 290-291d en la HVII

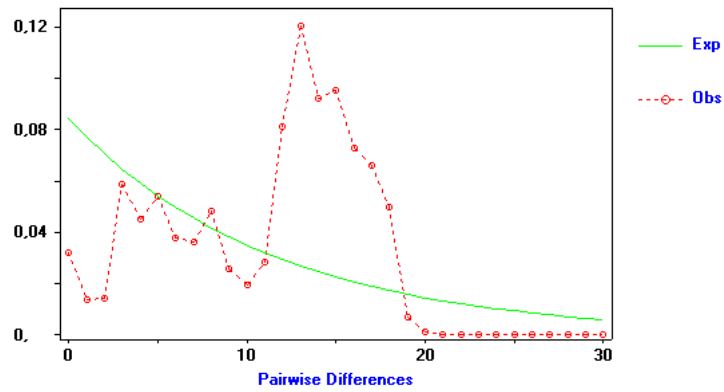


Figura VII.5. *Pairwise differences* para los Otomíes del Valle considerando la región D-Loop completa.

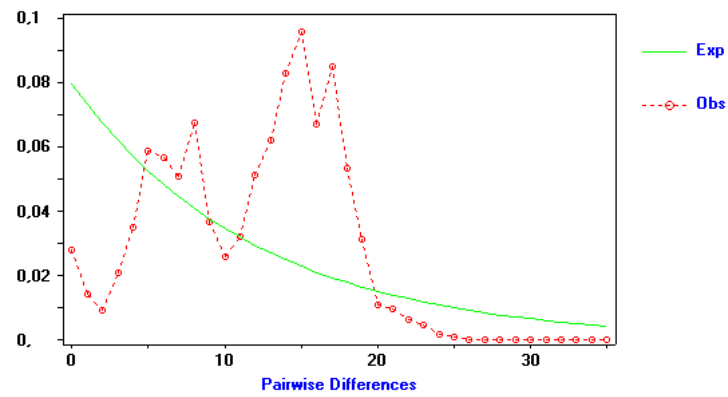


Figura VII.6. *Pairwise differences* para los Otomíes de la Sierra considerando la región D-Loop completa

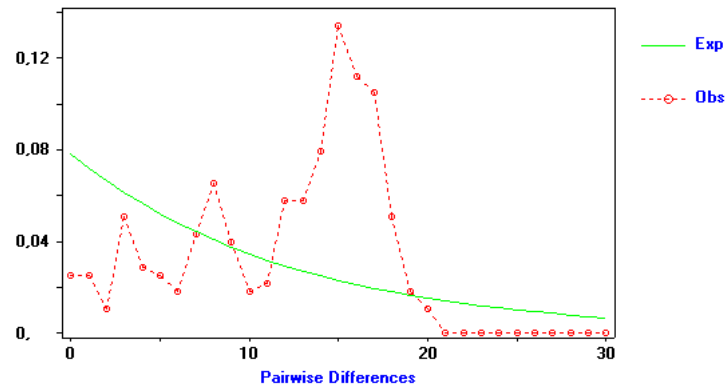


Figura VII.7. *Pairwise differences* en la comunidad de El Buena al estudiar la región D-Loop completa.

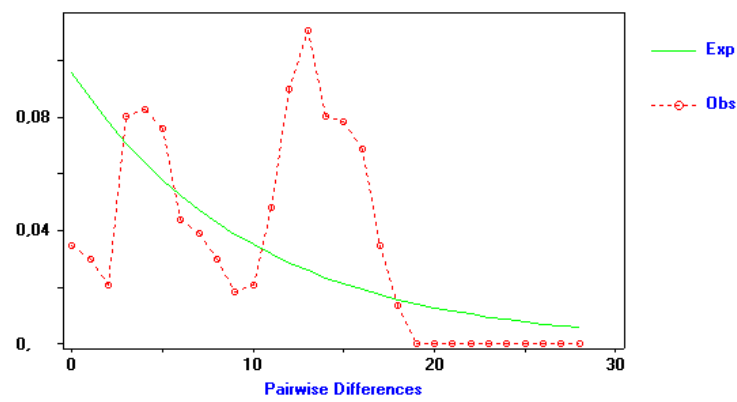


Figura VII.8. *Pairwise differences* en la comunidad de El Deca al estudiar la región D-Loop completa.

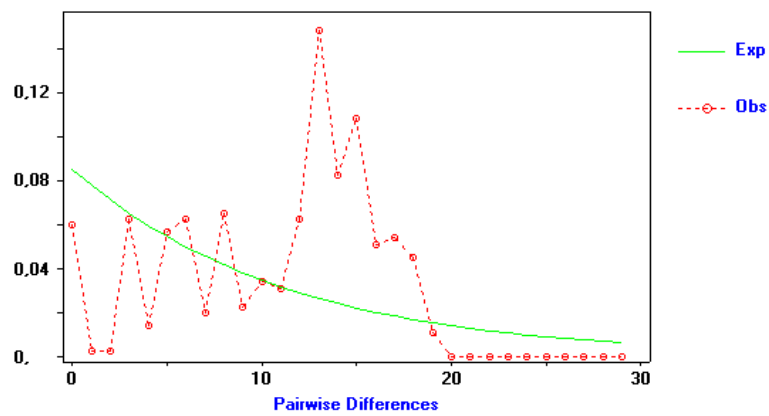


Figura VII.9. *Pairwise differences* en la comunidad de San Andrés Dabothá al estudiar la región D-Loop completa.

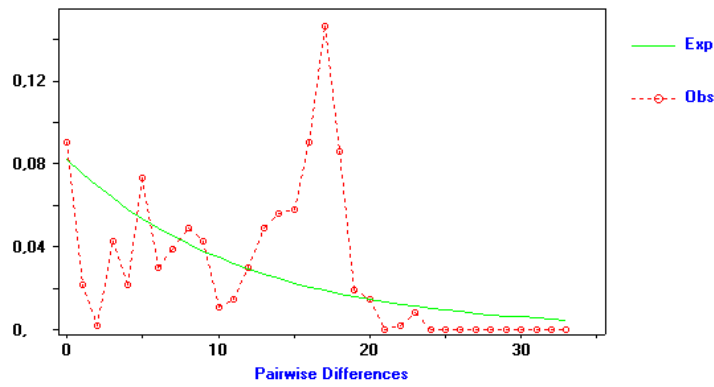


Figura VII.10. *Pairwise differences* en la comunidad de San Nicolás al estudiar la región D-Loop completa.

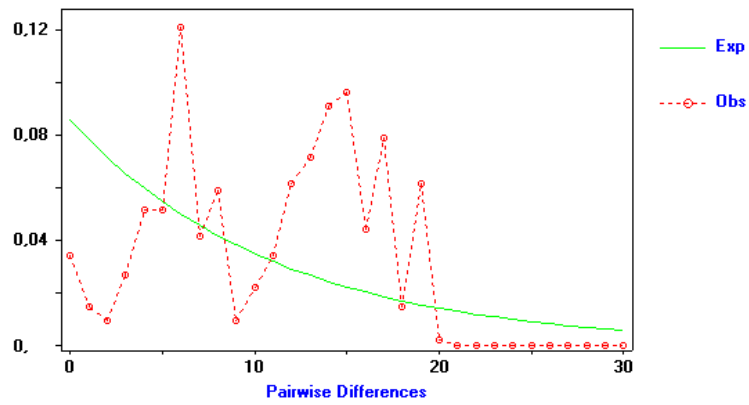


Figura VII.11. *Pairwise differences* en la comunidad de Santa Mónica al estudiar la región D-Loop completa.

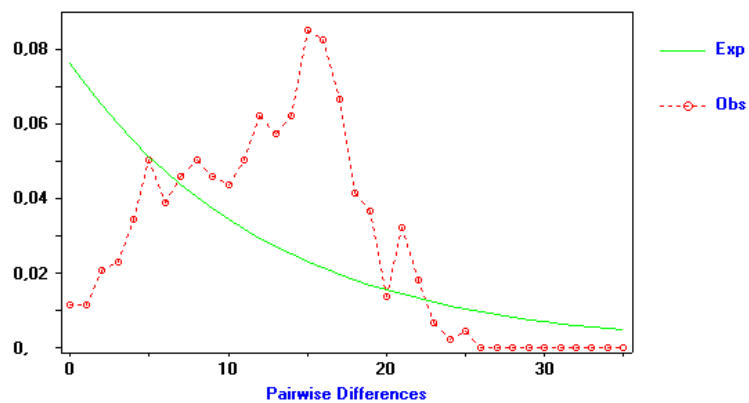


Figura VII.12. *Pairwise differences* en la comunidad de Santiago al estudiar la región D-Loop completa.

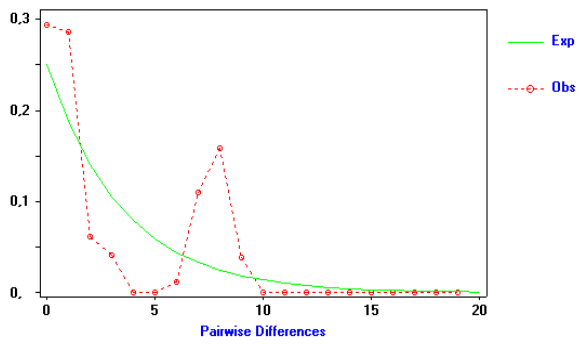


Figura VII.13. *Pairwise differences* de la población Pima para la región HVI.

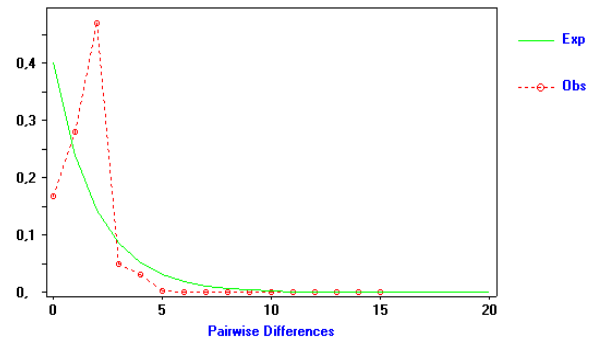


Figura VII.14. *Pairwise differences* de la población Pima para la región HVII.

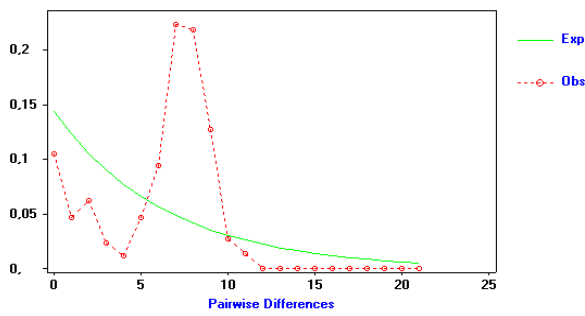


Figura VII.15. *Pairwise differences* de la población Mayo para la región HVI.

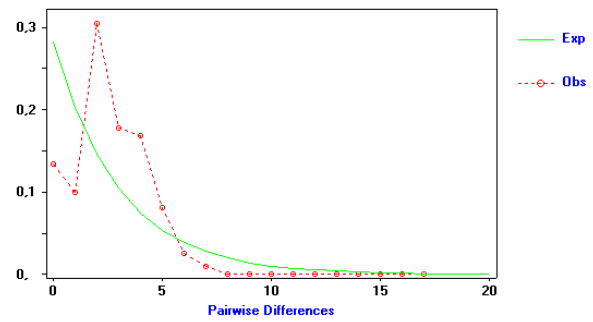


Figura VII.16. *Pairwise differences* de la población Mayo para la región HVII.

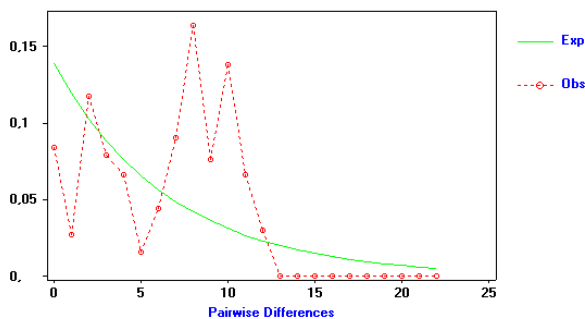


Figura VII.17. *Pairwise differences* de la población Huichol para la región HVI.

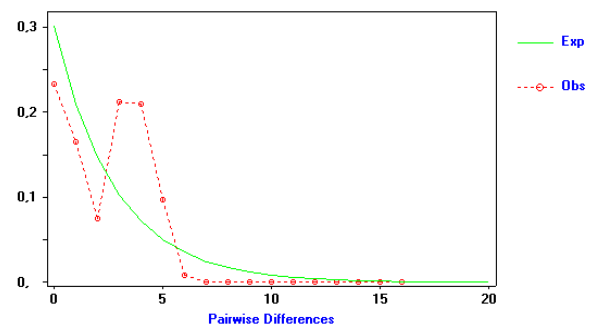


Figura VII.18. *Pairwise differences* de la población Huichol para la región HVII.

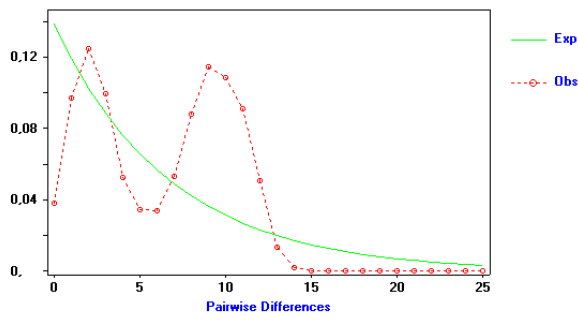


Figura VII.19. *Pairwise differences* de Nahuas de la Huasteca para la región HVI.

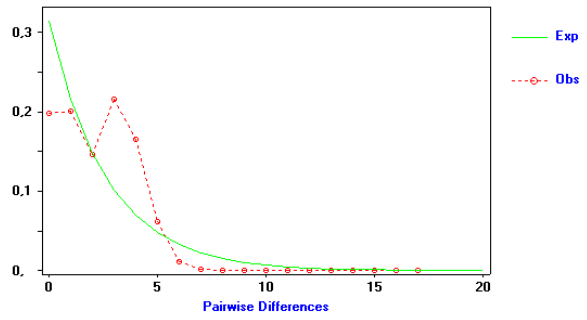


Figura VII.20. *Pairwise differences* de Nahuas de la Huasteca para la región HVII.

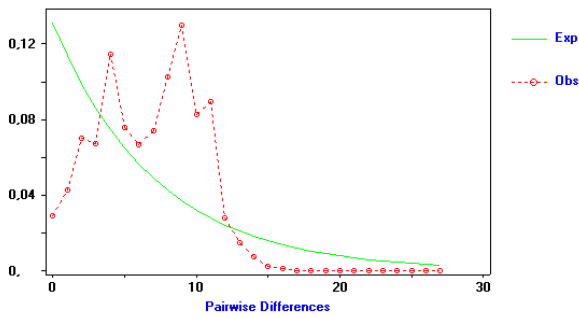


Figura VII.21. *Pairwise differences* de Otomíes para la región HVI.

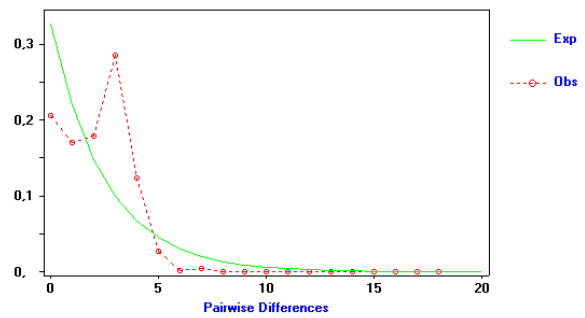


Figura VII.22. *Pairwise differences* de Otomíes para la región HVII.

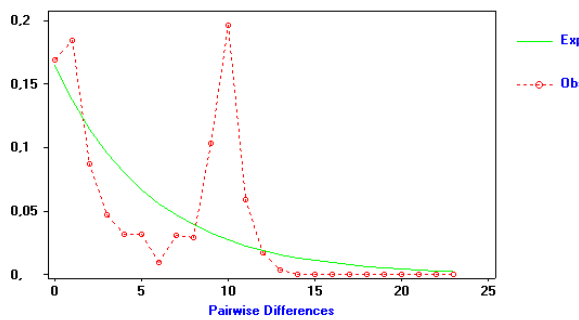


Figura VII.23. *Pairwise differences* de Tepehuas para la región HVI.

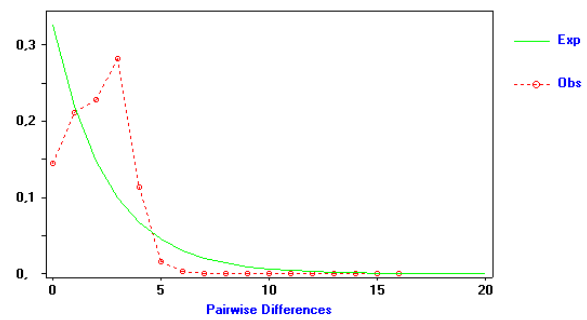


Figura VII.24. *Pairwise differences* de Tepehuas para la región HVII.

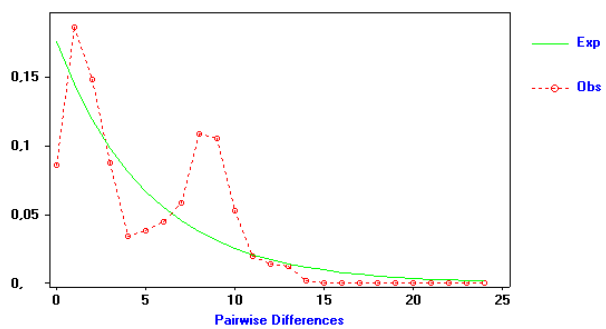


Figura VII.25. *Pairwise differences* de Mayas para la región HVI.

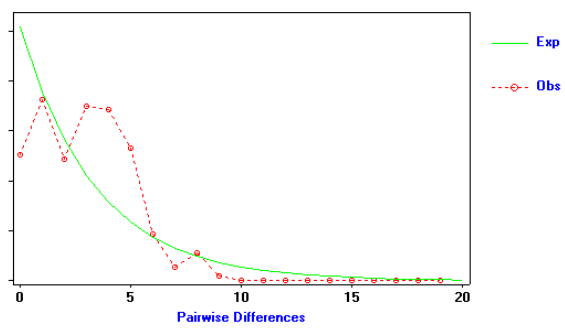


Figura VII.26. *Pairwise differences* de Mayas para la región HVII.

Tabla VII.2. Haplotipos únicos para cada una de las poblaciones indígenas considerando la D-Loop completa. La composición nucleotídica de los haplogrupos puede consultarse en la base de datos Haplotipos.xls (hoja: D-Loop)

Grupo	Haplotipos Únicos	Nº	%
Pimas	Hap_67, Hap_240, Hap_241, Hap_242, Hap_243, Hap_244, Hap_245, Hap_246, Hap_247, Hap_248, Hap_249, Hap_250, Hap_251, Hap_252, Hap_253, Hap_254	16	6,43
Mayos	Hap_1, Hap_139, Hap_140, Hap_141, Hap_142, Hap_143, Hap_144, Hap_145, Hap_146, Hap_147, Hap_148, Hap_149, Hap_150, Hap_151, Hap_152, Hap_153, Hap_154, Hap_155, Hap_156, Hap_157, Hap_158, Hap_159, Hap_160, Hap_161, Hap_162, Hap_163, Hap_164, Hap_165, Hap_166, Hap_167, Hap_168	31	12,45
Huicholes	Hap_93, Hap_94, Hap_95, Hap_96, Hap_97, Hap_98, Hap_99, Hap_101, Hap_102, Hap_103, Hap_104, Hap_105, Hap_106, Hap_107, Hap_108, Hap_109, Hap_110, Hap_111, Hap_112, Hap_113, Hap_114	21	8,43
Nahuas Huasteca	Hap_4, Hap_5, Hap_6, Hap_7, Hap_9, Hap_10, Hap_11, Hap_12, Hap_13, Hap_15, Hap_16, Hap_17, Hap_18, Hap_19, Hap_20, Hap_23, Hap_24, Hap_26, Hap_27, Hap_28, Hap_29, Hap_30, Hap_31, Hap_32, Hap_33, Hap_34, Hap_35, Hap_36, Hap_37, Hap_38, Hap_39, Hap_41, Hap_42, Hap_43, Hap_46, Hap_47, Hap_48, Hap_49, Hap_51, Hap_52, Hap_53, Hap_54, Hap_55, Hap_56, Hap_57, Hap_58, Hap_59, Hap_60, Hap_61, Hap_62, Hap_63, Hap_64, Hap_65, Hap_66, Hap_68, Hap_69, Hap_70, Hap_71, Hap_72, Hap_73, Hap_74, Hap_76, Hap_77, Hap_78, Hap_79, Hap_81, Hap_82, Hap_83, Hap_84, Hap_85, Hap_87, Hap_89, Hap_91, Hap_92	74	29,72
Otomíes Valle	Hap_169, Hap_170, Hap_171, Hap_172, Hap_173, Hap_174, Hap_175, Hap_176, Hap_177, Hap_178, Hap_179, Hap_180, Hap_181, Hap_182, Hap_183, Hap_184, Hap_185, Hap_186, Hap_187, Hap_188, Hap_193, Hap_194, Hap_195, Hap_196, Hap_197, Hap_198	26	10,44
Otomíes Sierra	Hap_200, Hap_201, Hap_202, Hap_203, Hap_204, Hap_205, Hap_206, Hap_207, Hap_208, Hap_209, Hap_210, Hap_211, Hap_212, Hap_213, Hap_214, Hap_215, Hap_216, Hap_217, Hap_218, Hap_219, Hap_220, Hap_221, Hap_222, Hap_223, Hap_225, Hap_226, Hap_227, Hap_228, Hap_229, Hap_230, Hap_231, Hap_232, Hap_233, Hap_234, Hap_235, Hap_236, Hap_237, Hap_238, Hap_239	39	15,66
Tepehuas	Hap_255, Hap_256, Hap_257, Hap_258, Hap_259, Hap_260, Hap_261, Hap_262, Hap_263, Hap_264, Hap_265, Hap_266, Hap_267, Hap_268, Hap_269, Hap_270, Hap_271, Hap_272	18	7,23
Mayas	Hap_115, Hap_116, Hap_117, Hap_118, Hap_119, Hap_120, Hap_121, Hap_122, Hap_123, Hap_124, Hap_125, Hap_126, Hap_127, Hap_128, Hap_129, Hap_130, Hap_131, Hap_132, Hap_133, Hap_134, Hap_135, Hap_136, Hap_137, Hap_138	24	9,64

Tabla VII.3. Haplotipos únicos para cada una de las poblaciones indígenas considerando la HVI. La composición nucleotídica de los haplogrupos puede consultarse en la base de datos Haplotipos.xls (hoja: HVI)

Grupo	Haplotipos Únicos	Nº	%
Pimas	Hap140, Hap141, Hap142	3	2,61
Mayos	Hap_88, Hap_89, Hap_90, Hap_91, Hap_92, Hap_94, Hap_95, Hap_96, Hap_97, Hap_98, Hap_99, Hap_100, Hap_101, Hap_102	14	12,17
Huicholes	Hap_60, Hap_61, Hap_62, Hap_63, Hap_64, Hap_65, Hap_66, Hap_67, Hap_68, Hap_69, Hap_70, Hap_71, Hap_72	13	11,30
Nahuas Huasteca	Hap_5, Hap_6, Hap_8, Hap_9, Hap_11, Hap_12, Hap_14, Hap_20, Hap_21, Hap_22, Hap_23, Hap_24, Hap_28, Hap_29, Hap_30, Hap_33, Hap_34, Hap_35, Hap_36, Hap_37, Hap_39, Hap_40, Hap_41, Hap_43, Hap_44, Hap_45, Hap_46, Hap_47, Hap_49, Hap_50, Hap_54, Hap_56, Hap_58, Hap_59	34	29,75
Otomíes Valle	Hap_103, Hap_104, Hap_106, Hap_107, Hap_108, Hap_109, Hap_110, Hap_111, Hap_112, Hap_114, Hap_115, Hap_119, Hap_120, Hap_121, Hap_123	15	13,04
Otomíes Sierra	Hap_124, Hap_125, Hap_126, Hap_127, Hap_128, Hap_129, Hap_130, Hap_131, Hap_132, Hap_133, Hap_134, Hap_135, Hap_136, Hap_137, Hap_138, Hap_139	16	13,91
Tepehuas	Hap_143, Hap_144, Hap_145, Hap_146, Hap_147, Hap_148	6	5,22
Mayas	Hap_73, Hap_74, Hap_75, Hap_76, Hap_77, Hap_79, Hap_80, Hap_81, Hap_82, Hap_83, Hap_84, Hap_85, Hap_86, Hap_87	14	12,17

Tabla VII.4. Haplotipos únicos para cada una de las poblaciones indígenas considerando la HVII. La composición nucleotídica de los haplogrupos puede consultarse en la base de datos Haplotipos.xls (hoja: HVII)

Grupo	Haplotipos Únicos	Nº	%
Pimas	Hap_70, Hap_71, Hap_71	3	4,84
Mayos	Hap_11, Hap_44, Hap_45, Hap_46, Hap_48, Hap_49, Hap_50, Hap_51, Hap_52, Hap_56	10	16,13
Huicholes	Hap_27, Hap_28, Hap_29, Hap_30, Hap_31, Hap_32, Hap_34,	7	11,29
Nahuas Huasteca	Hap_2, Hap_5, Hap_6, Hap_9, Hap_10, Hap_13, Hap_15, Hap_16, Hap_17, Hap_19, Hap_20, Hap_22, Hap_23, Hap_24, Hap_25, Hap_26	16	25,81
Otomíes Valle	Hap_57, Hap_58, Hap_59, Hap_60, Hap_62	5	8,06
Otomíes Sierra	Hap_63, Hap_64, Hap_65, Hap_66, Hap_67, Hap_68, Hap_69	7	11,29
Tepehuas	Hap_73, Hap_74, Hap_75, Hap_76, Hap_77	5	8,06
Mayas	Hap_35, Hap_36, Hap_37, Hap_38, Hap_39, Hap_40, Hap_41, Hap_42, Hap_43	9	14,52

Tabla VII.5. Haplotipos compartido por los ocho grupos indígenas para D-Loop. En la hemimatriz inferior se detallan los nombres de los haplotipos compartidos –para determinar su composición nucleotídica consultar la base de datos adjunta al trabajo Haplotipos.xls- en la superior el número absoluto

	Pimas	Mayos	Huicholes	Nahuas Huasteca	Otomíes Valle	Otomíes Sierra	Tepehuas	Mayas
Pimas		0	0	1	1	0	1	0
Mayos	-		2	2	0	1	0	0
Huicholes	-	Hap02,Hap100		3	1	1	1	0
Nahuas Huasteca	Hap67	Hap01, Hap02	Hap02, Hap44, Hap75		8	7	5	2
Otomíes Valle	Hap192	-	Hap44	Hap14, Hap21 Hap22, Hap44 Hap50, Hap80 Hap86, Hap88 Hap02, Hap03		7	3	1
Otomíes Sierra	-	Hap02	Hap02	Hap21, Hap25 Hap45, Hap80 Hap88	Hap21, Hap80 Hap88, Hap189 Hap190, Hap191 Hap199		1	1
Tepehuas	Hap192	-	Hap44	Hap08, Hap22 Hap40, Hap244 Hap90	Hap22, Hap44 Hap192	Hap224		1
Mayas	-	-	-	Hap08, Hap21	Hap21	Hap21	Hap08	

Tabla VII.6. Haplotipos compartido por los ocho grupos indígenas para HVI. En la hemimatriz detallan los nombres de los haplotipos compartidos –para determinar su composición nucleotídica consultar la base de datos adjunta al trabajo **Haplotipos.xls**- en la superior el número absoluto

	Pimas	Mayos	Huicholes	Nahuas Huasteca	Otomíes Valle	Otomíes Sierra	Tepehuas	Mayas
Pimas		3	1	2	2	1	2	2
Mayos	Hap32, Hap51 Hap57		3	6	3	4	3	4
Huicholes	Hap42	Hap02, Hap32 Hap42		5	3	4	3	3
Nahuas Huasteca	Hap42, Hap51	Hap01, Hap02 Hap32, Hap42 Hap51, Hap57	Hap02, Hap25 Hap32, Hap42 Hap48		13	14	8	9
Otomíes Valle	Hap51, hap118	Hap32, Hap51 Hap57	Hap25, Hap32 Hap48	Hap07, Hap13 Hap17, Hap19 Hap25, Hap27 Hap31, Hap32 Hap48, Hap51 Hap53, Hap55 Hap57		14	6	4
Otomíes Sierra	Hap51	Hap02, Hap32 Hap51, Hap57	Hap02, Hap25 Hap32, Hap48	Hap02, Hap03 Hap04, Hap07 Hap10, Hap13 Hap25, Hap27 Hap32, Hap38 Hap48, Hap51 Hap55, Hap57	Hap07, Hap13 Hap25, Hap27 Hap32, Hap48 Hap51, Hap55 Hap57, Hap105 Hap113, Hap116 Hap117, Hap122		7	4
Tepehuas	Hap51, Hap118	Hap02, Hap51 Hap57	Hap02, Hap25 Hap48	Hap02, Hap07 Hap25, Hap26 Hap38, Hap48 Hap51, Hap57	Hap07, Hap25 Hap48, Hap51 Hap57, Hap118	Hap02, Hap07 Hap25, Hap38 Hap48, Hap51 Hap57		3
Mayas	Hap42, Hap51	Hap02, Hap32 Hap42, Hap51	Hap02, Hap32 Hap42	Hap15, Hap16 Hap18, Hap32 Hap42, Hap51 Hap52	Hap07, Hap32 Hap51, Hap78	Hap02, Hap7 Hap32, Hap51	Hap02, Hap07 Hap51	

Tabla VII.7. Haplotipos compartido por los ocho grupos indígenas para HVII. En la hemimatriz detallan los nombres de los haplotipos compartidos –para determinar su composición nucleotídica consultar la base de datos adjunta al trabajo **Haplotipos.xls**- en la superior el número absoluto

	Pimas	Mayos	Huicholes	Nahuas Huasteca	Otomíes Valle	Otomíes Sierra	Tepehuas	Mayas
Pimas		5	4	5	3	4	3	3
Mayos	Hap01, Hap12 Hap18, Hap53 Hap55		2	4	3	3	4	2
Huicholes	Hap01, Hap12 Hap14, Hap33	Hap01, Hap12		3	2	2	2	3
Nahuas Huasteca	Hap01, Hap12 Hap14, Hap18 Hap21	Hap01, Hap_11 Hap12, Hap_18	Hap01, Hap12 Hap14		7	7	5	6
Otomíes Valle	Hap01, Hap12 Hap18	Hap01, Hap12 Hap18	Hap01, Hap12	Hap01, Hap03 Hap04, Hap07 Hap08, Hap12 Hap18		7	6	5
Otomíes Sierra	Hap01, Hap12 Hap18, Hap21	Hap01, Hap12 Hap18	Hap01, Hap12	Hap01, Hap04 Hap07, Hap08 Hap12, Hap18 Hap21	Hap01, Hap04 Hap07, Hap08 Hap12, Hap18 Hap61		5	4
Tepehuas	Hap01, Hap12 Hap18	Hap01, Hap12 Hap18, Hap47	Hap01, Hap12	Hap01, Hap03 Hap04, Hap12 Hap18	Hap01, Hap03 Hap04, Hap12 Hap18, Hap61	Hap01, Hap04 Hap12, Hap18 Hap61		4
Mayas	Hap01, Hap12 Hap14	Hap01, Hap12	Hap01, Hap12 Hap14	Hap01, Hap03 Hap04, Hap07 Hap12, Hap14	Hap01, Hap03 Hap04, Hap07 Hap12	Hap01, Hap04, Hap07, Hap12	Hap01, Hap03 Hap04, Hap12	

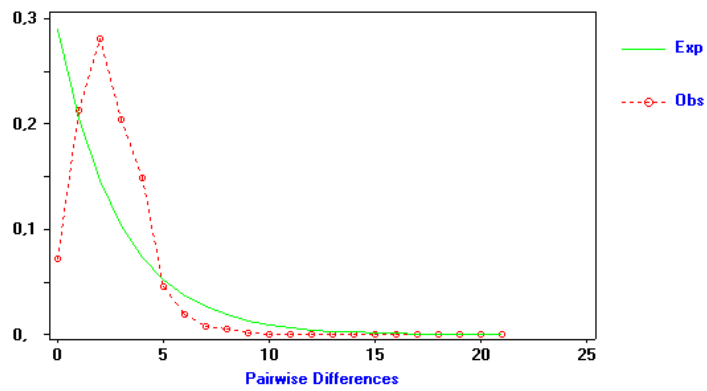
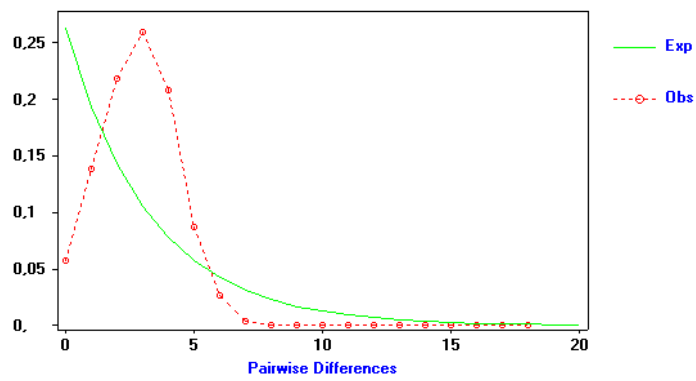
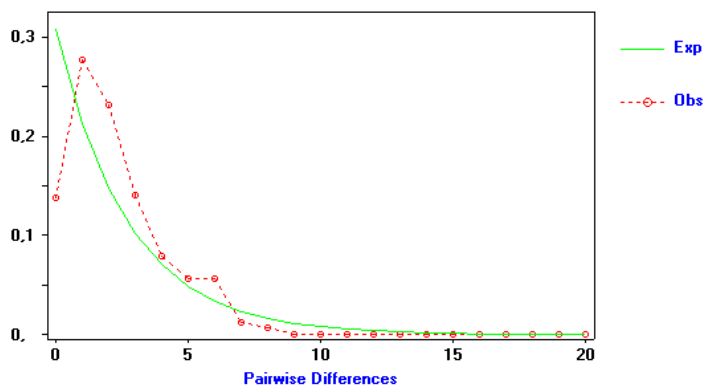
Figura VII.27. *Pairwise differences* para el haplogrupo A en base a datos de D-Loop.Figura VII.28. *Pairwise differences* para el haplogrupo B en base a datos de D-Loop.Figura VII.29. *Pairwise differences* para el haplogrupo C en base a datos de D-Loop.

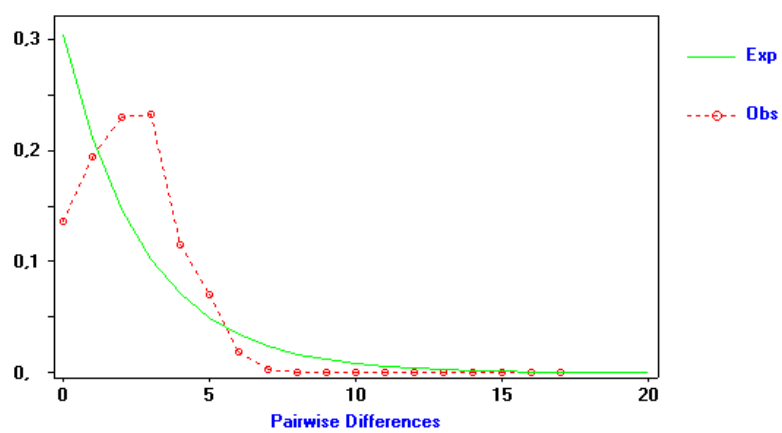
Figura VII.30. *Pairwise differences* para el haplogrupo D en base a datos de D-Loop.

Tabla VII.8. Nombre de los grupos indígenas, origen (comunidad tomada como referencia, estado y país), ubicación en regiones culturales, tamaño muestras, frecuencia de haplogrupos y referencia bibliográfica.

Grupo	Procedencia	Region	n	A	B	C	D	Otros	Latitud	Longitud	Referencia
Quechan	Yuman, (California, USA)	Aridoamerica	23	0,000	0,652	0,304	0,000	0,043	39,513	-121,557	Lorenz and Smith, 1996
Jemez	Los Álamos, (Nuevo México, USA)	Oasisamerica	36	0,000	0,889	0,028	0,000	0,083	35,888	-106,306	Lorenz and Smith, 1996; Malhi et al., 2003
Navajos	Window Rock, (Arizona, USA)	Oasisamerica	48	0,517	0,414	0,034	0,000	0,034	35,681	-109,052	Torrioni et al., 1993a; Lorenz and Smith, 1996
Zuni	Gallup, (Nuevo México, USA)	Oasisamerica	26	0,154	0,769	0,077	0,000	0,000	35,528	-108,742	Lorenz and Smith, 1996; Malhi et al., 2003
Apache	Peridot, (Arizona, USA)	Oasisamerica	29	0,621	0,172	0,138	0,069	0,000	33,310	-110,455	Torrioni et al., 1993a; Lorenz and Smith, 1996
Pima	Sacaton, (Arizona, USA)	Oasisamerica	43	0,047	0,535	0,395	0,000	0,023	33,083	-111,733	Torrioni et al., 1993; Lorenz and Smith, 1996; Malhi et al., 2003
Papagos	Caborca (Sonora, México)	Oasisamerica	37	0,000	0,568	0,378	0,054	0,000	30,717	-112,150	Malhi et al., 2003
Pimas	Yepachi (Chihuahua, México)	Aridoamerica	49	0,020	0,143	0,816	0,020	0,000	28,421	-108,377	Presente tesis
Tarahumaras	Bocoyna, (Chihuahua, México)	Aridoamerica	53	0,094	0,415	0,377	0,094	0,019	27,833	-107,583	Peñaloza-Espinosa et al., 2007
Mayos	Los Capomos (Sinaloa, México)	Aridoamerica	55	0,236	0,309	0,127	0,327	0,000	26,425	-108,515	Presente tesis
Huicholes	San Sebastian Teponahuastlan (Jalisco, México)	Mesoamérica	36	0,556	0,250	0,194	0,000	0,000	22,080	-104,063	Presente tesis
Nahua1	Coyulillo (Veracruz, Méxiuco)	Mesoamérica	38	0,684	0,079	0,000	0,158	0,079	21,750	-97,817	Peñaloza-Espinosa et al., 2007
Huicholes	Mesquitic (Jalisco, México)	Mesoamerica	50	0,460	0,260	0,260	0,000	0,000	21,750	-103,850	Sandoval et al., 2006
Nahuas Huasteca	Tlalchiyahualica (Hidalgo, México)	Mesoamerica	193	0,570	0,280	0,093	0,047	0,010	20,970	-98,399	Presente tesis
Nahua2	(México)	Mesoamerica	32	0,531	0,344	0,063	0,000	0,063	20,900	-98,200	Lorenz and Smith, 1996
Maya	Valladolid, (Campeche, México)	Mesoamerica	27	0,556	0,222	0,148	0,074	0,037	20,683	-88,200	Schurr et al., 1990; Torrioni et al., 1992
Otomies Valle	San Andrés Daboxtlá (Hidalgo, México)	Mesoamerica	81	0,494	0,148	0,272	0,086	0,000	20,525	-99,063	Presente tesis
Otomies1	Ixmiquilpan (Hidalgo, México)	Mesoamerica	68	0,405	0,300	0,230	0,065	0,000	20,483	-99,233	Sandoval et al., 2006
Tepehuas	Huehuetla (Hidalgo, México)	Mesoamerica	48	0,639	0,230	0,049	0,033	0,049	20,459	-98,076	Presente tesis
Otomies Sierra	San Bartolo Tutotepec (Hidalgo, México)	Mesoamerica	91	0,538	0,110	0,242	0,110	0,000	20,400	-98,218	Presente tesis
Otomies	Actopan (Hidalgo, México)	Mesoamérica	35	0,600	0,200	0,114	0,086	0,000	20,267	-98,933	Peñaloza-Espinosa et al., 2007
Mayas	Sabán (Quintana Roo, México)	Mesoamérica	58	0,741	0,155	0,069	0,035	0,000	20,033	-88,538	Presente tesis
Purepechas	Janitzió (Michoacan, México)	Mesoamerica	50	0,440	0,240	0,180	0,000	0,120	19,550	-101,667	Sandoval-Ramirez et al., 2006
Nahua3	Xochimilco (Edo Mexico, México)	Mesoamérica	43	0,721	0,186	0,093	0,000	0,000	19,262	-99,108	Peñaloza-Espinosa et al., 2007
Nahua4	Ixhuatlancillo (Veracruz, México))	Mesoamérica	47	0,553	0,277	0,000	0,170	0,000	18,900	-97,150	Peñaloza-Espinosa et al., 2007
Nahua5	Necoxtla (Puebla, México)	Mesoamérica	37	0,514	0,405	0,081	0,000	0,000	18,600	-98,367	Peñaloza-Espinosa et al., 2007
Nahua6	Chilacachapa (Queretaro, México)	Mesoamérica	41	0,463	0,341	0,073	0,122	0,000	18,283	-99,717	Peñaloza-Espinosa et al., 2007
Nahua7	Zitlala (Guerrero, México)	Mesoamérica	46	0,652	0,304	0,022	0,022	0,000	17,633	-99,083	Peñaloza-Espinosa et al., 2007
Triquis	San Miguel Copala (Oaxaca, México)	Mesoamerica	107	0,700	0,300	0,000	0,000	0,000	17,063	-97,930	Sandoval et al., 2006
Tzeltal	Pantelho, (Chiapas, México)	Mesoamérica	35	0,600	0,143	0,143	0,114	0,000	16,933	-92,550	Peñaloza-Espinosa et al., 2007
Mixteca Baja	Huajuapán, (Oaxaca, México)	Mesoamerica	29	0,924	0,071	0,000	0,000	0,000	16,183	-94,317	Torrioni et al., 1994c
Huetar	Quitirrisi (Costa Rica)	Centroamérica	27	0,704	0,037	0,000	0,259	0,000	10,467	-84,017	Santos et al., 1994
Bribri-Cabecar	Talamanca (Costa Rica)	Centroamérica	24	0,542	0,458	0,000	0,000	0,000	9,600	-82,783	Barrantes et al., 1990
Wounan	La Palma (Panamá)	Centroamérica	31	0,194	0,290	0,484	0,032	0,000	8,403	-78,145	Kolman and Bermingham, 1997
Ngöbe	Tolé (Panamá)	Centroamérica	46	0,674	0,326	0,000	0,000	0,000	8,233	-81,683	Kolman et al., 1995