

# **Métodos**



## MÉTODOS

### 1. OBTENCIÓN DE LAS MUESTRAS

En la presente tesis doctoral se trabajó preferentemente con piezas dentales puesto que éstas ofrecen mayores garantías a la hora de obtener información genética (KUROSAKI *et al.* 1993; DeGUSTA *et al.* 1994; WOODWARD *et al.* 1994b; OOTA *et al.* 1995; ZIERDT *et al.* 1996). Sin embargo en algunos casos también se obtuvo material genético a partir de huesos.

El material fue enviado directamente al laboratorio desde su ubicación original. Al no ser posible la participación directa del investigador en las excavaciones se informó al arqueólogo de las precauciones en materia de manipulación para minimizar la contaminación de los restos. En este sentido se evitó en la medida de lo posible tocar directamente los restos con las manos mediante el uso de guantes desechables. Una vez seleccionada la muestra ésta fue guardada en una bolsa de plástico o en un falcon de plástico estéril, rotulada y finalmente enviada al laboratorio para su análisis.

Una vez en el laboratorio se seleccionaron del conjunto enviado las muestras para el análisis de DNA. Siempre que fue posible se eligieron piezas dentales preferiblemente enteras, sin fisuras, y limpiadas superficialmente. No se escogió ningún tipo de pieza preferentemente, aunque se intentó trabajar con aquellas de desarrollo completo. Cuando no existían piezas dentales disponibles de un determinado resto de interés se seleccionaron zonas de hueso compacto, preferentemente de huesos largos, como tibias y fémures. En este caso se primó la selección de aquellos huesos que presentaban la superficie externa lisa, brillante y aparentemente bien conservada. Todo el proceso de manipulación fue realizado con guantes estériles.

La muestra definitiva fue introducida en un tubo falcon de plástico estéril convenientemente etiquetado con la referencia de laboratorio de la muestra. Las muestras así guardadas se conservaron en congelador a -20° C hasta su procesado.

## **2. PROSPECCIÓN MOLECULAR**

El objetivo de los métodos de “prospección molecular” (MONTIEL 2001) consiste en evaluar las posibilidades de éxito en la obtención de información genética de un determinado resto.

En la presente tesis doctoral fueron empleados tres de estos métodos de “prospección molecular”: la racemización de aminoácidos, la cuantificación de secuencias específicas de DNA mitocondrial y la detección de moléculas inhibidoras de la *PCR* en los extractos de DNA antiguo.

## 2.1. Contenido total y racemización de aminoácidos

### 2.1.1. Muestras estudiadas

El objetivo del presente estudio es comprobar empíricamente si existe una relación entre el contenido total de aminoácidos y su grado de racemización con la preservación del DNA tal y como ha sido sugerido por otros autores (POINAR *et al.* 1996). Adicionalmente se pretende evaluar la influencia de ciertas variables como el tipo de tejido, la antigüedad y procedencia geográfica de las muestras en estas estimas.

De las muestras ya estudiadas de las que existía un remanente fue seleccionado un subconjunto representativo de los diferentes yacimientos que comprendía aquellas de las que se había obtenido algún tipo de información genética (secuencias de mtDNA, medidas de cuantificación...).

Se estimaron las cantidades de las isoformas L y D de 9 aminoácidos: Alanina (Ala), Valina (Val), Isoleucina (Ile), Prolina (Pro), Leucina (Leu), Ácido Aspártico (Asp), Fenilalanina (Phe), Ácido Glutámico (Glu) y Serina (Ser) en 39 muestras de tres tipos de tejidos (hueso, diente total y dentina) de 31 individuos de diferente cronología y localización geográfica (**Tablas ME1 y ME2**).

En 7 casos (TM5, TM6, H70, H35, H37, H57 y H43) se analizó la cantidad de estos aminoácidos en dos tejidos de una misma muestra: diente y dentina. Adicionalmente se calcularon estas mismas cantidades en dos muestras de diente total de un mismo individuo (R 65-4 II).

Este proceso fue llevado a cabo en colaboración con los profesores Dr. José Eugenio Ortiz y Dr. Trinidad de Torres del Departamento de Ingeniería Geológica de la Escuela Superior de Ingenieros de Minas de la Universidad Politécnica de Madrid.

**Tabla ME1.** Muestras para las que se estimó el contenido total y la racemización de ciertos aminoácidos.

<b>Procedencia</b>	<b>Individuo</b>	<b>Muestra</b>	<b>Tejido</b>
Abauntz	AB23C	AB9	Diente total
Abauntz	AB17C	AB14	Diente total
Abauntz	AB35E	AB17	Diente total
<b>Total Abauntz</b>	<b>3</b>	<b>3</b>	
Caldeirão	O10 107	CA7	Diente total
<b>Total Caldeirao</b>	<b>1</b>	<b>1</b>	
Dja'de	SK-R 3	3DJ6	Hueso
<b>Total Dja'de</b>	<b>1</b>	<b>1</b>	
Mari	1	1MA1	Diente total
Mari	8	1MA8	Diente total
Mari	12	1MA12	Diente total
<b>Total Mari</b>	<b>3</b>	<b>3</b>	
Nerja	NE-1829	5NE	Dentina
Nerja	NE-NM82.2	2NE	Dentina
<b>Total Nerja</b>	<b>2</b>	<b>2</b>	
Tell Halula	H-25	2H25	Dentina
Tell Halula	H-26	1H26	Dentina
Tell Halula	H-30	2H30	Dentina
Tell Halula	H-35	1H1	Dentina
Tell Halula	H-35	1H1	Diente total
Tell Halula	H-37	1H20	Dentina
Tell Halula	H-37	2H20	Diente total
Tell Halula	H-43	2H31	Dentina
Tell Halula	H-43	2H31	Diente total
Tell Halula	H-47	2H18	Dentina
Tell Halula	H-48	1H33	Dentina
Tell Halula	H-49	2H15	Dentina
Tell Halula	H-52	1H19	Dentina
Tell Halula	H-53	2H14	Dentina
Tell Halula	H-54	2H34	Dentina
Tell Halula	H-57	2H21	Dentina
Tell Halula	H-57	1H21	Diente total
Tell Halula	H-64	1H13	Dentina
Tell Halula	H-65	1H10	Dentina
Tell Halula	H-70	1H11	Dentina
Tell Halula	H-70	1H11	Diente total
<b>Total Tell Halula</b>	<b>16</b>	<b>21</b>	
Tell Ramad	R65-4 II	TR9	Diente total
Tell Ramad	R65-4 II	TR11	Diente total
Tell Ramad	R65-C8-SE B	TR16	Diente total
Tell Ramad	R65-1S	TR18	Diente total
<b>Total Tell Ramad</b>	<b>3</b>	<b>4</b>	
Tres Montes	TM8	2TM6	Dentina
Tres Montes	TM8	1TM6	Diente total
Tres Montes	TM11	2TM5	Dentina
Tres Montes	TM11	1TM5	Diente total
<b>Total Tres Montes</b>	<b>2</b>	<b>4</b>	
<b>Total general</b>	<b>31</b>	<b>39</b>	

**Tabla ME2.** Tipos de muestras estudiadas de cada yacimiento.

Procedencia	Diente total	Dentina	Hueso	Total
Abauntz	3	-	-	3
Tres Montes	2	2	-	4
Nerja	-	1	1	2
Caldeirão	1	-	-	1
Tell Halula	5	16	-	21
Tell Ramad	2	-	2	4
Dja'de	-	-	1	1
Mari	3	-	-	3
<b>Total</b>	16	19	4	39

## 2.1.2. Precauciones

### 2.1.2.1. Precauciones en la obtención de la muestra

En condiciones normales no es necesario que la obtención de la muestra se realice en condiciones de esterilidad. Únicamente se requiere el uso de guantes para evitar la contaminación de la muestra con aminoácidos exógenos por contacto. Sin embargo, cuando se desea extraer material genético del resto en cuestión (sea polvo total de diente o hueso o dentina), deben emplearse las precauciones habituales para la manipulación de restos antiguos:

- trabajo en el interior de una campana de seguridad biológica
- uso de bata, mascarilla y guantes
- empleo de material estéril. A este efecto en la obtención de muestras de dentina las brocas de diamante se desechan tras cada uso.

### 2.1.2.2. Precauciones en el proceso de racemización

#### 2.1.2.2.1. Limpieza del material

La limpieza del material utilizado para el análisis pretende también evitar la contaminación por aminoácidos exógenos del manipulador o manipuladores. Para ello es suficiente seguir unas sencillas normas:

- La muestra y el material de laboratorio empleado para su análisis debe ser manipulado en todo momento con guantes de laboratorio. No es necesaria la manipulación en el interior de campanas de flujo biológico.

- El material de vidrio empleado para los análisis (excepto las pipetas Pasteur) se limpia por calcinación en un horno a 500°C durante 2 h.
- Los microtubos eppendorf de plástico y las pipetas Pasteur son de un solo uso y se desechan en cada muestra.
- Los tapones y septa de Teflon se lavan cuidadosamente con éter de petróleo, acetona y se enjuagan tres veces con agua ultralimpia.

#### 2.1.2.2.2. Pureza de los reactivos

- Toda el agua empleada en los análisis es de calidad Mili-Q de Millipore.
- Los ácidos fluorhídrico y clorhídrico y el anhídrido ácido trifluoracético son de Merck grado analítico. El cloruro de tionilo se obtuvo de Fluka AG.
- El alcohol isopropílico y el n-hexano son Merck calidad HPLC, el diclorometano es Merck calidad espectroscópica.

#### **2.1.3. Obtención de la muestra**

En la presente Tesis Doctoral se trabajó con tres tipos de material: **hueso** y/o **diente pulverizado** y **dentina**. Los dos primeros fueron obtenidos triturando la muestra en un molino refrigerado con nitrógeno líquido, tras limpieza superficial de la muestra mediante una arenadora (Sand- Blaster Modelo Base 1 Plus) que inyecta óxido de aluminio erosionando su parte más externa. La muestra de **dentina** se obtuvo perforando el diente en la raíz, justo por debajo de la línea de cimentación, con una broca de diamante acoplada a un taladro eléctrico marca Dremel. El polvo de cemento generado fue desechado. Para la recogida del polvo de la dentina se dispuso un trozo de papel de aluminio con la cara mate hacia arriba y se fue rascando con la broca dejando que el polvo cayera sobre el papel, procurando mover la broca a lo largo de la superficie del diente para evitar el calentamiento excesivo de una determinada región. Una vez se hubo obtenido una cantidad de polvo suficiente para el análisis (aproximadamente 20-40mg), la muestra fue envuelta en el mismo papel de aluminio y guardada hasta su procesamiento.



#### **2.1.4. Protocolo**

##### Diálisis

Una vez obtenidas las muestras se dializaron para eliminar los aminoácidos libres, no unidos a la matriz de colágeno. Las estimas de racemización corresponderán por tanto únicamente a aquel porcentaje de aminoácidos que forman parte de esta matriz (fracción colágeno insoluble o *IC*). Para la diálisis se siguió el siguiente procedimiento:

- Se cortaron tantos trozos de membrana de diálisis de 3.500 Daltons (Spectra/por mnco 3500 membrane) como muestras y se dejaron aproximadamente 2h en remojo en agua ultrapura.
- Con ayuda de una espátula se pesaron entre 20 y 40mg de muestra en una báscula de precisión dentro de tubos de ensayo de vidrio con tapón.
- Se añadieron 1ml de HCl 2N y 5ml de PBS (estabilizador de pH) a las muestras pulverizada agitando a continuación en el agitador vórtex.
- Se sacaron las membranas de diálisis del agua, se anudó uno de sus extremos, y se transfirió a su interior el contenido de cada tubo con ayuda de una pipeta Pasteur de vidrio. Se anudó entonces el otro extremo de la membrana y se introdujo en un matraz con agua ultrapura sujetándolo a su superficie con una pinza de madera o similar envuelta en parafilm.
- Se colocó cada matraz encima de un agitador magnético y se conectó de forma que la membrana se moviera en círculos.
- Se taparon los matraces con papel de aluminio y se dejaron 24h en agitación a temperatura ambiente.
- Se recuperó el volumen del interior de la membrana de diálisis y se transfirió con una pipeta Pasteur de vidrio a un matraz de vidrio.
- Se evaporó la muestra durante toda la noche en un rotavapor.

##### Hidrólisis

- Se disolvió el residuo seco en agua ultrapura.
- Se transfirió a tubos de ensayo con tapones de rosca recubiertos de Teflon y se añadió una mezcla de ácido clorhídrico fumante 12 N (2.9  $\mu$  l/mg de muestra) y de ácido clorhídrico 6 N (100  $\mu$ l).

- Se agitó la mezcla en el agitador vórtex.
- Se cerraron los tubos bajo atmósfera de nitrógeno.
- Se dejaron en un bloque calefactor a 100°C durante 20 h.

#### Desalinización

- Se transfirió la muestra a unos microtubos de análisis cónicos eppendorf de 1.5 ml con tapón y se añadió ácido fluorhídrico concentrado (1.25 µl/mg de muestra).
- Se mezcló bien con el vórtex.
- Se centrifugó la muestra durante 4 min en una centrífuga Eppendorf a máxima velocidad.
- Se transfirió el sobrenadante a otro Eppendorf nuevo de 1.5 ml y se congeló en nitrógeno líquido.
- Se evaporó el sobrenadante a continuación en un desecador de plástico.
- Se redisolvió el residuo en 100 µl de agua ultrapura y se mezcló a continuación en un agitador vórtex.
- Se centrifugó y se transfirió a un vial de vidrio de 2 ml con septa de Teflon y tapón de rosca.
- Se evaporó el agua en un desecador de plástico dejando los tapones sin acabar de desenroscar.

#### Derivatización

- La primera derivatización de los aminoácidos consiste en su esterificación. Para ello, se añadieron al residuo seco 250µl de cloruro de tionilo 3M. Debido a la elevada toxicidad del compuesto esta reacción debe ser llevada a cabo en el interior de una campana de seguridad biológica.
- Se creó una atmósfera de nitrógeno dentro de los viales y se cierran fuertemente.
- Se transfirieron los viales a un bloque calefactor ajustado a 100°C, dónde se dejaron durante 1 h.
- Posteriormente los viales, abiertos pero sin destapar, se llevaron a sequedad en el desecador de plástico al vacío.
- La segunda derivatización consiste en la N-trifluoracetilación. Para llevarla a cabo se añadieron 200 µl de anhídrido trifluoracético (25% en diclorometano) a cada muestra.
- Se creó una atmósfera de nitrógeno dentro de los viales y se cerraron fuertemente.

- Se calentaron a 100°C durante 5 min en el bloque calefactor.
- El diclorometano y el anhídrido trifluoroacético sin reaccionar se evaporaron bajo flujo de nitrógeno.
- Posteriormente las muestras se disolvieron en 100 µl de n-hexano y se transfirieron a los viales de inyección.

#### Inyección de las muestras en el Cromatógrafo de Gases

- Se inyectaron 0.2-4 µl de la muestra en un Cromatógrafo de Gases (GC) Hewlett-Packard 5890. Durante la inyección de las muestras se mantuvo el puerto a una temperatura constante de 215°C. Durante los primeros 75 s el aparato funcionó en modo *splitless* y a continuación se pasó a modo *split*.

Se empleó Helio como gas portador a una presión en cabeza de columna de 5.8 psi, y una columna de sílice fundida Chirasil-L-Val (0.25 µm x 25 m) de Chrompack. Los gradientes de temperatura empleados fueron los siguientes: 50°C (1 min), calentamiento a 40°C/min hasta 115°C, permanencia a 115°C durante 12 min, calentamiento a 3°C/min hasta 190°C, permanencia a 190°C durante 10 min, enfriamiento hasta 50°C permaneciendo a esta temperatura (80°C si el tiempo entre pinchazos es largo-noche). La detección se realizó con un detector del tipo NPD mantenido a 300°C.

#### Evaluación de los resultados

La integración de las áreas de los picos se realizó usando el programa de integración PEAK96 de Hewlett-Packard para PC.

## **2.2. Cuantificación de secuencias específicas de DNA mitocondrial por *Real Time PCR***

Diversos autores han propuesto que la obtención de información genética a partir de extractos de DNA antiguo que contengan un número de copias de DNA inferior al millar no proporciona resultados fiables ni reproducibles (GILBERT *et al.* 2003b; HANDT *et al.* 1996). Para tratar de determinar la fiabilidad de la información genética recuperada de las muestras antiguas estudiadas, se estimó el número de moléculas de DNA mitocondrial presentes en diferentes extractos mediante cuantificación absoluta por *Real Time PCR*, empleando para ello una sonda TaqMan MGB específica. Con este objeto se construyó un diseño específico que permitía la cuantificación de las mismas secuencias de la región Hipervariable I que posteriormente serían amplificadas por *PCR* y secuenciadas para determinar su variabilidad. La ventaja de este nuevo diseño de cuantificación es que permite estimar con bastante exactitud el número de las mismas moléculas que posteriormente serán “diana” de la reacción de *PCR*.

### **2.2.1. Principios teóricos de la cuantificación por *Real Time PCR***

La cuantificación por *Real Time PCR* permite monitorizar la reacción de *PCR* al mismo tiempo que éste tiene lugar. Otros métodos de cuantificación basados en la *PCR* miden la cantidad de producto generado al final de la reacción de amplificación y luego estiman la concentración de DNA de los extractos. Sin embargo en la cuantificación por *Real Time PCR* la detección comienza durante los primeros ciclos de amplificación, cuando la cantidad de producto amplificado sobrepasa un determinado umbral. Cuanto mayor es el número de copias presentes en el extracto de DNA, antes se sobrepasa este umbral y se inicia la detección.

Existen dos tipos de químicas disponibles para realizar ensayos de cuantificación por *Real Time PCR*, SYBR-GREEN I y TaqMan. El diseño de cuantificación empleado en el presente trabajo utiliza el segundo tipo de química. En este tipo de ensayos se diseña una sonda específica complementaria a una región del producto que se desea amplificar (sonda TaqMan). Esta sonda presenta un fluorocromo unido en el extremo 5' ó *reporter* (FAM<sup>TM</sup>, TET<sup>TM</sup> o VIC<sup>TM</sup>) y otro unido en el extremo 3' ó *quencher* (TAMRA<sup>TM</sup>). Cuando la sonda está intacta, la fluorescencia emitida por el *reporter* es

absorbida por el *quencher* mediante un proceso conocido como energía de transferencia de Förster (FÖRSTER 1948; LAKOWICZ 1983). Si la secuencia de DNA complementaria a la sonda está presente, se produce la unión entre ambos. Cuando la polimerasa inicia el proceso de copiado desde el extremo 3' del cebador degrada la sonda gracias a su actividad 5' nucleásica. Esta degradación hace que el *reporter* y el *quencher* se separen, produciéndose un incremento en la fluorescencia que es recogido y almacenado por el secuenciador. La separación de la sonda del DNA posibilita que el proceso de extensión del *primer* continúe hasta el final. Este proceso se repite en cada ciclo de la *PCR*, dando lugar a un incremento en la intensidad de la fluorescencia proporcional a la cantidad de producto generado (**Figura FME1**).

En ensayos de cuantificación que requieren de una elevada especificidad resulta recomendable el empleo de una variante de estas sondas TaqMan: las sondas TaqMan MGB. A diferencia de las sondas convencionales, éstas presentan un segundo *quencher* no fluorescente en el extremo 3' e incorporan además en este mismo extremo un grupo MGB (*Minor Groove Binder*). La función de este grupo es aumentar la temperatura de fusión ( $T_m$ ) de la sonda, permitiendo así el empleo de sondas cortas.

Además de los cebadores y la sonda la mezcla de reacción de *Real Time PCR* incorpora una “referencia pasiva”, que es un marcador que permite normalizar la señal emitida por el *reporter* durante la recogida de datos, y corregir así fluctuaciones en la fluorescencia debidas a cambios en la concentración o en el volumen de las diferentes reacciones.

El parámetro  $R_n$  se define como el cociente entre la intensidad de emisión del *reporter* y la intensidad de emisión de la referencia pasiva.

El término  $R_n^+$  define al parámetro  $R_n$  cuando la reacción incorpora moléculas de template, mientras que  $R_n^-$  es el valor de  $R_n$  en ausencia de template o durante los primeros ciclos de *PCR*, previamente al inicio del incremento en la señal de fluorescencia. La magnitud real de la señal generada durante un ensayo viene definida por incremento de  $R_n$  ( $\Delta R_n$ ), que es la diferencia entre los valores de  $R_n^+$  y  $R_n^-$ .

La relación entre estos tres términos se expresa de la siguiente forma:

$$\Delta R_n = (R_n^+) - (R_n^-)$$

dónde

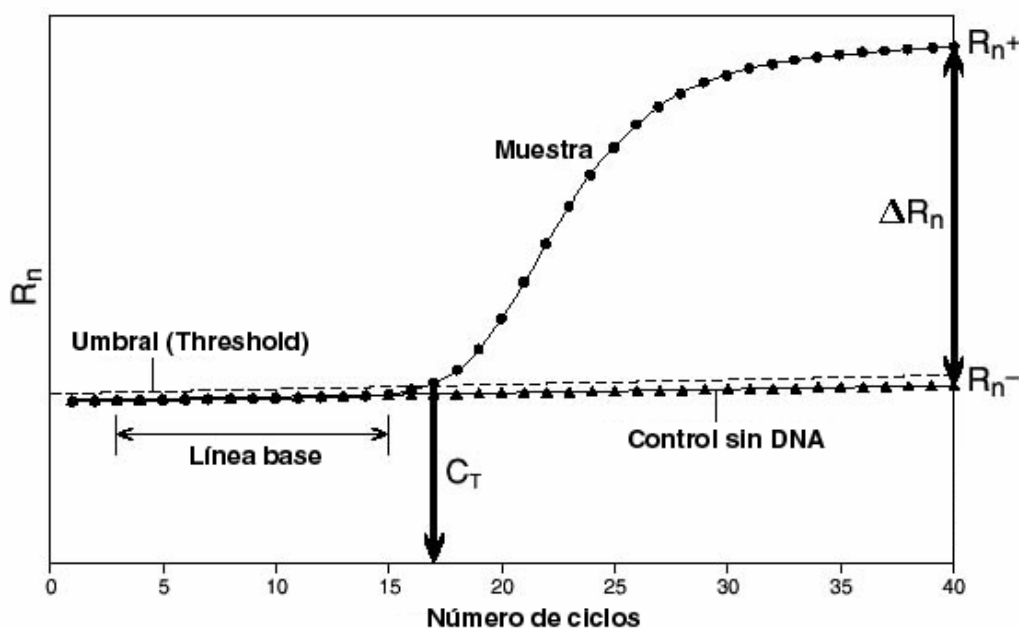
$$R_n^+ = \frac{\text{Intensidad de emisión del } reporter}{\text{Intensidad de emisión de la "referencia pasiva"}} \quad \text{en presencia de DNA molde}$$

$$R_n^- = \frac{\text{Intensidad de emisión del } reporter}{\text{Intensidad de emisión de la "referencia pasiva"}} \quad \text{en ausencia de DNA molde o durante los primeros ciclos de } PCR$$

Durante los primeros ciclos de *PCR* se produce muy poco cambio en la señal de fluorescencia, formándose la **línea base** para la curva de amplificación. Cuando se produce un incremento en la intensidad de la fluorescencia sobre esta línea base se inicia el proceso de detección. Ligeramente por encima de esta línea base puede fijarse un **umbral de detección de fluorescencia** (*Threshold*). El parámetro  $C_T$  (*Threshold Cycle*) o valor umbral se define como el ciclo de amplificación en el que tiene lugar este incremento significativo en la intensidad de la señal de fluorescencia ( $\Delta R_n$ ) y en el que se inicia la detección. Este umbral puede calcularse como la desviación estándar de  $R_n$  durante estos primeros ciclos multiplicada por un factor de ajuste.

Existen dos maneras de estimar la concentración de DNA de una muestra problema mediante *Real Time PCR*: la cuantificación absoluta y la cuantificación relativa. En la cuantificación absoluta la concentración de DNA de las muestras que se desea amplificar se estima intrapolando sus valores de  $C_T$  en una recta patrón construida con estándares de concentración conocida. Este método requiere de la cuantificación previa del stock de DNA utilizado para la construcción de la recta patrón mediante otros métodos como la espectrofotometría (medida de la Absorbancia a 260nm) o la espectrofluorimetría.

**Figura FME1.** Cinética de amplificación y parámetros empleados en la cuantificación por *Real Time PCR*. Modificada de "TaqMan® PCR Reagent Kit With AmpliTaq Gold® DNA Polymerase: Protocol: Rev D".(2003).



### 2.2.2. Muestras estudiadas

Se cuantificaron mediante *Real Time PCR* un total de 103 extractos de DNA antiguo pertenecientes a 85 individuos de diferente procedencia y cronología (**Tabla ME3**). Fueron seleccionados aquellos extractos de los que existía un volumen remanente que ya habían sido amplificados por *PCR*, independientemente del resultado obtenido.

Cada muestra fue cuantificada por duplicado dentro de una misma amplificación de *Real Time PCR*. La amplificación se repitió en aquellos casos en los que existía una gran dispersión entre los valores de cuantificación obtenidos de cada una de las réplicas (muestras 5NE y TM4), o en aquellos en los que había fallado la amplificación de una o ambas réplicas (muestra 3NE).

**Tabla ME3.** Muestras en las que se realizó una cuantificación del DNA por *Real Time PCR*.

Procedencia	Individuo	Muestra	Procedencia	Individuo	Muestra	Procedencia	Individuo	Muestra
Abauntz	AB27	AB3	Nerja	NE-1829	5NE	Tell Halula	H-70	3H11
Abauntz	ABb1.1	AB5	Nerja	NE-NM82.2	2NE	Tell Ramad	R65-7I	TR1
Abauntz	ABb1.2	AB6	Nerja	NE-NT79	3NE	Tell Ramad	R63-1	TR2
Abauntz	AB29F	1AB8	Nerja	NE-NAP	4NE	Tell Ramad	R65(8)	TR3
Abauntz	AB23C	AB9	El Pirulejo	1PI	1PI	Tell Ramad	R65(8)	TR4
Abauntz	AB21E	AB11	El Pirulejo	2PI	2PI	Tell Ramad	R65-10	TR5
Abauntz	AB21D	AB13	Sant Pau	2 7608	SP2	Tell Ramad	R65-10	TR6
Abauntz	AB17C	AB14	Sant Pau	6 7214	1SP6	Tell Ramad	R65-1	TR7
Abauntz	AB35F	AB15	Sant Pau	6 7214	2SP6	Tell Ramad	R65-3I	TR8
Abauntz	AB35E	AB17	Sant Pau	7 7720	1SP7	Tell Ramad	R65-4II	TR10
Abauntz	AB25E	AB20	Sant Pau	7 7720	2SP7	Tell Ramad	R65-4II	TR11
Abauntz	AB25D	AB21	Sant Pau	18 7758	1SP18	Tell Ramad	R65-14	TR12
Açikli	7 AH89	10SK4	Sant Pau	18 7758	2SP18	Tell Ramad	R65-C8-SEA	TR13
Açikli	8a AH89	16SK2	Sant Pau	19 7928	SP19	Tell Ramad	R69(2)	TR14
Açikli	AH95 2c-1	SK106	Tell Halula	H-3	TH1	Tell Ramad	R69(2)	TR15
Açikli	AH95 2c-3	SK109	Tell Halula	H-3	TH2	Tell Ramad	R65-C8-SEB	TR16
Caldeirão	O12 194	CA4	Tell Halula	H-4	TH3	Tell Ramad	R65-1S	TR18
Caldeirão	O10 107	CA7	Tell Halula	H-4	TH4	Tell Ramad	R66-N4-400	TR19
Dja'de	SK-A2	DJ1	Tell Halula	H-7	TH5	Toledo	Toledo K-13	TO1
Dja'de	SK-D4 16	DJ3	Tell Halula	H-7	TH6	Tres Montes	TM-2	2TM1
Dja'de	SK-N4	1DJ4	Tell Halula	H-8	TH7	Tres Montes	TM-3	1TM2
Dja'de	SK-R3	1DJ6	Tell Halula	H-8	TH8	Tres Montes	TM-5	TM3
Dja'de	SK-R3	2DJ6	Tell Halula	H-37	2H20	Tres Montes	TM-6	1TM4
Mari	1	1MA1	Tell Halula	H-45	1H17	Tres Montes	TM-6	2TM4
Mari	2	2MA2	Tell Halula	H-47	1H18	Tres Montes	TM-8	1TM5
Mari	3	1MA3	Tell Halula	H-49	1H15	Tres Montes	TM-11	1TM6
Mari	4	1MA4	Tell Halula	H-53	1H14	Tres Montes	TM-14	TM7
Mari	4	2MA4	Tell Halula	H-56	1H16	Tres Montes	TM-15	TM8
Mari	5	1MA5	Tell Halula	H-57	1H21	Tres Montes	TM-17	TM10
Mari	6	1MA6	Tell Halula	H-64	2H13	Tres Montes	TM-18	TM11
Mari	8	1MA8	Tell Halula	H-65	2H10	Tres Montes	TM-19	TM12
Mari	10	2MA10	Tell Halula	H-66	1H9	Tres Montes	TM-20	TM13
Mari	11	3MA11	Tell Halula	H-66	2H9	<b>Total</b>	<b>85</b>	<b>103</b>
Mari	12	1MA12	Tell Halula	H-68	H12			
Nerja	NE-1829	1NE	Tell Halula	H-70	2H11			

### 2.2.3. Diseño de los cebadores y de la sonda

La pareja de cebadores de la reacción de amplificación y la sonda TaqMan MGB fueron diseñados mediante el programa “Primer-Express” para MAC de Applied Biosystems en la Unitat de Sequenciació de los antiguos Serveis Científico-Tècnics, actualmente Unitat de Genòmica del Parc Científic de la Universitat de Barcelona. Dada la sensibilidad de la técnica, para su diseño fueron tenidos en cuenta un conjunto de parámetros:



- Longitud del amplicón comprendida entre 150-160 pares de bases, dentro de la zona de la región Hipervariable I del mtDNA amplificada en las muestras antiguas de la presente tesis doctoral.
- Mínima distancia entre la sonda y los cebadores sin llegar al solapamiento.
- Contenido de G-C de los cebadores y de la sonda entre 20-80%. Un elevado contenido en G-C dificulta la desnaturalización de las cadenas durante la amplificación, resultando en una disminución de la eficiencia de la reacción.
- Temperatura de fusión de los cebadores elevada, entre 58-60°C, para garantizar la especificidad de su unión.
- Temperatura de fusión ( $T_m$ ) de la sonda también elevada, de 8 a 10°C superior a la de los cebadores, entre 68-70°C.
- Ausencia de regiones de autocomplementariedad dentro de los cebadores, para evitar la formación de estructuras secundarias dentro de cada cebador y el apareamiento entre la pareja de cebadores (*primer-dimer*).
- Longitud de los cebadores entre 20-25 pares de bases.
- Ausencia de G en el extremo 5' de la sonda, puesto que dificulta su separación del DNA. Esto puede resultar en una disminución de los valores de fluorescencia normalizados ( $\Delta R_n$ ).
- Ausencia de más de dos G o C en los últimos 5 nucleótidos del extremo 3' de cualquiera de los dos cebadores, ya que favorece la formación de subproductos inespecíficos de la reacción de amplificación.
- Selección de la cadena de DNA que proporciona una sonda con un mayor contenido de Cs respecto a Gs debido a que observaciones empíricas han demostrado que esta combinación produce valores más elevados de  $\Delta R_n$ .

Tras introducir los datos en el programa, éste genera multitud de combinaciones primers-sonda y las ordena en base al valor óptimo. Para esta selección el programa tiene en cuenta la mayor parte de los parámetros mencionados. Sin embargo otros, como la selección de la cadena que proporcione un mayor contenido de Cs respecto de Gs en la sonda y la presencia de más de dos de estas Cs ó Gs en los últimos 5 nucleótidos de los cebadores, deben comprobarse manualmente. De forma adicional se evitó especialmente que en la zona de hibridación de los cebadores y de la sonda existieran puntos mutacionales frecuentes en poblaciones europeas, puesto que esto podía disminuir la

eficiencia de su hibridación y del proceso en general. Finalmente se seleccionó la siguiente combinación (**Tabla ME4**):

**Tabla ME4.** Características de los cebadores y de la sonda empleados para la cuantificación por *Real Time PCR*.

	Secuencia	Longitud	Tm	%GC	Longitud amplicón
<b>L16124 (forward)</b>	ACTGCCAGCCACCATGAATATT	22	59.3°C	45.5	149pb
<b>H16231 (reverse)</b>	TGGAGTTGCAGTTGATGTGTGA	22	58.8°C	45.5	
<b>Sonda</b>	AATACTTGACCACCTGTAGTAC	22	69°C		

En la **Figura FME2** se detallan la región de hibridación de estos cebadores y sonda respecto a los cebadores diseñados en esta Tesis Doctoral para el análisis de la variabilidad mutacional de la región Hipervariable I en muestras antiguas.

```

1
6
0
9
3   100   110   120   130   140   150   160
TTCGTACATTACTGCCAGCCACCATGAATATTGTACGGTTACCATAAATACTTGACCACCTGTAGTACCAT
AAGCATGTAATGACGGTCCGGTGGTACTTATAACATGCCATGGTATTTATGAACTGGTGGACATCATGTA
    170    180    190    200    210    220    230
AAAAACCCAATCCACATCAAAACCCCTCCCCATGCTTACAAGCAAGTACAGCAATCAACCCCTCAACTA
TTTTTGGGTTAGGTGTAGTTTTGGGGGAGGGGTACGAATGTTTCGTTTCATGTCGTTAGTTGGGAGTTGAT
    240    250    260    270    280    290    300
TCACACATCAACTGCAACTCCAAAGCCACCCCTCACCCACTAGGATACCAACAAACCTACCCACCCTTA
AGTGTGTAGTTGACGTTGAGGTTTCGGTGGGGAGTGGGTGATCCTATGGTTGTTGGATGGGTGGGAAT
0    310    320    330    340    350    360    370
ACAGTACATAAGTACATAAAGCCATTTACCGTACATAGCACATTACAGTCAAATCCCTTTCTCGTCCCCAT
TGTCATGTATCATGTATTTCCGGTAAATGGCATGTATCGTGTAATGTCAGTTTAGGGAAGAGCAGGGGTA
70    380    390    400    410    420    430
GGATGACCCCCCTCAGATAGGGGTCCCTTGACCACCATCCTCCGTGAAATCAATATCCCGCACAAAGAGT
CTACTGGGGGAGTCTATCCCCAGGGAACCTGGTGGTAGGAGGCACTTTAGTTATAGGGCGTGTCTCA

```

**Figura FME2.** Detalle de la zona de la región Hipervariable I del DNA mitocondrial estudiada en la presente tesis doctoral. Los ocho cebadores utilizados en este estudio aparecen marcados en cursiva y negrita (cebadores externos) y en cursiva (cebadores internos). Los cebadores utilizados para la cuantificación de DNA por *Real Time PCR* aparecen subrayados y la sonda enmarcada. Se resaltan en negrita en la cadena *forward* aquellas posiciones más variables dentro de las poblaciones europeas.

## 2.2.4. Obtención de estándares

### 2.2.4.1. Clonación del producto de amplificación

Para producir un stock de DNA de concentración conocida para la fabricación de la recta patrón se clonó el producto generado tras la amplificación de una muestra de DNA extraída de sangre con los cebadores externos L16112 y H16380, que proporcionan un

producto de amplificación de 352 pb. Este amplicón contiene la secuencia de 150 pb amplificada durante la *Real Time PCR*. Para ello se siguió el protocolo de clonación descrito en el correspondiente apartado de esta tesis doctoral.

#### 2.2.4.1.1. Cuantificación del producto

Una vez producido el producto clonado (plásmido + fragmento mtDNA) fue cuantificado mediante dos métodos independientes:

##### 2.2.4.1.1.1. Cuantificación en gel

El proceso experimental es el siguiente:

- Se lineariza el producto digiriéndolo con la enzima Sac I 1 hora a 37°C.
- Se fabrica un gel de agarosa al 0.7%, 0.5x TBE.
- Se cargan diferentes volúmenes de la muestra (1 y 2µl) junto con uno o dos marcadores de peso molecular (1µl λ HindIII 250ng y 1µl λ HindIII/Eco RI 250ng) en los que se conozca la cantidad de DNA de cada una de las bandas de peso molecular.
- Se corre el gel (50V, 2:30 horas) y se visualiza en un transiluminador.

La cuantificación se realiza comparando la intensidad de banda de las diferentes cantidades de muestra con la intensidad de las diferentes bandas de los marcadores. La cantidad de DNA presente en la muestra será aproximadamente igual a la que contenga aquella o aquellas bandas de los marcadores de intensidad similar.

Según este método, se estimó que la concentración del stock de DNA resultante de la clonación era de aproximadamente **80-90 nanogramos** (ng).

##### 2.2.4.1.1.2. Cuantificación por espectrofluorimetría

La cuantificación por espectrofluorimetría requiere de la fabricación de una recta patrón a partir de diluciones de un stock de DNA de concentración conocida.

- A partir de un stock de pGEM3Z de 50ng/µl se fabrican diluciones de 0, 100, 200, 300 y 400ng.

- Las diluciones del estándar y aquellas muestras que se desean cuantificar se incuban con DABA 200mg/μl durante 45 minutos a 60°C.
- Una vez transcurrido el tiempo, se para la reacción añadiendo HCl 1N.
- Se leen las muestras en un espectrofluorímetro.
- A partir de las diferentes diluciones de concentraciones conocidas y de los valores de espectrofluorimetría obtenidos se construye una recta patrón. La intrapolación de los valores de las muestras dentro de esta recta proporciona una estima de su concentración de DNA. En la **Tabla ME5** se muestran los diferentes valores de fluorimetría y concentración obtenidos para el conjunto de diluciones y muestras.

**Tabla ME5.** Valores de fluorimetría y concentración de DNA de los estándares y de las muestras.

	Volumen	Valor fluorimetría	Cantidad DNA (ng)	ng/ μl	Promedio ng/ μl
Dilución 1	1 μl	138.054	0	0	
Dilución 2	1 μl	183.649	100	100	
Dilución 3	1 μl	229.88	200	200	
Dilución 4	1 μl	276.41	300	300	
Dilución 5	1 μl	338.105	400	400	
Muestra	1 μl	175.135	82.148	82.148	
Muestra	1 μl	176.505	84.928	84.928	<b>87.195</b>
Muestra	2 μl	227.806	189.02	94.51	

#### 2.2.4.1.2. Estima del número de copias

Los dos métodos de cuantificación proporcionaron estimas similares de la concentración del stock de DNA, entre 80 y 90 ng/ μl el primero y en torno a 87 ng/ μl el segundo. Se tomó el valor de 87 ng/ μl como promedio de ambas medidas de cuantificación, cantidad que fue utilizada para el cálculo del número de moléculas por microlitro.

El DNA stock está formado por el plásmido más el inserto de DNA mitocondrial. En pares de bases esto es :

Plásmido: pBluescript + SK = 2961pb

Inserto: Región Control mtDNA = 305pb (16095-16399)

Plásmido recombinante = 2961 + 605 = 3266pb

Teniendo en cuenta que 1 mol de pares de bases pesa 3265g y que cada mol contiene  $6.023 \times 10^{23}$  moléculas, puede concluirse que el peso de cada molécula de plásmido recombinante es de aproximadamente  $3.58 \times 10^{-6}$  picogramos (pg).

Según esto, un stock original de concentración 87 ng/ $\mu$ l, contiene aproximadamente  **$24.3 \times 10^9$  moléculas**.

#### 2.2.4.2. Diluciones

A partir de este stock de  $24.3 \times 10^9$  moléculas se construyeron stocks de 200  $\mu$ l que contenían  $10^9$ ,  $10^8$ ,  $10^7$ ,  $10^6$ ,  $10^5$ ,  $10^4$ ,  $10^3$ ,  $10^2$ , 10 y 1 copias. Estos stocks fueron alicuotados y almacenados en el congelador a  $-20^\circ\text{C}$ . La concentración en nanogramos de cada uno de estos stocks es (**Tabla ME6**):

**Tabla ME6.** Concentración y número de moléculas de DNA de cada uno de los estándares empleados en la cuantificación por *Real Time PCR*.

Nº moléculas/ $\mu$ l	ng/ $\mu$ l
$10^9$	3.5
$10^8$	$3.5 \times 10^{-1}$
$10^7$	$3.5 \times 10^{-2}$
$10^6$	$3.5 \times 10^{-3}$
$10^5$	$3.5 \times 10^{-4}$
$10^4$	$3.5 \times 10^{-5}$
$10^3$	$3.5 \times 10^{-6}$
$10^2$	$3.5 \times 10^{-7}$
10	$3.5 \times 10^{-8}$
1	$3.5 \times 10^{-9}$
0	0

#### 2.2.5. Construcción de la recta patrón

Los diferentes estándares de concentración conocida (0- $10^9$  copias) se amplificaron siguiendo el protocolo descrito en el apartado 2.2.7.2. del presente capítulo. Una vez finalizada la reacción de amplificación se analizaron los resultados con el programa “SDS”. Este programa calcula una recta de regresión entre el número de copias de los diferentes estándares y sus valores correspondientes de Ct . Sobre esta recta se intrapolan los valores de Ct obtenidos para aquellas muestras que se desea cuantificar, obteniéndose el número de copias correspondiente. Para que esta estima resulte lo más exacta posible es

condición necesaria que exista una buena relación entre el número de copias de los estándares y sus valores de Ct. La medida de esta relación viene determinada por el valor del coeficiente de correlación ( $r^2$ ) calculado para cada recta de regresión.

Tras sucesivas amplificaciones de los diferentes estándares se comprobó que para valores inferiores a  $10^4$  copias se perdía la linearidad de la relación entre el número de copias y el valor de Ct. Esto puede estar relacionado con el límite de detección de la técnica, que para este tipo de material genético no permite cuantificar por debajo de  $10^4$  copias ( $3.5 \times 10^{-5}$  ng). También puede deberse a que existe un error en el proceso de construcción de los diferentes estándares mediante diluciones seriadas que se acentúa en aquellas diluciones de orden inferior, en este caso por debajo de  $10^4$  copias. Ante la imposibilidad de obtener valores fiables de cuantificación por debajo de  $10^4$  copias se optó por construir una recta patrón a partir de este orden de dilución, es decir, formada por los estándares  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$ .

#### 2.2.6. Optimización de las concentraciones de cebadores y de sonda

Mediante el ajuste de las concentraciones de cada uno de los cebadores y de la sonda empleados en la cuantificación por *Real Time PCR* puede determinarse aquella combinación que proporciona los menores valores de  $C_T$  y los mayores valores de  $\Delta R_n$ . Aunque los valores de  $C_T$  son el parámetro por el que se asignan los valores de cuantificación, los valores de  $\Delta R_n$  son importantes cuando se intenta obtener la máxima sensibilidad y reproducibilidad.

Para determinar la concentración óptima de cada uno de los cebadores fueron ensayadas por triplicado las siguientes 9 combinaciones de concentración con unas concentraciones fijas de sonda (250nM) y de DNA ( $10^8$  copias):

**Tabla ME7.** Concentraciones de cebadores ensayadas para optimizar la reacción de amplificación.

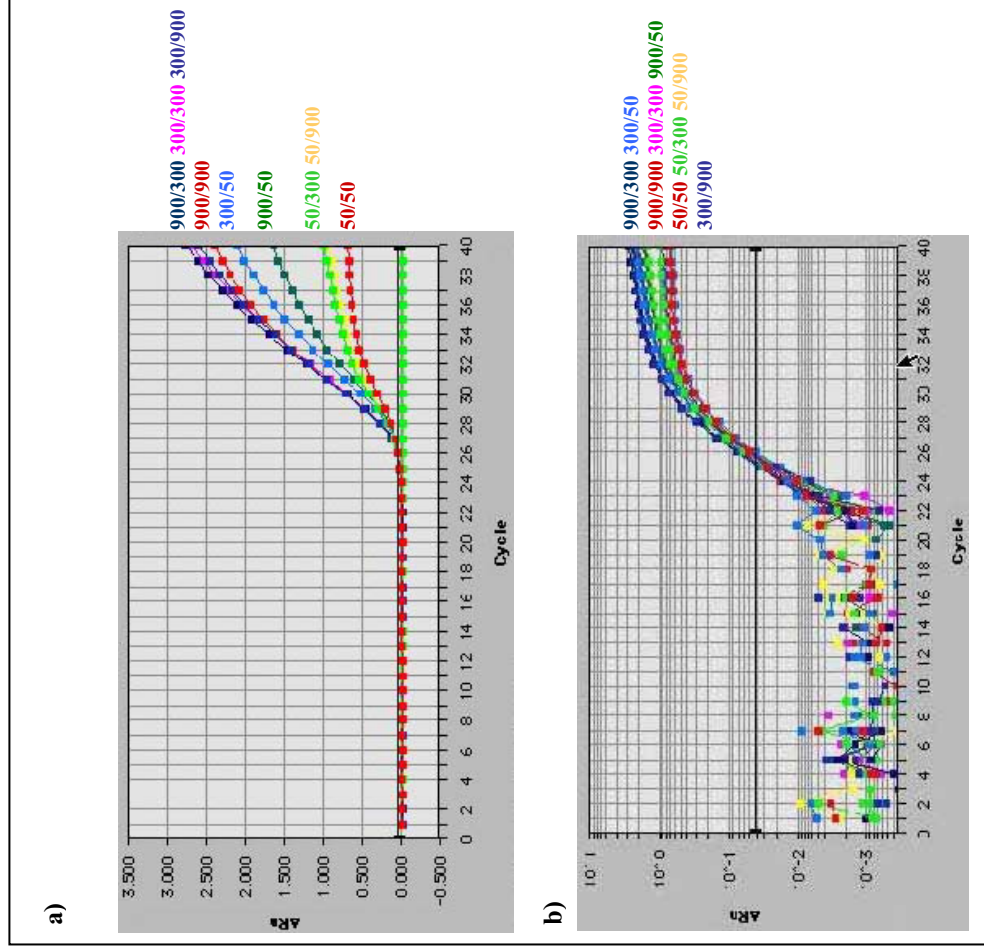
Primer Reverse	Primer Forward		
	50nM	300nM	900nM
50nM	50/50	50/300	50/900
300nM	300/50	300/300	300/900
900nM	900/50	900/300	900/900

Según los proveedores, el empleo de las concentraciones óptimas de cada uno de los cebadores definidas a partir del ensayo junto con una concentración de sonda de 250nM proporciona resultados inmejorables. En algunos casos, sin embargo, puede conseguirse la misma eficiencia con concentraciones menores de sonda, con un consecuente abaratamiento de los costes, de ahí la utilidad de definir un experimento de optimización de la concentración de so

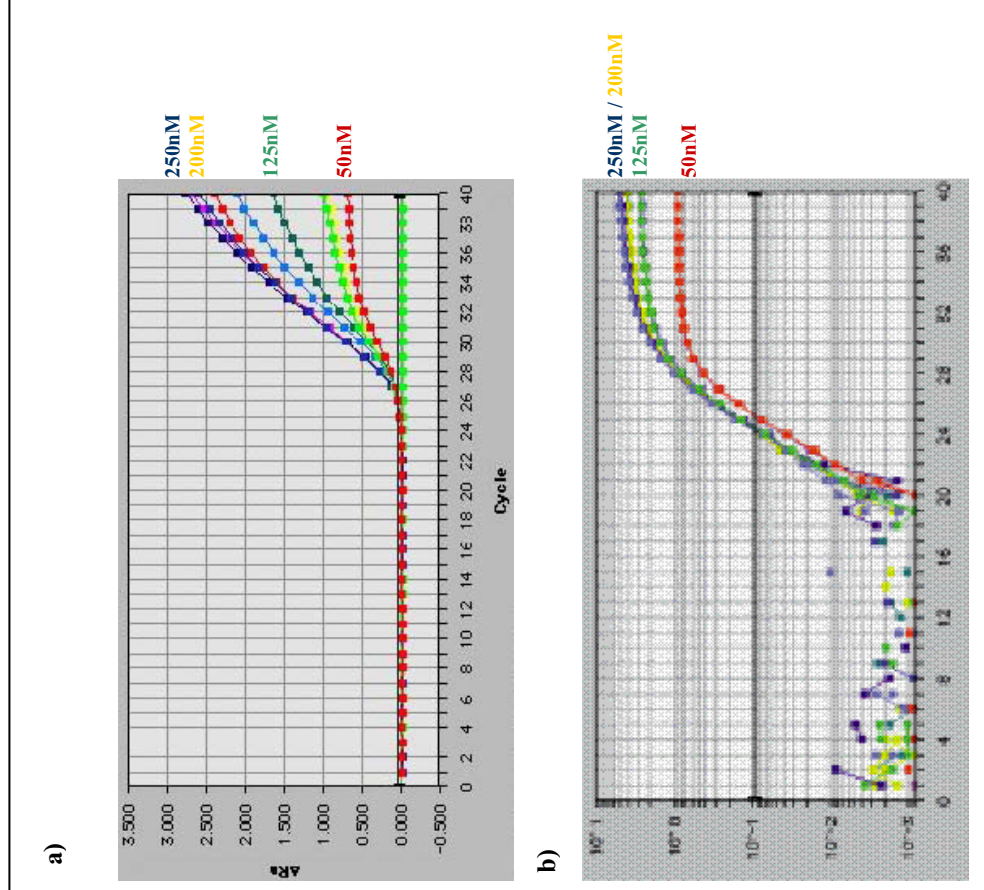
Para determinar la concentración óptima de sonda fueron ensayadas por triplicado concentraciones desde 50 a 250nM a intervalos de 50 con una concentración fija de DNA ( $10^8$  copias) y la cantidad óptima de cada uno de los cebadores definida en el experimento anterior.

En las **Figuras FME3 y FME4** pueden observarse los resultados de ambos experimentos de optimización. Aunque todas las combinaciones de concentración de los cebadores proporcionan valores de  $C_T$  similares, los valores de  $\Delta R_n$  varían ampliamente. En general, los valores menores de  $\Delta R_n$  corresponden a aquellas combinaciones en las que uno de los dos cebadores está a una concentración inferior a 300nM, especialmente cuando ésta corresponde al primer Forward. Las combinaciones que proporcionan mejores valores de  $\Delta R_n$  son aquellas en las que la concentración de ambos cebadores está entre 300 y 900nM. De todas ellas la combinación 900nM de primer Forward / 300nM de primer Reverse es la que ofrece una mayor intensidad de la señal de fluorescencia.

Todas las concentraciones ensayadas de sonda, excepto la menor (50nM), proporcionan valores de  $C_T$  similares, sin embargo, tal y como sucedía para los cebadores, existen diferencias significativas en la intensidad de la señal de fluorescencia ( $\Delta R_n$ ), correspondiendo los mayores niveles a la concentración más elevada (250nM) de sonda.



**Figura FME3.** Curvas de amplificación de las 9 combinaciones de concentración de los cebadores de la PCR. (a) Escala lineal. (b) Escala logarítmica.



**Figura FME4.** Curvas de amplificación de concentraciones de sonda entre 50-250nM. (a) Escala lineal. (b) Escala logarítmica.



## 2.2.7. Diseño experimental

### 2.2.7.1 Componentes del kit

Para minimizar la manipulación, se utiliza una master mix específica (*TaqMan Universal PCR Master Mix*) que contiene todos los componentes de la mezcla de reacción a excepción de la sonda y los cebadores. Su composición es la siguiente:

- *AmpliTaq Gold* DNA Polimerasa
- *Amp Erase UNG*
- dNTPs con dUTP
- Referencia pasiva
- Tampón de *PCR*

La *AmpliTaq Gold Polimerasa* es una enzima termoestable que presenta actividad exonucleásica en el sentido 5' 3' pero no en 3' 5' (INNIS *et al.* 1988; HOLLAND *et al.* 1991). Esta polimerasa se activa únicamente cuando se alcanza la temperatura de 94°C y permite el empleo de la estrategia de *Hot Start PCR*, que puede resultar muy útil cuando se prevé que el tiempo transcurrido entre la preparación de las reacciones y su depósito en el termociclador o secuenciador va a ser más o menos largo. Esta técnica hace innecesario el uso de hielo durante la preparación de estas reacciones y minimiza las amplificaciones inespecíficas que pueden tener lugar durante este lapso temporal.

El kit incorpora además otra enzima, la *AmpErase UNG* que ayuda a prevenir la contaminación por *carry over* durante la fase de preparación de las reacciones de amplificación. Debido a la elevada sensibilidad de la *PCR*, los amplicones generados en anteriores reacciones pueden servir como sustrato de nuevas amplificaciones, resultando en la contaminación de nuevas muestras. Junto con esta encima, se suministra una mezcla de dNTPs que presenta en lugar de dTTP su análogo dUTP. Durante la amplificación este nucleótido sustituirá al dTTP en las nuevas cadenas en formación, de manera que el producto final presentará dUTP en lugar de dTTP en estos lugares. La *AmpErase UNG* es una Uracil-N-glicosilasa que rompe las cadenas de DNA en aquellas posiciones en las que existe un Uracilo. El tratamiento con esta encima evita la contaminación por *carry over* ya que degrada específicamente aquellos productos

resultantes de reacciones anteriores pero no el DNA *template*, puesto que sólo los primeros contienen dUTP.

### 2.2.7.2. Preparación de las reacciones de amplificación

Las reacciones de amplificación se preparan teniendo en cuenta todas las precauciones relativas a la contaminación para muestras de DNA antiguo (apartado 9.2.2.1. del presente capítulo). Cada muestra fue amplificada por duplicado en la misma placa.

Se fabrican dos mezclas de reacción independientes para las muestras y la sonda en dos tubos eppendorf de 1.5ml autoclavados. La composición de la mezcla de reacción es la siguiente:

**Tabla ME8.** Composición de la mezcla de reacción de *Real Time PCR*.

	Concentración stock	Concentración final	Volumen/ reacción muestras	Volumen/ reacción patrones
<i>TaqMan Universal PCR Master Mix</i>	2x	1x	12.25 µl	12.25 µl
<i>Primer Forward</i>	20 µM	900 nM	1.125 µl	1.125 µl
<i>Primer Reverse</i>	20 µM	300 nM	0.375 µl	0.375 µl
Sonda	5 µM	250 nM	1.25 µl	1.25 µl
DNA			5 µl	10 µl
Agua			5 µl	-
Total			25 µl	25 µl

### 2.2.7.3. Carga y cerrado de la placa

Para este tipo de ensayo se utilizaron placas de 96 pocillos (*96-Well Reaction Plate with Barcode*) tapadas con una cubierta óptica adhesiva. El proceso de carga y cerrado de la placa se realiza de la siguiente forma:

- Se coloca la placa en un soporte de 96 pocillos (una caja vacía de puntas de pipeta de 200µl puede servir).
- Una vez preparadas las mezclas de amplificación sin DNA, se reparten en los diferentes pocillos de la placa a razón de 15 µl para las patrones y 20 µl para las muestras, cargando primero los pocillos de la recta patrón.

- Se añade en cada pozo el DNA de los patrones (10µl) comenzando por los de menor orden de dilución .
- Se añade entonces el DNA de las muestras que se desean cuantificar (5 µl).
- Se tapa la placa con la cubierta óptica adhesiva, para ello se dobla por los extremos y se levanta cuidadosamente el papel que la cubre. Se coloca entonces la cubierta encima de la placa sujetándola por los extremos y se ajusta presionando sobre la superficie con la ayuda de un aplicador suministrado con la placa. Presionando con el aplicador, se cortan los extremos que sobresalen de la cubierta por las líneas de puntos.
- Una vez tapada la placa se coloca encima la almohadilla de compresión, también suministrada con la placa.
- Se coloca la placa en el bloque térmico del sistema de detección correspondiente.

#### 2.2.7.4. Preparación del Sistema de Detección

Para la detección se utilizó en régimen de usuario un Sistema de Detección de Secuencia ABI Prism 7700 de la Unitat de Seqüenciació de los antiguos Serveis Científico-Tècnics de la Universitat de Barcelona, actualmente Unitat de Genòmica del Parc Científic de la misma universidad. El instrumento debe encenderse al menos dos horas antes de su uso. El proceso empleado es el siguiente:

- Se inicia el programa SDS (MAC). Se selecciona “New Plate” y se definen las siguientes opciones:

Plate Type: Single Reporter.

Plate Format: Standard Plate.

Run: Real Time, O.K.

- Se revisan las condiciones del programa de *PCR (Thermal Cycle Conditions)*.
- Se introduce el volumen de reacción: 25 µl.
- Se revisa la recolección de datos (*Data Collection*).
- Se introduce la placa en el bloque térmico del aparato, colocando encima la almohadilla de compresión.
- Se cierra apretando la cubierta térmica.
- En el programa se pulsa “Show Analysis”. Cuando el estado del aparato (Instrument status) marque “idle” y la temperatura de la tapa (cover temperature) alcance 104°C puede iniciarse la amplificación marcando “Run”. El programa de amplificación dura aproximadamente dos horas.

#### 2.2.7.5. Obtención de los resultados en bruto

Una vez el programa ha finalizado, los resultados brutos se guardan en el disco duro del ordenador.

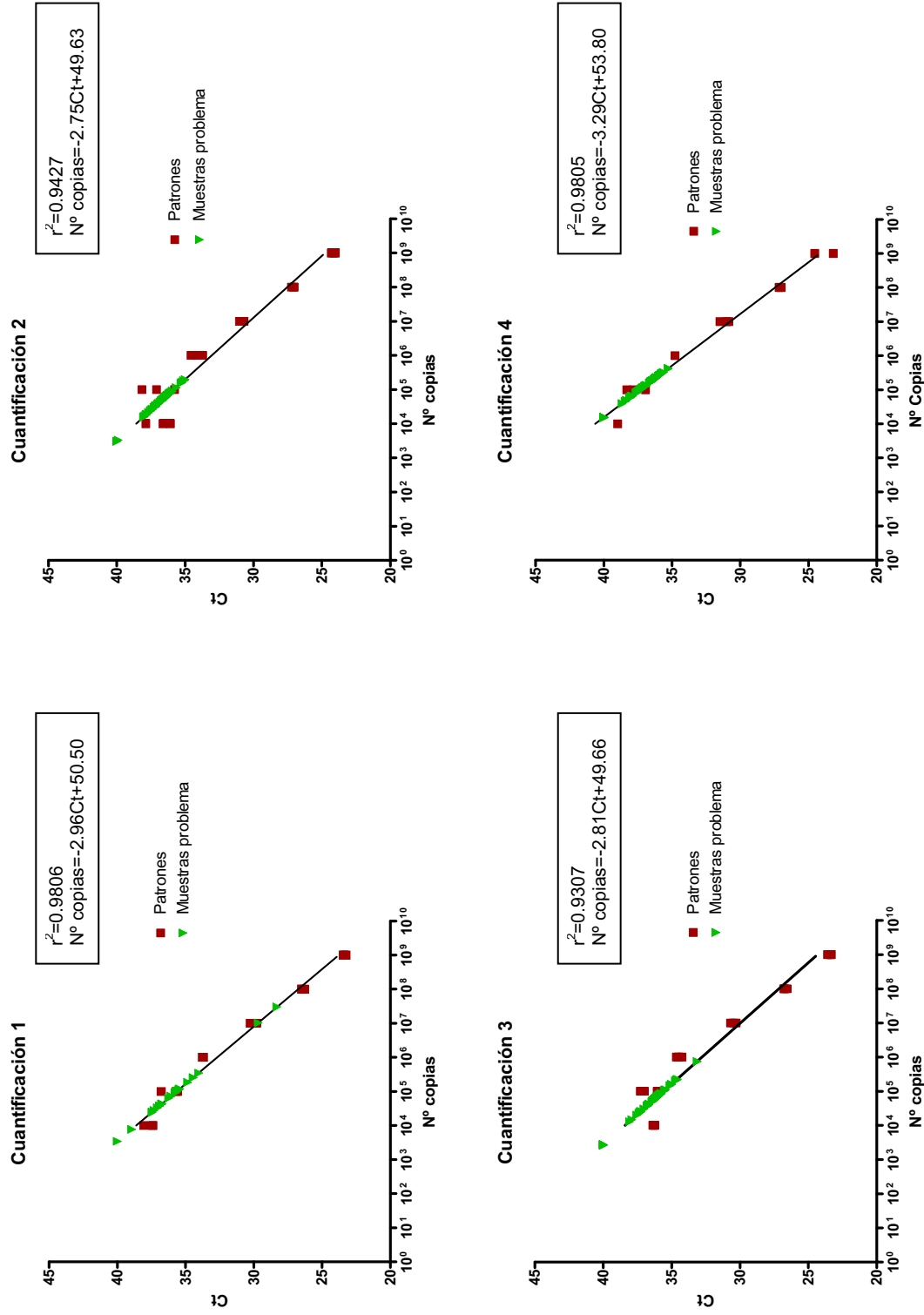
- Se abre el programa SDS.
- Se revisan las condiciones del programa de amplificación y de la recogida de datos.
- Se selecciona la placa y se selecciona la opción “Analyze” del menú de análisis. El programa analiza los datos recogidos por el sistema de detección y elabora la curva de amplificación.
- Se visualizan los datos en escala linear haciendo doble click sobre el eje vertical y cambiando la escala a linear.
- Se ajusta la línea base si es necesario. El área de la curva comprendida dentro de la línea base (por defecto entre los ciclos 3-15) debe ser igual a 0. Si no es así puede estrecharse el rango de ciclos modificando el valor final. Una buena aproximación consiste en determinar en qué ciclo comienza la amplificación, restarle dos unidades e introducir este valor como valor final.
- Se vuelve a la visión logarítmica haciendo doble click sobre el eje vertical y cambiando la escala de linear a logarítmica.
- Se define el umbral de detección (*Threshold*). Éste debe posicionarse aproximadamente en la mitad del tercio inferior de la parte visible de la fase exponencial de la amplificación.
- Una vez realizados estos cambios, se guarda el fichero mediante la opción “Save” del menú “File”. El programa calcula una recta de regresión entre el número de copias de cada muestra patrón y el valor de Ct obtenido para cada una de ellas. Una vez obtenida la recta de regresión se intrapolan los valores de Ct obtenidos para cada una de las muestras, obteniéndose el número de copias correspondiente.
- Para exportar los resultados al programa Excel se selecciona la opción “Export results” del menú “File”

#### 2.2.7.6. Edición y análisis de los resultados

El fichero Excel generado por el programa SDS contiene los valores de  $\Delta R_n$  y de Ct obtenidos para cada reacción (muestras y patrones) así como los valores de cuantificación de las diferentes muestras (expresados en número de copias) resultantes de la intrapolación de sus valores de Ct en la recta patrón.

Si la relación entre el número de copias de DNA de los patrones y los valores de Ct se ajusta a una recta de regresión de primer orden con un buen coeficiente de regresión (superior a 0.9), el número de copias estimado por el programa SDS en los extractos de DNA antiguo puede considerarse correcto. En la **Figura FME5** se representan las rectas patrón empleadas para calcular el número de copias en los extractos de DNA antiguo.

**Figura FME5.** Rectas patrón empleadas para estimar el número de copias de los extractos de DNA antiguo. Cada recta corresponde a una reacción de amplificación diferente. Se detallan también la ecuación de la recta y el coeficiente de regresión de cada una de las rectas patrón.



## **2.3. Detección de inhibidores de la PCR**

### **2.3.1. Inhibición de PCR con DNA fresco**

#### *2.3.1.1. Muestras estudiadas*

Mediante la detección de la presencia de inhibidores de la PCR en algunos de los extractos de DNA antiguo analizados se pretendía relacionar la capacidad inhibidora con ciertas variables tales como la eficiencia de amplificación, la procedencia del extracto o su antigüedad.

Con este objetivo se seleccionaron 112 extractos pertenecientes a 99 individuos representativos de los diferentes yacimientos de los que existía un volumen remanente y con los que se había intentado la amplificación de secuencias de mtDNA independientemente de su resultado. En la **Tabla ME9** se detallan los individuos y extractos analizados de cada yacimiento.

**Tabla ME9.** Muestras e individuos en los que se evaluó la presencia de inhibidores mediante la inhibición de una PCR con “DNA fresco”.

Procedencia	Individuo	Muestra	Procedencia	Individuo	Muestra	Procedencia	Individuo	Muestra
Abauntz	AB29F	1AB8	Mari	7	1MA7			
Abauntz	AB23C	AB9	Mari	8	1MA8			
Abauntz	AB21E	AB11	Mari	10	2MA10	Tell Halula	H-57	1H21
Abauntz	AB21D	AB13	Mari	11	3MA11	Tell Halula	H-64	2H13
Abauntz	AB17C	AB14	Mari	12	1MA12	Tell Halula	H-65	2H10
Abauntz	AB35F	AB15	<b>Total</b>	<b>11</b>	<b>12</b>	Tell Halula	H-66	1H9
Abauntz	AB35E	AB17	Muge	TAS 168	TA1	Tell Halula	H-66	2H9
Abauntz	AB25E	AB20	Muge	TAS 214	TA2	Tell Halula	H-68	H12
<b>Total</b>	<b>8</b>	<b>8</b>	Muge	TAS 216	TA9	Tell Halula	H-70	2H11
Açikli	8a AH89	16SK2	Muge	TAS 216	TA3	<b>Total</b>	<b>26</b>	<b>31</b>
Açikli	8b AH89	13SK3	Muge	TAS168	TA7	Tell Ramad	R65-1	TR7
Açikli	A95 2c-2	SK107-3	Muge	TAS168	TA8	Tell Ramad	R65-C8-SE A	TR13
<b>Total</b>	<b>3</b>	<b>3</b>	<b>Total</b>	<b>4</b>	<b>6</b>	Tell Ramad	R65-4II	TR9
Amir Abdallah	AA397	AA1	Nerja	NE-1829	1NE	Tell Ramad	R65(8)	TR4
Amir Abdallah	AA418	AA12	Nerja	NE-NM82.2	2NE	Tell Ramad	R65-7I	TR1
<b>Total</b>	<b>2</b>	<b>2</b>	Nerja	NE-NT79	3NE	Tell Ramad	R69(2)	TR14
Atxuri	AX-12	AT24	Nerja	NE-NAP	4NE	Tell Ramad	R69(2)	TR15
Atxuri	AX-14	AT26	Nerja	NE-NM82.1	5NE	<b>Total</b>	<b>6</b>	<b>7</b>
Atxuri	AX-15	AT27	<b>Total</b>	<b>4</b>	<b>5</b>	Toledo	K-13	TO1
<b>Total</b>	<b>3</b>	<b>3</b>	Sant Pau	2 7608	SP2	<b>Total</b>	<b>1</b>	<b>1</b>
Caldeirão	O11 77b	CA1	Sant Pau	6 7214	1SP6	Tres Montes	TM-2	2TM1
Caldeirão	O10 107	CA7	Sant Pau	7 7720	1SP7	Tres Montes	TM-3	1TM2
Caldeirão	Q12 207	CA8	Sant Pau	7 7720	2SP7	Tres Montes	TM-6	1TM4
<b>Total</b>	<b>3</b>	<b>3</b>	<b>Total</b>	<b>3</b>	<b>4</b>	Tres Montes	TM-6	2TM4
Dja'de	SK-A2	DJ1	Tell Halula	H-3	TH1	Tres Montes	TM-8	1TM5
Dja'de	SK-D4 16	DJ3	Tell Halula	H-3	TH2	Tres Montes	TM-11	1TM6
Dja'de	SK-N4	2DJ4	Tell Halula	H-4	TH3	Tres Montes	TM-14	TM7
Dja'de	SK-R 3	1DJ6	Tell Halula	H-4	TH4	Tres Montes	TM-15	TM8
Dja'de	SK-R 3	2DJ6	Tell Halula	H-7	TH5	Tres Montes	TM-17	TM10
<b>Total</b>	<b>4</b>	<b>5</b>	Tell Halula	H-8	TH7	Tres Montes	TM-18	TM11
El Pirulejo	2PI	2PI	Tell Halula	H-8	TH8	Tres Montes	TM-19	TM12
<b>Total</b>	<b>1</b>	<b>1</b>	Tell Halula	H-24	1H24	Tres Montes	TM-20	TM13
Garai	G3	GA3	Tell Halula	H-25	2H25	<b>Total</b>	<b>11</b>	<b>12</b>
Garai	G11	GA11	Tell Halula	H-26	1H26	Zafarraya	ZA13	ZA13
Garai	G15	GA15	Tell Halula	H-28	1H28	Zafarraya	ZA16	ZA16
Garai	G16	GA16	Tell Halula	H-30	1H30	<b>Total</b>	<b>2</b>	<b>2</b>
Garai	G23	GA23	Tell Halula	H-35	1H1	<b>Total general</b>	<b>99</b>	<b>112</b>
Garai	G24	GA24	Tell Halula	H-36	1H2			
Garai	G50	GA50	Tell Halula	H-37	2H20			
<b>Total</b>	<b>7</b>	<b>7</b>	Tell Halula	H-43	2H31			
Mari	1	1MA1	Tell Halula	H-44	1H32			
Mari	2	2MA2	Tell Halula	H-47	1H18			
Mari	3	1MA3	Tell Halula	H-48	1H33			
Mari	4	1MA4	Tell Halula	H-49	1H15			
Mari	4	2MA4	Tell Halula	H-53	1H14			
Mari	5	1MA5	Tell Halula	H-54	2H34			
Mari	6	1MA6	Tell Halula	H-56	1H16			



2.3.1.2. *Protocolo experimental*

Se prepararon reacciones de *PCR* de 25µl de volumen final con una cantidad fija de DNA obtenido de individuos vivos (“DNA fresco”) a las que se añadieron tres volúmenes diferentes de cada uno de los extractos de DNA antiguo. En la **Tabla ME10** se detallan las condiciones de la reacción de amplificación.

**Tabla ME10.** Condiciones de las reacciones de amplificación empleadas para detectar la presencia de moléculas inhibitoras en los extractos de DNA antiguo. Reactivos de Biotools®.

Reactivos	Concentración stock	Concentración en reacción	Volumen añadido
Tampón de PCR	10x	1x	2.5µl
MgCl <sub>2</sub>	50mM	2mM	1µl
dNTPs	10mM	0.2mM	0.5µl
cebadores	100pmol/µl	5pmol/µl	0.1µl
<i>Taq</i> polimerasa	1u/µl	0.05u/µl	1.25µl
DNA “fresco”	100ng/µl	-	1µl
DNA antiguo	-	-	5, 2.5 ó 1µl
Agua	-	-	13.55, 16.05 ó 17.55µl
Total			25µl

En la amplificación se emplearon los cebadores externos del primer segmento de la HVRI del mtDNA analizado en la presente tesis doctoral (posiciones 16112-16262). Los resultados se visualizaron en geles de agarosa al 2%, interpretándose la ausencia o disminución de la banda de amplificación como signo de la presencia de inhibición para dicha cantidad y dicho extracto de DNA antiguo.

Para preparar las reacciones de amplificación no se siguieron las medidas de esterilidad propuestas para las muestras antiguas. Todas las reacciones fueron preparadas en la mesa de laboratorio con material autoclavado pero no radiado y con guantes.

### **2.3.2. Espectros de absorción de luz**

Para determinar si el patrón espectrofotométrico normal de un DNA fresco se ve alterado por la presencia de inhibidores se seleccionaron tres muestras procedentes de diferentes yacimientos y con diferente capacidad inhibidora: TA3 (Muge); ZA16 (Zafarraya) y GA3 (Garai) y se determinó su patrón espectrofotométrico. Se diluyó una alícuota de cada uno de los extractos para obtener un volumen final de 100µl y se midió su espectro de absorción de luz para una longitud de onda comprendida entre 200 y 400nm en un espectrofotómetro Beckman de doble haz.

### 3. LIMPIEZA SUPERFICIAL DE LAS MUESTRAS

La limpieza superficial permite eliminar la mayor parte de la suciedad procedente del enterramiento que suele acompañar a la pieza dentaria u ósea. Así mismo, elimina posibles moléculas de DNA contaminante procedentes de los arqueólogos encargados de la excavación y/o de los antropólogos responsables del estudio morfológico.

#### 3.1. Limpieza con arenadora

La limpieza se realiza con una arenadora o “sandblaster” Dentalfarm modelo Base-1 Plus de C.I.E. (**Figura FME6**) que inyecta un material abrasivo a presión sobre la superficie de la pieza provocando la eliminación de la capa más externa. La muestra se introduce por una abertura lateral de la cámara y se manipula introduciendo las manos enfundadas en unos guantes de goma a través de dos orificios situados en la parte frontal. Un compresor de aire incorporado al aparato (modelo M-220-S) inyecta a presión al interior de la cabina un material abrasivo que puede variar según el grado de abrasión deseado (**Figura FME7**). En nuestro caso el abrasivo utilizado fue óxido de aluminio del tipo Al 203-150. Éste se coloca en un depósito lateral que contiene una resistencia eléctrica dónde debe calentarse y deshidratarse a 50°C durante como mínimo 3 horas antes de su uso.

El polvo eliminado y el material abrasivo son adsorbidos desde la cabina de trabajo hasta un depósito de agua situado en la parte posterior del aparato, donde se inmovilizan en forma de sedimento. El fondo de la cabina de trabajo se rellena con 6-7 kg de corindón, que actúa inmovilizando los detritos producidos en la cámara durante el proceso de arenación y que puede extraerse accionando un conmutador situado en el interior de la cabina de trabajo. Este material es aún más abrasivo que el óxido de aluminio, y puede usarse en casos determinados, por ejemplo si se desea eliminar un depósito de cálculo dental o incrustaciones de carbonatos de la superficie del diente.

La limpieza se produce por abrasión directa del óxido de aluminio, que impacta en la pared del diente o hueso y elimina la capa más superficial. Conviene que la abrasión sea intensa, pero no demasiado, ya que podría perforar la superficie

permitiendo el acceso a la cámara radicular, en el caso del diente, o al interior del hueso, aumentando el riesgo de contaminación con DNA exógeno.

El proceso es el siguiente:

- Se rellena el depósito con el material abrasivo hasta la medida recomendada y se conecta a la corriente aproximadamente tres horas antes de su utilización para su deshidratación.
- Se abre la válvula del tubo de alimentación neumática de la arenadora y se pone en marcha el compresor de aire.
- Se activa el interruptor de iluminación de la cabina de trabajo situado en la parte superior de la cabina.
- Se activa el interruptor de puesta en marcha del sistema de evacuación (y destrucción) del polvo generado durante el proceso de limpieza situado también en la parte superior del aparato.
- Por la abertura lateral se introduce el tubo falcon que contiene la muestra en el interior de la cabina de trabajo. Se introducen entonces las manos en los guantes acoplados a la cabina, se saca la muestra del falcon y se deposita éste en el exterior a través de la abertura lateral. Una de las manos sujetará la muestra con cuidado mientras que la otra sostendrá el proyector de abrasivo.
- Para limpiar la muestra se aprieta con el pie el pedal de puesta en marcha del sistema de arenación de manera que el aire del compresor llegue al proyector mezclado con el abrasivo. Dirigiendo el proyector sobre la superficie del diente o hueso se erosiona su superficie en mayor o menor grado dependiendo de su suciedad.
- Se guarda la muestra en el interior de un nuevo falcon estéril que se introduce por la abertura lateral.

### **3.2. Irradiación con luz ultravioleta**

Una vez realizada la limpieza superficial con la arenadora conviene exponer la muestra a luz ultravioleta (254nm) durante 15 minutos por cada cara. Se abren para ello los tubos falcon en el interior de la campana de seguridad biológica, se depositan las muestras dentro de los correspondientes tapones sin tocarlas y se conecta la luz

ultravioleta situada en el interior de la campana. Pasados 15 minutos se voltean las muestras con ayuda de unas pinzas previamente lavadas con lejía e irradiadas con UV.



**Figuras FME6 y FME7.** Aparataje empleado para la limpieza superficial de las muestras. Arenadora modelo Base 1 Plus (izquierda) y compresor de aire (derecha).



**Figura FME8.** Molino refrigerado con nitrógeno líquido (Spex 6700) empleado para la tritución de las muestras de la presente tesis doctoral.

## **4. PULVERIZACIÓN DE LAS MUESTRAS**

### **4.1. Limpieza del material**

Los viales de plástico así como la barra percutora y las piezas de los extremos de acero inoxidable se lavan de la siguiente manera:

- Un primer lavado con detergente, preferiblemente aniónico, tras el cual se aclaran con agua del grifo.
- El percutor y las piezas de los extremos de acero inoxidable se limpian entonces con un trozo de papel de filtro impregnado primero en etanol 70% y luego en lejía 70%. Tras cada limpieza conviene secar bien las piezas para que no queden residuos de ninguno de los dos compuestos. Tras la última limpieza con lejía se aclaran con agua del grifo.
- Se enjuagan abundantemente todas las piezas con agua mili-Q y se secan con papel de filtro.
- Se enjuagan abundantemente todas las piezas con agua estéril para inyectables (Braun).
- Tras la limpieza se autoclavan las piezas y se depositan en una cámara con luz UV, donde permanecen durante al menos 3 horas.

### **4.2. Trituración**

Una vez limpiadas las muestras, se pulverizan mediante un molino de impactación electromagnética modelo “Freezer Mill, Spex 6700” refrigerado con nitrógeno líquido (**Figura FME8**). El nitrógeno líquido congela la muestra facilitando así su pulverización y evita la posible degradación del material genético del hueso por el exceso de calor generado con la trituración mecánica.

Las muestras se depositan en el interior de unos viales especiales que se introducen en el molino. Cada uno de los viales consta de un tubo transparente de plástico de unos 10 cm de largo con un diámetro aproximado de 2 cm, dos extremos metálicos que se encajan a ambos lados del vial y un percutor también metálico que se coloca en el interior del mismo junto con la muestra. Al poner en marcha el aparato se crea una corriente magnética

entre los dos extremos metálicos del vial de manera que el percutor se mueve impactando contra los mismos y pulverizando la muestra situada en el interior. La velocidad y el tiempo de impactación pueden regularse mediante unos mandos situados en el exterior del molino. Un minuto a máxima frecuencia suele ser suficiente para triturar completamente una muestra.

El nitrógeno líquido se recoge de los servicios centrales de la facultad y se almacena en un depósito exterior. De ahí pasa al interior del molino gracias a la presión generada por una bombona rellena de nitrógeno gaseoso que conecta con el depósito de nitrógeno líquido a través de un tubo de plástico.

La presión se regula a través de tres válvulas situadas una en la bombona de nitrógeno, otra a su salida y la tercera al final del tubo que conecta con el depósito de nitrógeno líquido. Dos manoreductores permiten ver la presión en el interior de la bombona y a su salida, en el tubo que conecta con el recipiente de nitrógeno líquido. La presión del nitrógeno en el interior del recipiente se compensa por medio de una pequeña válvula que se mantiene abierta en reposo, y se cierra cuando se insufla gas.

Durante el uso del molino y especialmente durante el manejo de nitrógeno líquido es preciso tomar diversas precauciones:

- Evitar el contacto del nitrógeno líquido con la piel ya que puede provocar quemaduras. Para la manipulación se recomienda usar guantes gruesos y protegerse los ojos y la cara con una máscara de plástico.
- Utilizar las pinzas metálicas suministradas con el molino para introducir y extraer los viales del interior del aparato.
- No poner en marcha el aparato sin la muestra o sin el impactador.
- No poner en marcha el aparato sin nitrógeno líquido, o con un nivel inferior al señalado.

El proceso a seguir es el siguiente:

- En el interior de la campana de seguridad biológica se ensamblan los viales depositando la muestra en su interior con ayuda de unas pinzas.

- Se comprueba que la válvula situada al final del tubo de plástico que conecta con el depósito de nitrógeno líquido esté en posición paralela al mismo (abierta).
- Se abre la válvula de la bombona que contiene el nitrógeno en estado gaseoso. Deberá observarse entonces un aumento de presión en el manoreductor de la izquierda.
- Se abre la válvula situada a la salida de la bombona de nitrógeno gaseoso. De esta manera la presión en el interior de la bombona se compensa y el nitrógeno es impulsado a través del tubo. Se llena el depósito del molino con nitrógeno líquido hasta la marca roja, que se encuentra a unos centímetros por encima del depósito donde se alojan los viales, procurando no sobrepasarla. Si el depósito se llena en exceso el nitrógeno líquido puede derramarse por encima del aparato, inmovilizando momentáneamente las bisagras de la tapa por congelación. En este caso conviene esperar hasta que el nitrógeno se haya evaporado, de lo contrario la tapa del molino puede llegar a resquebrajarse si se intenta cerrar en estas condiciones.
- Se colocan los viales en el interior del molino lleno de nitrógeno líquido con ayuda de una pinza metálica que se encaja en uno de los extremos metálicos del vial permitiendo su manipulación sin tocarlo. Resulta conveniente mantener el orden de las muestras a la hora de triturar, por lo tanto se coloca la primera muestra en el receptáculo cilíndrico y las otras dos en el depósito lateral para que se enfríen, quedando así preparadas para una trituración posterior.
- Se baja la tapa del molino con cuidado, de manera que los viales queden sumergidos en el nitrógeno líquido y se cierra.
- Se deja enfriar la muestra unos 5 minutos hasta que remita el vapor de nitrógeno líquido expulsado por la válvula de escape de la parte superior de la tapa del molino.
- Se ajusta el regulador de tiempo a 1 minuto y se va subiendo paulatinamente la frecuencia de impactos hasta llegar al máximo.
- Se abre la tapa del molino con ayuda de unos guantes y se extraen los viales con ayuda de la pinza metálica.
- Se limpia cuidadosamente la superficie de los viales con papel de filtro impregnado en lejía al 70% y se depositan en el interior de la campana de seguridad biológica procurando seguir el orden de trituración. Los viales se sujetan por uno de los extremos metálicos a un pequeño aparato de metal suministrado junto con el molino. Se sujeta entonces el extremo libre a la pinza metálica y se extrae girando un tornillo incorporado a la misma.
- Una vez abierto el vial se extrae el percutor metálico con ayuda de unas pinzas metálicas estériles y se decanta el polvo en el interior de un tubo falcon de plástico



estéril previamente rotulado. El polvo de hueso puede almacenarse en el congelador a –20°C.

## **5. LAVADO DEL POLVO DE HUESO**

El polvo de hueso obtenido tras la trituración se lava repetidas veces con EDTA 0.5M pH 8 estéril. El EDTA es un quelante iónico que actúa secuestrando las sales iónicas contenidas en la muestra, incluyendo el calcio de los huesos.

- Se añaden aproximadamente 10ml de EDTA 0.5M pH 8 al polvo de hueso contenido en tubos falcon estériles de 50ml, se tapa y se agita enérgicamente.
- Se centrifugan los tubos falcon 5 minutos a 2500 r.p.m. Tras la centrifugación el polvo queda adherido al fondo del tubo.
- Se decanta el sobrenadante, que contiene EDTA y sales de la muestra y se desecha.
- Se repite la operación tantas veces como sea necesario hasta que el sobrenadante se vuelva incoloro. De dos a cuatro lavados suelen ser suficientes para la mayoría de las muestras aunque algunas de ellas pueden necesitar lavados adicionales. Conviene llegar a un compromiso entre el número de lavados y la suciedad de la muestra, ya que en cada lavado se pierde cierta cantidad de DNA contenido en la hidroxiapatita, que pasa al sobrenadante. Una vez lavado el sedimento, se guarda en nevera a 4°C hasta llevar a cabo la siguiente fase.

## 6. DIGESTIÓN CON SOLUCIÓN DE LISIS

La “solución de lisis” actúa rompiendo las membranas de las células óseas (osteoclastos) y liberando el material genético contenido en el núcleo y las mitocondrias.

La solución de lisis se compone de EDTA, Tris-HCl, SDS y proteinasa K. El EDTA es un quelante iónico que se encarga de eliminar las sales que puedan haber quedado después del lavado y aquellas liberadas del interior de la célula tras la rotura de membranas. El Tris-HCl actúa como estabilizador de pH. El SDS es un detergente aniónico que actúa disolviendo los lípidos orgánicos y la proteinasa K es una enzima que hidroliza las proteínas.

Todas las soluciones empleadas en la fabricación de la solución de lisis a excepción del SDS son soluciones estériles prefabricadas. EL SDS se fabrica diluyendo la cantidad correspondiente de polvo en agua estéril (Braun). Una vez disuelta se filtra mediante el sistema “Steritop” (Millipore) y se almacena en botellas de vidrio estériles.

Puede prepararse un único stock de solución de lisis añadiendo a continuación la cantidad correspondiente a cada tubo falcon con la muestra, o bien añadir cada uno de los componentes de la solución por separado a cada muestra. La primera opción implica el manejo de volúmenes más grandes, con lo que es necesario el empleo de pipetas de vidrio que han de ser previamente esterilizadas. En el segundo caso se manejan volúmenes más pequeños, por lo que pueden usarse micropipetas de 1ml con puntas desechables “*DNAsa y RNAsa free*”, filtro para aerosoles y garantía de esterilidad de fábrica.

Las concentraciones de los stocks y la solución de lisis para una y para cinco muestras es la siguiente:

**Tabla ME11.** Concentraciones y volúmenes de los componentes de la solución de lisis.

	<b>Concentración stock</b>	<b>Concentración en reacción</b>	<b>Volumen (5 muestras)</b>	<b>Volumen (1 muestra)</b>
EDTA	50mM	5mM	5ml	1ml
Tris	100mM	10mM	5ml	1ml
SDS	5%	0.5%	5ml	1ml
Prot.K	20mg/ml	50µg/ml	125µl	25µl
H <sub>2</sub> O	-	-	34.875ml	7ml

- Después de añadir la solución de lisis a las muestras se tapan los tubos falcon, se agitan y se incuban de 12 a 24 horas a 37°C con agitación dentro de un digestor termostatzado.
- Se centrifugan 5 minutos a 2500 r.p.m.
- Se decanta el sobrenadante sobre nuevos tubos falcon y se guarda para la extracción.

## **7. EXTRACCIÓN DEL DNA**

La extracción del DNA se realiza mediante un protocolo Fenol/Cloroformo estándar con ciertas modificaciones. El fenol se mezcla con el sobrenadante resultante de la lisis (fase acuosa) y desnaturaliza las proteínas de la muestra, que pasan a la fase fenólica mientras que los ácidos nucleicos permanecen solubles en la fase acuosa. El cloroformo disuelve los lípidos y mejora la eficiencia de las extracciones debido a su capacidad para desnaturalizar proteínas. La última parte de la extracción se realiza sólo con cloroformo para eliminar posibles restos de fenol en la fase acuosa. La diferente densidad de las fases acuosa y orgánica (fenólica o clorofórmica) permite su separación por centrifugación de una manera fácil y rápida.

### **7.1. Preparación del material**

El material de vidrio utilizado para la extracción, desechable (pipetas Pasteur) y no desechable (pipetas de 5 y 10ml) se lava de la siguiente manera:

- Se realiza un primer lavado con detergente, preferiblemente aniónico y se aclara con agua del grifo.
- Se lava a continuación con etanol al 70%, se aclara con agua del grifo y se seca con papel de filtro limpio, procurando que no queden restos de etanol.
- Se lava entonces con lejía al 70% y se procede de la misma manera.
- Se enjuaga dos o tres veces con agua mili-Q.
- Se aclara dos o tres veces con agua estéril para inyectables (Braun) y se seca con papel de filtro limpio.
- Se envuelven las pipetas en papel de aluminio, individualmente las de 5 y 10 ml, y en grupos de 12 las Pasteur.
- Se autoclavan durante 15 minutos a 210°C y posteriormente se secan en una estufa a 60°C.
- Una vez autoclavadas, se depositan en el interior de una cámara con luz UV, dónde permanecen un mínimo de 3 horas.

En cuanto a los reactivos utilizados para la extracción, tanto el fenol como el cloroformo son soluciones preparadas de la casa comercial Iberlabo con garantía de esterilidad y ausencia de RNAsa y DNAsa.

Previamente a su utilización el fenol debe saturarse con una solución de Tris-base 0.02 M que se suministra con él. Esta solución se añade directamente a la botella de fenol en el interior de la campana de seguridad biológica y se deja reposar hasta que la mezcla sea total. Una vez preparado el fenol debe guardarse en nevera a -4° C, en una botella opaca, o bien envuelto en papel de aluminio para prevenir su oxidación.

## **7.2. Extracción de DNA**

La extracción del DNA de las muestras de hueso o diente se realiza en grupos de siete muestras y un control de extracción. Cada grupo se procesa habitualmente en un mismo día, extrayendo cuatro muestras por la mañana y cuatro por la tarde.

- Se limpia la campana de seguridad biológica con lejía al 70% y se dispone el material necesario en su interior: fenol y cloroformo, dos pipetas de vidrio de 10ml, 12 pipetas Pasteur, 4 peritas de goma, una gradilla con 4 tubos falcon estériles de 50ml, un dispensador de volúmenes para pipetas de vidrio, una mascarilla y dos pares de guantes limpios. Antes de introducir en la campana los tubos falcon, la gradilla y el dispensador de volúmenes se limpian superficialmente con papel de filtro impregnado en lejía al 70%.
- Se conecta el fluorescente de luz UV de la campana y se deja expuesto el material durante 30 minutos.
- Se añade a cada muestra (10ml) un volumen de fenol con pipetas de vidrio y se agita enérgicamente para que se mezcle.
- Se centrifuga 2 minutos a 3000 r.p.m. Tras la centrifugación se forman dos fases debido a la diferencia de densidad entre el fenol y la muestra (fase acuosa).
- Se extraen los tubos falcon de la centrífuga y se limpian superficialmente con papel de filtro impregnado en lejía al 70%. Se depositan en el interior de la campana de seguridad biológica.
- Se recupera la fase inferior, que es la que contiene el DNA, con pipetas Pasteur de vidrio y se pasa a unos nuevos tubos falcon de 50ml.

- Se añade a cada muestra (aproximadamente 10ml) medio volumen de fenol (5ml) y medio volumen de cloroformo (5ml) con pipetas de vidrio y se agita.
- Se centrifuga 2 minutos a 3000 r.p.m. Tras la centrifugación se forman de nuevo dos fases debido a la diferencia de densidades.
- Se extraen los tubos falcon de la centrifuga, se limpian de la manera descrita y se depositan en el interior de la campana de seguridad biológica.
- La fase acuosa que contiene el DNA en este caso es la superior ya que la mezcla de fenol y cloroformo es más densa que el agua. Se recupera pues la fase superior con ayuda de una pipeta Pasteur.
- Se añade un volumen de cloroformo (aproximadamente 5ml) a cada tubo Falcon con una pipeta de vidrio y se agita.
- Se centrifuga 2 minutos a 3000 r.p.m. Tras la centrifugación se forman de nuevo dos fases.
- Igual que en el paso anterior se recupera la fase superior (acuosa) con pipetas Pasteur de vidrio y se pasa a unos nuevos tubos falcon. Debe procurarse extraer la mayor parte de la fase acuosa pero sin absorber la fase que contiene el cloroformo, ya que éste podría disolver el filtro del Centriplus obturándolo e impidiendo que pudiera llevarse a cabo la posterior concentración de la muestra.

## **8. CONCENTRACIÓN DEL DNA**

La concentración es una técnica de microseparación que permite eliminar las sales del extracto de DNA remanentes de los lavados con EDTA, de la digestión enzimática o de la propia muestra. Estas sales dificultan la disolución del DNA y pueden llegar a inhibir las reacciones enzimáticas posteriores. La concentración permite a su vez reducir el volumen de la muestra aumentando la concentración del DNA.

La concentración se lleva a cabo mediante microconcentradores Centriplus 30.000 (Amicon). Cada microcentrador consta de un compartimento en el que se deposita el extracto de DNA y que presenta una membrana de filtración. En la parte inferior del mismo puede acoplarse un depósito que recoge el producto resultante de la filtración, que en este caso es desechado (SAMBROOK *et al.* 1989). Para recoger el volumen no filtrado existe un tercer depósito que puede acoplarse a la parte superior del compartimento central.

En este dispositivo la concentración se produce por ultrafiltración de la solución a través de una membrana anisotrópica de tipo YM hidrofílica y de baja absorción con un tamaño de poro de 30000 Daltons. La fuerza centrífuga actúa sobre la superficie de la membrana en un ángulo fijo de manera que los solutos de bajo peso molecular contenidos en el extracto pueden traspasarla mientras que aquellos de diámetro superior al del poro quedan retenidos. Los solutos que van siendo retenidos durante el proceso se depositan en los márgenes de la membrana de manera que no obturan el filtro. Por lo tanto el poder de filtración de éste se mantiene teóricamente durante todo el proceso de filtrado, o disminuye ligeramente.

Después de la filtración el volumen aproximado que queda sin filtrar en la superficie de la membrana es de 25 a 50  $\mu\text{l}$ . Para recogerlo sólo hay que acoplar el depósito de recogida de la muestra a la parte superior del Centriplus que contiene la membrana, invertirlo y centrifugar brevemente. Se ha calculado que diluyendo la muestra 10 veces y concentrándola a continuación mediante este sistema se consiguen eliminar el 99% de las sales contenidas en la muestra.



El proceso detallado es el que sigue:

- Se disponen en el interior de la campana dos alícuotas de agua estéril en tubos falcon de plástico de 50ml, cuatro Centriplus, un tubo falcon de plástico de 50ml sin abrir, una mascarilla y dos pares de guantes limpios. Los Centriplus y los tubos falcon se limpian superficialmente con papel de filtro impregnado en lejía al 70% antes de ser introducidos en la campana.
- Se conecta el fluorescente de luz UV del interior de la campana y se irradia el material durante 20 minutos.
- Se decanta cada uno de los extractos de DNA sobre el depósito superior del Centriplus y se cierra.
- Se centrifuga a 4000 r.p.m. hasta que la totalidad de la muestra haya atravesado el filtro. 30 minutos suelen ser suficientes.
- Se extraen los Centriplus de la centrífuga, se limpian cuidadosamente con papel de filtro impregnado en lejía al 70% y se introducen en el interior de la campana de seguridad biológica.
- Se desecha el volumen contenido en el depósito inferior del Centriplus desmontándolo y decantando su contenido. Se añaden entonces 15ml de agua estéril en el depósito superior de cada Centriplus.
- Se centrifuga a 4000 r.p.m. hasta que se haya filtrado prácticamente todo el agua y tan sólo quede una pequeña cantidad (entre 50 y 100  $\mu$ l) mojando el filtro. De 15 a 20 minutos suelen ser suficientes, aunque algunas muestras pueden necesitar tiempos de centrifugación mayores. Conviene pues ajustar la centrífuga a un tiempo bajo (por ejemplo 15 minutos) y centrifugar a tiempos cortos las veces que sea necesario si queda demasiado volumen mojando el filtro.
- Durante este tiempo se disponen en el interior de la campana de seguridad biológica tres micropipetas Gilson de 200  $\mu$ l-1ml, 20-200  $\mu$ l y 2-20  $\mu$ l, una caja de puntas estériles de pipeta de 2-200  $\mu$ l con filtro para aerosoles, 5 puntas de pipeta estériles con filtro para aerosoles de 200  $\mu$ l-1ml y dos pares de guantes limpios. Sobre una gradilla para microtubos se disponen 8 tubos de 1.5ml con el tapón abierto.
- Se conecta el fluorescente de luz UV del interior de la campana y se irradia el material mientras dura la centrifugación.
- Una vez centrifugados se limpian los Centriplus de la forma descrita y se introducen en el interior de la campana.

- Se desecha el volumen contenido en el depósito inferior del Centriplus desmontándolo y decantando su contenido. Se acopla a la cámara superior que contiene la membrana el depósito de recogida de la muestra y se invierte el Centriplus.
- Se centrifuga del Centriplus invertido de esta manera 5 minutos a 2500 r.p.m.
- Se recupera el volumen contenido en el depósito de recogida y se reparte equitativamente en dos microtubos eppendorf de 1'5ml con las micropipetas. En la superficie de la tapa se apunta el volumen recuperado.
- Se almacena el DNA disuelto en el congelador a -20° C. Una de las alícuotas se utiliza en las reacciones de amplificación mientras que la otra se guarda como reserva.

## 9. AMPLIFICACIÓN DE SECUENCIAS DE LA REGIÓN HIPERVARIABLE I DEL DNA MITOCONDRIAL

### 9.1. Muestras estudiadas

Se intentó la amplificación de secuencias específicas de la región Hipervariable I del DNA mitocondrial en 169 de los 183 extractos de DNA antiguo obtenidos.

Con estos extractos se llevaron a cabo un total de 1618 intentos de amplificación en 248 grupos distintos de amplificación empleando dos diseños experimentales diferentes (para más detalles ver apartado 9.2 de este capítulo). Con el primero de ellos (GARCÍA-BOUR *et al.* 2004) se realizaron un total de 325 intentos de amplificación, de los cuales 113 correspondían al segmento completo de 278 pb y 212 a un segmento parcial de 180 pb. Con el segundo modelo de amplificación, diseñado específicamente para la presente tesis doctoral, se produjeron 1274 intentos de amplificación, 730 del primer segmento (133 pb legibles), 556 del segundo (112 pb legibles) y 7 del segmento completo (246 pb legibles).

Del total de las muestras sometidas a una primera amplificación, 147 fueron reamplificadas y 116 purificadas en gel y sometidas a reacción de secuenciación. El número de intentos de cada reacción por extracto de DNA se resume en la **Tabla ME12**.

**Tabla ME12.** Número de reacciones de amplificación, reamplificación y secuenciación llevadas a cabo en cada uno de los extractos de DNA y en cada yacimiento arqueológico (en negrita).

<b>Procedencia</b>	<b>Individuo</b>	<b>Muestra</b>	<b>Amplificación</b>	<b>Reamplificación</b>	<b>Secuenciación</b>
Abauntz	AB27	AB3	1	0	0
Abauntz	ABb1.1	AB5	7	3	8
Abauntz	ABb1.2	AB6	1	0	0
Abauntz	AB29F	1AB8	7	1	0
Abauntz	AB23C	AB9	7	3	7
Abauntz	AB21E	AB11	10	1	0
Abauntz	AB21D	AB13	20	3	5
Abauntz	AB17C	AB14	12	2	9
Abauntz	AB35F	AB15	11	2	7
Abauntz	AB35E	AB17	12	3	6
Abauntz	AB25E	AB20	20	2	6
Abauntz	AB25D	AB21	1	0	0
<b>Total Abauntz</b>			<b>109</b>	<b>20</b>	<b>48</b>
Açikli	12 AH89	19SK5	14	8	8
Açikli	7 AH89	10SK4	7	0	0
Açikli	8a AH89	16SK2	19	15	12
Açikli	8b AH89	13SK3	6	7	5
Açikli	AH95 2c-1	SK106	5	4	8
Açikli	AH95 2c-2	SK107-3	22	18	9
Açikli	AH95 2c-2	SK107-4	6	5	2
Açikli	AH95 2c-3	SK109	7	2	3
<b>Total Açikli</b>			<b>86</b>	<b>59</b>	<b>47</b>
Amir Abdallah	AA370	AA2	1	0	0
Amir Abdallah	AA393	AA3	1	0	0
Amir Abdallah	AA269	AA4	7	0	0
Amir Abdallah	AA303	AA5	1	0	0
<b>Total Amir Abdallah</b>			<b>10</b>	<b>0</b>	<b>0</b>
Atxuri	AX-3	AT15	4	5	6
Atxuri	AX-4	AT16	7	5	2
Atxuri	AX-5	AT17	6	5	2
Atxuri	AX-6	AT18	6	5	2
Atxuri	AX-7	AT19	5	5	2
Atxuri	AX-8	AT20	12	12	4
Atxuri	AX-9	AT21	6	5	0
Atxuri	AX-10	AT22	2	1	0
Atxuri	AX-11	AT23	4	3	2
Atxuri	AX-12	AT24	6	7	0
Atxuri	AX-13	AT25	11	9	3
Atxuri	AX-14	AT26	4	4	0
Atxuri	AX-15	AT27	11	12	8
<b>Total Atxuri</b>			<b>84</b>	<b>78</b>	<b>31</b>
Caldeirão	O11 77b	CA1	34	33	2
Caldeirão	R12 sc 71	CA3	17	18	8
Caldeirão	O12 194	CA4	36	47	66
Caldeirão	O10 107	CA7	10	4	7
Caldeirão	O12 207	CA8	8	2	0
<b>Total Caldeirão</b>			<b>105</b>	<b>104</b>	<b>83</b>
Dja'de	SK-A2	DJ1	7	4	8
Dja'de	SK-D4 16	DJ3	9	3	11
Dja'de	SK-N4	1DJ4	6	3	9
Dja'de	SK-R3	1DJ6	7	3	3

Tabla ME12. (Continuación).

Procedencia	Individuo	Muestra	Amplificación	Reamplificación	Secuenciación
Dja'de	SK-R3	2DJ6	6	2	3
<b>Total Dja'de</b>			<b>35</b>	<b>15</b>	<b>34</b>
Garai	G3	GA3	10	9	2
Garai	G8	GA8	2	2	2
Garai	G11	GA11	2	2	0
Garai	G15	GA15	10	9	10
Garai	G16	GA16	3	4	0
Garai	G21	GA21	1	1	0
Garai	G22	GA22	5	6	2
Garai	G23	GA23	1	2	4
Garai	G24	GA24	2	2	4
Garai	G49	GA49	1	1	0
Garai	G50	GA50	6	6	13
<b>Total Garai</b>			<b>43</b>	<b>44</b>	<b>37</b>
Mari	1	1MA1	12	5	6
Mari	2	2MA2	9	4	5
Mari	3	1MA3	6	2	7
Mari	4	1MA4	3	2	6
Mari	4	2MA4	14	7	8
Mari	5	1MA5	8	2	8
Mari	6	1MA6	4	3	9
Mari	7	1MA7	7	2	2
Mari	8	1MA8	14	6	10
Mari	10	2MA10	9	4	7
Mari	11	3MA11	7	3	13
Mari	12	1MA12	9	4	8
<b>Total Mari</b>			<b>102</b>	<b>44</b>	<b>89</b>
Muge	TAS168	TA1	30	32	18
Muge	TAS168	TA7	7	0	0
Muge	TAS168	TA8	7	0	0
Muge	TAS214	TA2	40	43	15
Muge	TAS216	TA3	44	52	16
Muge	TAS216	TA9	5	1	0
<b>Total Muge</b>			<b>133</b>	<b>128</b>	<b>49</b>
Nerja	NE-NM82.1	1NE	8	8	0
Nerja	NE-1829	5NE	12	11	4
Nerja	NE-NM82.2	2NE	14	13	4
Nerja	NE-NT79	3NE	14	13	4
Nerja	NE-NAP	4NE	22	15	7
<b>Total Nerja</b>			<b>70</b>	<b>60</b>	<b>19</b>
El Pirulejo	1PI	1PI	13	13	6
El Pirulejo	2PI	2PI	10	10	4
<b>Total El Pirulejo</b>			<b>23</b>	<b>23</b>	<b>10</b>
Sant Pau	2 7608	SP2	6	0	0
Sant Pau	6 7214	1SP6	5	0	0
Sant Pau	6 7214	2SP6	5	0	0
Sant Pau	7 7720	1SP7	4	0	0
Sant Pau	7 7720	2SP7	5	0	0
Sant Pau	18 7758	1SP18	2	1	2
Sant Pau	18 7758	2SP18	5	0	0
Sant Pau	19 7928	SP19	5	0	0
<b>Total Sant Pau</b>			<b>37</b>	<b>1</b>	<b>2</b>

**Tabla ME12.** (Continuación).

<b>Procedencia</b>	<b>Individuo</b>	<b>Muestra</b>	<b>Amplificación</b>	<b>Reamplificación</b>	<b>Secuenciación</b>
Tell Halula	H-3	TH1	7	4	2
Tell Halula	H-3	TH2	19	15	14
Tell Halula	H-4	TH3	9	6	0
Tell Halula	H-4	TH4	16	12	19
Tell Halula	H-7	TH5	19	17	5
Tell Halula	H-7	TH6	10	5	8
Tell Halula	H-8	TH7	10	10	4
Tell Halula	H-8	TH8	8	5	8
Tell Halula	H-24	1H24	15	2	3
Tell Halula	H-25	2H25	12	1	2
Tell Halula	H-26	1H26	9	0	0
Tell Halula	H-28	1H28	15	2	2
Tell Halula	H-30	2H30	5	1	4
Tell Halula	H-35	1H1	7	0	0
Tell Halula	H-36	1H2	15	2	2
Tell Halula	H-37	2H20	11	4	4
Tell Halula	H-43	2H31	14	2	2
Tell Halula	H-44	1H32	14	2	2
Tell Halula	H-45	1H17	26	4	6
Tell Halula	H-47	1H18	18	6	13
Tell Halula	H-48	1H33	8	3	2
Tell Halula	H-49	1H15	8	1	6
Tell Halula	H-53	1H14	16	2	5
Tell Halula	H-54	2H34	5	0	0
Tell Halula	H-56	1H16	16	2	0
Tell Halula	H-57	1H21	28	7	7
Tell Halula	H-64	2H13	17	3	9
Tell Halula	H-65	2H10	14	2	6
Tell Halula	H-66	1H9	9	3	0
Tell Halula	H-66	2H9	16	3	3
Tell Halula	H-68	H12	15	2	6
Tell Halula	H-70	2H11	29	2	9
Tell Halula	H-70	3H11	13	3	4
Tell Halula	H-71	H35	3	2	0
Tell Halula	H-72	1H36	3	2	4
Tell Halula	H-73	1H37	3	2	4
Tell Halula	H-74	2H38	3	2	0
Tell Halula	H-76	1H45	2	1	0
Tell Halula	H-78	1H40	3	2	0
Tell Halula	H-78	2H40	3	2	0
Tell Halula	H-79	2H41	2	1	0
Tell Halula	H-80	H42	2	1	0
Tell Halula	H-82	H44	2	1	0
Tell Halula	H-85	1H43	2	1	0
Tell Halula	H-89	1H46	2	1	0
Tell Halula	H-90	2H47	2	1	0
Tell Halula	H-91	1H39	3	2	0
<b>Total Tell Halula</b>			<b>488</b>	<b>154</b>	<b>165</b>
Tell Ramad	R65-7I	TR1	4	4	2
Tell Ramad	R63-1	TR2	2	2	0
Tell Ramad	R65(8)	TR3	3	3	0

Tabla ME12. (Continuación).

Procedencia	Individuo	Muestra	Amplificación	Reamplificación	Secuenciación
Tell Ramad	R65(8)	TR4	4	4	0
Tell Ramad	R65-10	TR5	3	3	0
Tell Ramad	R65-10	TR6	6	6	2
Tell Ramad	R65-1	TR7	16	15	23
Tell Ramad	R65-3I	TR8	7	6	6
Tell Ramad	R65-4II	TR9	24	13	13
Tell Ramad	R65-4II	TR10	3	3	0
Tell Ramad	R65-4II	TR11	23	13	12
Tell Ramad	R65-14	TR12	8	7	7
Tell Ramad	R65-C-8 SE A	TR13	2	2	0
Tell Ramad	R69(2)	TR14	12	10	6
Tell Ramad	R69(2)	TR15	5	5	0
Tell Ramad	R65-C-8 SE B	TR16	10	9	6
Tell Ramad	R65-1S	TR18	9	11	6
Tell Ramad	R66-N4-Nº 400	TR19	15	12	4
<b>Total Tell Ramad</b>			<b>156</b>	<b>128</b>	<b>87</b>
Toledo	K-13	TO1	24	3	8
<b>Total Toledo</b>			<b>24</b>	<b>3</b>	<b>8</b>
Tres Montes	TM-2	2TM1	15	3	7
Tres Montes	TM-3	1TM2	14	2	9
Tres Montes	TM-6	1TM4	13	3	2
Tres Montes	TM-6	2TM4	7	2	4
Tres Montes	TM-8	1TM5	11	2	4
Tres Montes	TM-11	1TM6	14	2	8
Tres Montes	TM-14	TM7	1	0	0
Tres Montes	TM-18	TM11	7	2	7
Tres Montes	TM-19	TM12	2	0	0
<b>Total Tres Montes</b>			<b>84</b>	<b>16</b>	<b>41</b>
Zafarraya	ZA13	ZA13	15	16	15
Zafarraya	ZA16	ZA16	14	15	4
<b>Total Zafarraya</b>			<b>29</b>	<b>31</b>	<b>19</b>
<b>Total general</b>			<b>1618</b>	<b>908</b>	<b>769</b>

## 9.2. Protocolo experimental

### 9.2.1. Cebadores de la reacción de amplificación

En la presente tesis doctoral se estudiaron los polimorfismos de secuencia de un fragmento de la Región Hipervariable I (HVRI o HVSI) del D-loop mitocondrial (también región Control o *CRS*).

Debido a que el material genético de partida está presente en los extractos a muy baja concentración y pureza se utilizó la estrategia de la “PCR anidada” (*Nested PCR*) para alcanzar la cantidad de producto necesaria para la secuenciación. El proceso consiste en una primera amplificación con “cebadores externos” seguida de una reamplificación del producto generado con otra pareja de “cebadores internos” que hibridan con regiones internas al producto de la primera reacción. El empleo de cebadores de amplificación internos en la segunda reacción en lugar de la reamplificación con el mismo juego de cebadores resulta mucho más eficiente (ALBERT y FENYO 1990) puesto que de esta manera se asegura la copia de aquellas cadenas de DNA terminadas prematuramente.

En un principio los cebadores utilizados fueron aquellos diseñados para la amplificación de muestras de Tierra del Fuego (GARCÍA-BOUR *et al.* 2004). La secuencia, localización y características de estos cebadores así como del fragmento amplificado se describen en la siguiente tabla (**Tabla ME13**).

**Tabla ME13.** Características de los cebadores diseñados para amplificar un segmento de la Región Hipervariable I de la Región Control del DNA mitocondrial de muestras antiguas de Tierra del Fuego (GARCÍA-BOUR *et al.* 2004).

Nombre	Tipo de cebador	Secuencia del cebador	Posiciones del cebador	Fragmento amplificado	Longitud secuencia legible
L16154	Externo	5' AAT ACT TGA CCA CCT GT 3'	16138-16154	16138-16416 279 pb	
H16400	Externo	5' TTC ACG GAG GAT GGT GG 3'	16416-16400		
L16252	Externo	5' CAT CAA CTG CAA CTC CA 3'	16236-16252	16236-16416 180 pb	
L16158	Interno	5' CTT GAC CAC CTG TAG TA 3'	16142-16158	16142-16410 269 pb	16159-16393 235 pb
H16394	Interno	5' GAG GAT GGT GGT CAA GG 3'	16410-16394		
L16283	Interno	5' CCA CTA GGA TAC CAA CA 3'	16267-16283	16267-16410 144 pb	16284-16393 110 pb



Si bien estos cebadores habían proporcionado una buena eficiencia de amplificación de las muestras de Tierra del Fuego (GARCÍA-BOUR *et al.* 2004), sus características no se adecuaban a las muestras del presente trabajo. Los primeros ensayos realizados con estos *primers* revelaron una muy baja eficiencia de amplificación y la aparición de “smear” y dobles y triples bandas en los geles de electroforesis de las amplificaciones. Nos planteamos entonces diseñar un nuevo juego de cebadores para amplificar las muestras de la presente tesis doctoral.

Para ello se seleccionó un fragmento de 416pb (posiciones 16024→16440) de la HVRI del mtDNA. Debido al estado de fragmentación del DNA antiguo (PÄÄBO *et al.* 1998; PÄÄBO 1989) se planteó amplificar esta región mediante dos amplificaciones que proporcionaran fragmentos solapantes de longitud no superior a los 200pb.

Los nuevos cebadores se diseñaron mediante el programa “Primers for the World Wide Web”, disponible en internet de forma gratuita . Este programa permite seleccionar la pareja de cebadores más adecuada en una determinada región de acuerdo con diferentes parámetros establecidos por el usuario. Los criterios de selección introducidos en el programa fueron:

- Longitud de los cebadores: 20pb.
- Mínima Tm: 57; Máxima Tm: 63.
- Bases ambiguas: 0.

Para escoger los cebadores del primero de los dos fragmentos solapantes se analizaron las posiciones 16024→16300. La zona de localización de estos cebadores se acotó dentro de las regiones 16024→16200 (cebador “Forward”) y 16230→16300 (cebador “Reverse”). Con estas condiciones el programa generó un total de 51 cebadores “Forward” y 17 cebadores “Reverse”. De la misma manera, se buscaron los cebadores del segundo fragmento de amplificación dentro de las posiciones 16221→16300 (“Forward”, 17 cebadores) y 16380→16440 (“Reverse”, 24 cebadores).

El conjunto de cebadores generados por el programa “Primers for the WWW” fueron exportados al programa “CGC FindPatterns” (Accelrys. San Diego, USA). Este

programa permite analizar los cebadores por parejas (Forward-Reverse) y realizar una selección de acuerdo con ciertos parámetros. Se seleccionaron las siguientes opciones:

- Longitud del cebador: 18-20pb.
- Mínima complementariedad en extremo 3’.
- Mínima complementariedad entre parejas de cebadores.
- Ausencia de homología con otras regiones.
- Contenido G+C: 45-65%.
- Tm: 45-65°C.
- Diferencia de Tm entre pareja de cebadores: < 2°C.
- Longitud del producto de amplificación: 130-190pb.
- Contenido G+C del producto de amplificación: 40-55%.

De las 30 parejas de cebadores que cumplían los requisitos establecidos, se seleccionaron 8 en base a sus características, su localización y a la longitud del segmento amplificado. Estos cebadores se describen en la **Tabla ME14** y en la **Figura FME9**.

**Tabla ME14.** Características de los cebadores utilizados para amplificar dos fragmentos solapantes de la Región Hipervariable I de la Región Control del DNA mitocondrial.

	Nombre	Tipo de cebador	Secuencia del cebador	Posiciones del cebador	Fragmento amplificado	Longitud secuencia legible
Primer fragmento	L16112	Externo	5’ CGT ACA TTA CTG CCA GCC 3’	16095-16112	16095-16280 186pb	16126-16258 133pb
	H16262	Externo	5’ TGG TAT CCT AGT GGG TGA G 3’	16280-16262		
	L16125	Interno	5’ GCC AGC CAC CAT GAA TAT TG 3’	16106-16125	16106-16276 171pb	
	H16259	Interno	5’ ATC CTA GTG GGT GAG GGG 3’	16276-16259		
Segundo fragmento	L16251	Externo	5’ CAC ACA TCA ACT GCA ACT CC 3’	16232-16251	16232-16399 168pb	16258-16369 112pb
	H16380	Externo	5’ TCA AGG GAC CCC TAT CTG AG 3’	16399-16380		
	L16257	Interno	5’ TCA ACT GCA ACT CCA AAG CC 3’	16238-16257	16238-16389 152pb	
	H16370	Interno	5’ CCT ATC TGA GGG GGG TCA TC 3’	16389-16370		

Una vez diseñados los cebadores, se analizó la presencia de regiones de autocomplementariedad en cada uno y la posible formación de dímeros entre ellos mediante el programa “DNASTAR<sup>®</sup>”, disponible también gratuitamente en internet. Los resultados de este análisis y las características de los cebadores se describen en la **Tabla ME15**.

**Tabla ME15.** Características del nuevo diseño de cebadores diseñados con el programa “CGC FindPatterns” (Accelrys. San Diego, USA) (\*) Resultados obtenidos mediante el programa DNASTAR<sup>®</sup>.

Cebador	L16112	H16262	L16125	H16259	L16251	H16380	L16257	H16370
%GC	55.6	52.6	50	61.1	50	55.0	50	60
Tm	50	48.9	51.4	50.9	51.6	51.6	53.2	52.4
Diferencia Tm	1.1		0.6		0.1		0.8	
<i>T annealing</i>	52.3		52.8		53.7		53.4	
%GC producto	45.7		46.2		48.8			
Auto-complementariedad*	no		no		no		no	
Complementariedad* entre cebadores	no		no		no		no	

```

1
6
0
9
3
100 110 120 130 140 150 160
TTCGTACATTACTGCCAGCCACCATGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTACAT
AAGCATGTAATGACGGTCGGTGGTACTTATAACATGCCATGGTATTTATGAACTGGTGGACATCATGTA
170 180 190 200 210 220 230
AAAAACCCAATCCACATCAAACCCCTCCCCATGCTTACAAGCAAGTACAGCAATCAACCCTCAACTA
TTTTTGGGTTAGGTGTAGTTTTGGGGAGGGGTACGAATGTTTCGTTTCATGTCGTTAGTTGGGAGTTGAT
240 250 260 270 280 290 300
TCACACATCAACTGCAACTCCAAAGCCACCCTCACCCACTAGGATACCAACAACCTACCCACCTTA
AGTGTGTAGTTGACGTTGAGGTTTTCGGTGGGGAGTGGGTGATCCTATGGTGTGTTGGATGGGTGGGAAT
0 310 320 330 340 350 360 370
ACAGTACATAGTACATAAAGCCATTTACCGTACATAGCACATTACAGTCAAATCCCTTCTCGTCCCCAT
TGTCATGTATCATGTATTTTCGGTAAATGGCATGTATCGTGTAATGTCAGTTTAGGGAAGAGCAGGGGTA
70 380 390 400 410 420 430
GGATGACCCCCCTCAGATAGGGGTCCCTTGACCACCATCCTCCGTGAAATCAATATCCCGCACAAAGAGT
CCTACTGGGGGAGTCTATCCCAGGGAACTGGTGGTAGGAGGCACTTTAGTTATAGGGCGTGTCTCA

```

**Figura FME9.** Detalle de la zona de la región Hipervariable I del DNA mitocondrial estudiada en la presente tesis doctoral. Los ocho cebadores utilizados en este estudio aparecen marcados con cajas (cebadores externos) y en negrita (cebadores internos) en versales (primer segmento) y cursiva (segundo segmento).

## 9.2.2. Reacción de amplificación

### 9.2.2.1. Precauciones

La preparación de las reacciones de amplificación y la extracción del DNA son las fases de análisis de material antiguo que más riesgo de contaminación entrañan por parte del manipulador. Durante la amplificación puede producirse además la contaminación por amplicones de anteriores reacciones (*carry over*).

Para tratar de minimizar la contaminación durante la amplificación se emplearon las siguientes medidas:

- Las reacciones de amplificación se prepararon en el interior de una campana de flujo laminar vertical (Telstar Micro-V o Heraeus modelo Gemini).
- El interior de la campana y el material no desechable se limpió con lejía al 70% antes de cada experimento.
- Se irradió el interior de la campana y el material de trabajo con luz ultravioleta durante 20 y 30 minutos antes de cada experimento. En el laboratorio de Genética Forense y Genética de Poblaciones de la U.C.M. se irradió también durante el mismo tiempo la habitación destinada a la preparación de las reacciones de amplificación.
- Durante todo el proceso se usaron bata y guantes de laboratorio.
- Se empleó material estéril de un solo uso (puntas de pipeta y microtubos eppendorf).
- Se usaron pipetas de desplazamiento positivo con puntas de pipeta con filtro para aerosoles.
- Se preparó una única mezcla de reacción (“mezcla maestra” o “*master mix*”) para todas las muestras en lugar de una reacción de amplificación independiente para cada una de ellas. La mezcla maestra contiene todos los reactivos de *PCR* a excepción del DNA, que se añade una vez ésta ha sido repartida en los diferentes tubos. La “*mix*” se prepara manteniendo constante la concentración final de los diferentes reactivos pero variando su volumen en función del número de muestras que se deseen analizar. Generalmente se calcula el “*mix*” para una o dos muestras más de las que se pretende amplificar para compensar posibles errores de pipeteo.
- En cada ensayo se amplificaron como máximo 12 muestras.

### 9.2.2.2. Preparación de las reacciones de amplificación

- Se limpia el interior de la campana de flujo laminar con lejía al 70% y el material no desechable necesario para la PCR (micropipetas automáticas, gradilla para microtubos, recipiente para desechar puntas).
- Se conecta el dispositivo de flujo de la campana pulsando el interruptor situado en su parte exterior superior derecha.
- Se disponen, abiertos sobre la gradilla, un microtubo eppendorf de 1.5 ml y tantos eppendorfs de PCR (0.66  $\mu$ l) como reacciones pretendan realizarse. Se dispone también sobre la gradilla una alícuota de agua estéril para PCRs.
- Se colocan un par de guantes de laboratorio (de látex o cloruro de polivinilo) sobre la superficie de la campana.
- Una vez dispuesto el material, se conecta la luz UV de la campana pulsando el interruptor situado en su interior y se deja durante 20-30 minutos.
- Se carga primero el control negativo de amplificación (5 $\mu$ l) cerrando a continuación el tubo.
- Se prepara la mezcla de amplificación en el interior del eppendorf de 1.5 ml. El volumen a añadir de cada uno de los reactivos depende de la concentración del *stock* y del número de reacciones de amplificación. Para los diferentes reactivos utilizados durante el transcurso de esta tesis doctoral las cantidades correspondientes a una reacción son las que siguen:

#### GIBCO-BRL

	Concentración stock	Concentración final en reacción	Volumen (1 reacción)
Agua	-	-	11.05
PCRb	10x	1x	2.5 $\mu$ l
MgCl <sub>2</sub>	50mM	2mM	1 $\mu$ l
W1 (jabón)	1%	0.04%	1 $\mu$ l
DNTPs (GIBCO)	1.25 mM	0.2 mM	4 $\mu$ l
Primers	100 $\mu$ M	0.4 $\mu$ M	0.1 $\mu$ l/cada
Taq	5 u/ $\mu$ l	0.05 u/ $\mu$ l	0.25 $\mu$ l
DNA	-	-	5 $\mu$ l
Total	-	-	25 $\mu$ l

Taq Expand High Fidelity (Roche®)

	Concentración stock	Concentración final en reacción	Volumen (1 reacción)
Agua	-	-	14.61 µl
PCRb	10x	1x	2.5 µl
MgCl <sub>2</sub>	25mM	2mM	2 µl
dNTPs (Biotools)	10 mM	0.2 mM	0.5 µl
Primers	100 µM	0.4 µM	0.1 µl/cada
Taq	3.5 u/µl	0.0026 u/µl	0.1857 µl
DNA	-	-	5 µl
Total			25 µl

Biotools®

	Concentración stock	Concentración final en reacción	Volumen (1 reacción)
Agua	-	-	14.55 µl
PCRb	10x	1x	2.5 µl
MgCl <sub>2</sub>	50 mM	2 mM	1 µl
DNTPs (Biotools)	10 mM	0.2 mM	0.5 µl
Primers	100 µM	0.4 µM	0.1 µl/cada
Taq	1 u/µl	0.05 u/µl	1.25 µl
DNA	-	-	5 µl
Total	-	-	25 µl

El cálculo del volumen total de la mezcla se realiza multiplicando el número de reacciones por el volumen calculado para una reacción de 25 µl.

- Se reparte la mezcla de amplificación a razón de 20 µl por tubo. Una vez realizado el reparto deben cerrarse los tubos para evitar que durante la adición del DNA exista traspaso de uno a otro.
- Se añade aceite mineral para PCR (*Mineral Oil for PCR*, Braun) al control negativo de amplificación si ésta va a realizarse en un termociclador que lo precise (Perkin Elmer TC1).
- Se añade la cantidad correspondiente de DNA (normalmente 5µl, aunque puede variar en función del volumen final de reacción y del volumen disponible de muestra). Se añade aceite mineral para PCR si es necesario y se cierra el tubo.
- Si existe, se carga el control negativo de *carry over* con agua estéril. Se añade aceite mineral para PCR si es necesario.

- El tubo en el que se ha fabricado la mezcla de reacción se amplifica junto con las muestras como control negativo adicional para comprobar la posible contaminación de los reactivos y del proceso experimental. Se añade aceite a esta muestra si es necesario.
- Se trasladan las muestras al termociclador, dónde se depositan en el interior del bloque térmico.
- Se conecta el programa de amplificación correspondiente.

#### 9.2.2.3. Programas de amplificación

En la realización del presente trabajo se utilizaron los siguientes modelos de termocicladores:

- Termociclador Perkin Elmer modelo TC1 (Unidad de Antropología. Universidad de Barcelona).
- Termociclador Perkin Elmer modelo TC1 (Departamento de Toxicología y Legislación Sanitaria. Universidad Complutense de Madrid).
- Termociclador Eppendorf modelo *Mastercycler personal* (Departamento de Toxicología y Legislación Sanitaria. Universidad Complutense de Madrid).
- Termociclador Eppendorf modelo *Mastercycler personal* (Departamento de Toxicología y Legislación Sanitaria. Universidad Complutense de Madrid).

Los programas de amplificación de secuencias de DNA mitocondrial se detallan en la **Tabla ME16**.

**Tabla ME16.** Programas de amplificación de la HVRI del mtDNA empleados en el presente estudio.

	<b>Cebadores</b>	<b>Perkin Elmer TC1</b>	<b>Eppendorf</b>
<b>Amplificación</b>	<b>L16154, H16400</b>	94°C 5min	-
		40°C 1min 1 ciclo	
		72°C 1min	
		94°C 1min	
<b>Amplificación</b>	<b>L16252, H16400</b>	40°C 1min 35 ciclos	
		72°C 1min	
		94°C 5min	-
		48°C 30seg 1 ciclo	
<b>Reamplificación</b>	<b>L16158, H16394 ó L16253, H16394</b>	72°C 30seg	
		91°C 30seg	
		48°C 30seg 39 ciclos	
		72°C 30seg	
<b>Amplificación</b>	<b>L16112-H16262</b>	94 °C 5min 1 ciclo	94 °C 5min 1 ciclo
		94°C 1min	94°C 45 seg
		52°C 1min 40 ciclos	52°C 45 seg 40 ciclos
		72°C 1min	72°C 45 seg
<b>Amplificación</b>	<b>L16251-H16380</b>	72°C 5min 1 ciclo	72°C 5min 1 ciclo
		94 °C 5min 1 ciclo	94 °C 5min 1 ciclo
		94°C 1min	94°C 45seg
		53°C 1min 40 ciclos	53°C 45 seg 40 ciclos
<b>Reamplificación</b>	<b>L16125-H16259</b>	72°C 1min	72°C 45seg
		72°C 5min 1 ciclo	72°C 5min 1 ciclo
		94 °C 5min 1 ciclo	94 °C 5min 1 ciclo
		94°C 1min	94°C 45 seg
<b>Reamplificación</b>	<b>L16257-H16370</b>	52°C 1min 30 ciclos	52°C 45seg 30 ciclos
		72°C 1min	72°C 45seg
		72°C 5min 1 ciclo	72°C 5min 1 ciclo
		94 °C 5min 1 ciclo	94 °C 5min 1 ciclo



## 10. ELECTROFORESIS EN GEL DE AGAROSA

### 10.1. Fabricación de los geles de agarosa

Los geles de agarosa para electroforesis se preparan en el interior de unos moldes que se encajan en la cubeta de electroforesis. En función del número de muestras, de la velocidad de separación y del grado de resolución deseados pueden emplearse cubetas de diferente capacidad (KELLER y MANAK 1989).

Para el análisis de los resultados de amplificación se fabricaron geles de agarosa al 2% empleando cubetas de 50 y 70ml de capacidad. Para la preparación de los geles se siguieron las siguientes proporciones:

	Concentración stock	Concentración final	Cantidades a añadir según volumen de gel	
			50ml	70ml
Agarosa	-	2%	1.4 g	1g
H <sub>2</sub> O	-	-	65.1 ml	46.5ml
TBE (10x)	10x	0.5x	3.5 ml	2.5ml
BrEt	100mg/ml	5 µg/ml	3.5 µl	2.5µl

Los geles se preparan de la manera siguiente:

- En un matraz erlenmeyer se pesa la agarosa y se le añade la cantidad correspondiente de agua destilada, se mezcla brevemente por agitación y se calienta en el microondas durante unos pocos minutos vigilando que no hierva. Antes de que esté totalmente disuelta se extrae del microondas, se agita y se introduce unos minutos más.
- Se enfría por debajo de los 60° en un agitador magnético o poniendo las paredes del matraz debajo del grifo de agua fría.
- Una vez enfriado, se pesa de nuevo y se añade la cantidad de agua que se ha evaporado durante la ebullición. A continuación se añaden el TBE y por último el Bromuro de Etidio.
- Se vierte la solución en el soporte de la cubeta de electroforesis, que se ha sellado previamente por los lados con celo o cinta aislante y se coloca el peine que formará los pocillos del gel. Se deja solidificar a temperatura ambiente durante aproximadamente 20 minutos.

- Se extrae el peine, se retiran las tiras de cinta que sellan el gel y se deposita el soporte en el interior de la cubeta de electroforesis. Se llena la cubeta con tampón de electroforesis procurando cubrir el gel.

## 10.2. Carga de las muestras en el gel

- Se transfieren de 8 a 12  $\mu$ l de los productos de amplificación a unos nuevos eppendorfs de PCR o a la superficie de un trozo de *parafilm*. Cuando estas muestras proceden de la primera amplificación este proceso se realiza en el interior de la campana de flujo laminar tomando todas las precauciones necesarias para evitar la contaminación. Se traspasan en primer lugar los controles negativos y a continuación las muestras para evitar que éstas los contaminen por *carry over*. Si existe control positivo, éste debe ser manipulado en el exterior de la campana. Sin embargo, los productos resultantes de la segunda amplificación no corren el riesgo de contaminarse debido a que el número de moléculas endógenas es muy superior a las posibles moléculas contaminantes que pudieran introducirse durante esta fase. Por ello su manipulación puede realizarse fuera de la campana bajo condiciones no estériles.

Antes de cargar las muestras en los pocillos del gel es necesario añadirles un tampón de carga. El tampón de carga contiene una sustancia de elevada densidad (glicerol, sucrosa...) que ayuda a retener las muestras en los pocillos del gel evitando que difundan y algún tipo de colorante que permita visualizar la migración de la muestra en el gel. La composición del tampón utilizado es la siguiente:

- Glicerol                    50%
- Azul de bromofenol    0.2%
- xilen-cianol                0.2%

El tampón de carga se añade en una proporción aproximada de 1/10 respecto al volumen de muestra y se centrifuga brevemente para facilitar su mezcla.

El **azul de bromofenol** migra a través de la agarosa aproximadamente 2.2 veces más rápido que el **xilen-cianol** independientemente de la concentración de agarosa (SAMBROOK *et al.* 1989). En geles con una concentración final de TBE 0.5x, la

velocidad de migración del azul de bromofenol es similar a la de un fragmento lineal de DNA bicatenario de unos 300 pb, mientras que la del xilen-cianol puede equipararse a la de un fragmento de las mismas características pero de 4 kb. En una electroforesis de estas características los productos de amplificación del presente trabajo, de aproximadamente 150 pb, se situarían entre ambos frentes de colorante, más próximos al azul de bromofenol.

- Se añaden aproximadamente 2µl de tampón de carga a cada muestra y se centrifuga brevemente en una microcentrífuga para facilitar su mezcla.
- Se cargan las muestras con el colorante en los pocillos con una micropipeta y puntas autoclavadas desechables.
- Se cubre la cubeta de electroforesis con una tapa de metacrilato que conecta los electrodos a la fuente de alimentación y se ajusta la fuente de alimentación al voltaje deseado (entre 100 y 60V).
- Se deja transcurrir la electroforesis hasta que el frente del colorante más próximo al ánodo haya recorrido más de la mitad de la matriz del gel.

### **10.3. Visualización de los geles**

- Una vez se ha producido la migración de los productos de amplificación en el gel, se desconecta la fuente de alimentación, se extrae el gel del molde con cuidado y se coloca en un transiluminador de rayos UV. El bromuro de etidio es una molécula plana que se intercala en la doble hélice de DNA. Cuando esta molécula se expone a luz UV se excita emitiendo fluorescencia de color azul. Para observar el gel deben protegerse los ojos de la exposición a la luz UV mediante unas gafas o una máscara o bien situar una pantalla transparente protectora sobre el transiluminador. El gel puede fotografiarse mediante el sistema digital de análisis acoplado al transiluminador o mediante un film Polaroid de alta sensibilidad (Tipo 55) tras una exposición de entre 40 y 60 segundos dependiendo de la intensidad de las bandas de DNA.

## **11. AISLAMIENTO DEL PRODUCTO DE AMPLIFICACIÓN**

Una vez amplificado el DNA debe aislarse del resto de los componentes de la reacción para poder ser secuenciado. Existen diversos protocolos que permiten aislar el producto de amplificación, directamente de la reacción de *PCR* o a partir de los geles de agarosa tras la electroforesis.

La eficiencia de los diferentes métodos es variable y la elección de uno u otro dependerá en cada caso de las características del producto amplificado. Para el caso concreto del estudio de secuencias cortas de DNA antiguo, se optó por un protocolo de purificación a partir de agarosa de bajo punto de fusión debido a que es el que proporciona una mayor eficiencia de recuperación. Sin embargo, este método plantea ciertos problemas. El más grave es la copurificación junto con el DNA de moléculas inhibitoras de reacciones enzimáticas que pueden estar presentes en la agarosa (SAMBROOK *et al.* 1989). Para evitar o minimizar este proceso se utilizó agarosa de grado biotecnológico, libre de moléculas contaminantes y con garantía de ausencia de actividad RNA-ásica y DNA-ásica.

### **11.1. Purificación de DNA de agarosa de bajo punto de fusión**

#### ***11.1.1. Corte de las bandas del gel***

La agarosa utilizada para purificar el DNA amplificado fue una agarosa de bajo punto de fusión (NuSieve GTG, FMC Iberlabo) de grado biotecnológico y con las garantías citadas anteriormente.

La agarosa NuSieve posee un bajo punto de fusión ( $\leq 65^{\circ}$  C) y de gelificación ( $\leq 35^{\circ}$  C). Por lo tanto es aconsejable que tanto la solidificación del gel como la posterior electroforesis se realicen en frío, a  $4^{\circ}$  C.

Este tipo de agarosa permite resolver fragmentos comprendidos entre 1000 y 8pb. Sin embargo, el rango de separación eficiente de moléculas lineales de DNA depende del tipo de tampón utilizado en la fabricación del gel. En general, los fragmentos de DNA más pequeños se separan mejor en geles con TBE mientras que aquellos de mayor tamaño se

separan mejor en geles con TAE. En la **Tabla ME17** se muestra el rango de separación del DNA con diferentes concentraciones de agarosa empleando TAE o TBE a concentración 1x.

**Tabla ME17.** Rango de separación del DNA en diferentes concentraciones de agarosa empleando TAE 1x o TBE 1x.

Rango de separación de DNA en pb	% agarosa en el gel	
	TAE	TBE
500-1000	2.5	2.0
150-700	3.0	2.5
100-450	3.5	3.0
70-300	4.0	3.5
10-100	4.5	4.0
8-50	5.0	4.5

Para la resolución de un único fragmento de longitud comprendida entre 150 y 200pb se recomienda trabajar con concentraciones entre 3 y 3.5% de agarosa en TBE. Sin embargo pueden emplearse concentraciones menores, de hasta un 2%. Cuanto menor sea la concentración de agarosa menos cantidad de residuos de la misma quedarán en el producto final.

El proceso a seguir es el siguiente:

- Se fabrica un gel de agarosa NuSieve GTG al mínimo porcentaje de agarosa compatible con la resolución deseada. Para un gel de 70 ml al 2% se añaden:

1´4 g de agarosa NuSieve GTC  
 3´5 ml de TBE 10x  
 3´5 µl de Bromuro de Etidio  
 H<sub>2</sub>O hasta 70ml

El bajo punto de fusión de este tipo de agarosa hace que su ebullición se produzca rápidamente y de manera brusca. Por ello se recomienda utilizar un recipiente de capacidad 2-4 veces mayor que el volumen final de la solución, disolver la agarosa en el microondas a baja temperatura y extraer el recipiente antes de que comience la ebullición.

- Se deja solidificar en nevera durante aproximadamente 20 minutos.  
 - Se carga todo el volumen (aproximadamente 17µl) de aquellas muestras que han proporcionado banda tras la reamplificación. La carga debe realizarse dejando un pocillo

libre entre una y otra para evitar el traspaso de volumen entre muestras adyacentes. Una vez cargadas, se realiza la electroforesis a 70 V durante aproximadamente 1 hora.

- Se visualiza el gel en el transiluminador y se cortan las bandas correspondientes a la amplificación con un bisturí, introduciéndolas en microtubos eppendorf de 1.5 ml previamente autoclavados y radiados durante 30 minutos en luz UV. El corte de las bandas debe realizarse en un tiempo inferior a 10-12 segundos para prevenir su degradación (ORREGO y KING 1990).

- Tras cortar las bandas, se observa de nuevo el gel en el transiluminador para ver si quedan restos de alguna de las muestras.

Una vez aislada la banda del gel, el DNA puede purificarse precipitándolo con sales o mediante kits comerciales, que minimizan el tiempo de manipulación. En el presente trabajo inicialmente se purificaron las muestras por precipitación y posteriormente se emplearon sistemas de purificación comerciales.

### 11.1.2. Aislamiento del DNA

#### 11.1.2.1. Precipitación con sales

- Se perfora la base de los microtubos eppendorf de 600µl que contienen la banda cortada con una aguja hipodérmica estéril procurando no tocar con ella la agarosa.
- Se colocan estos eppendorfs dentro de otros de 2ml previamente rotulados, autoclavados y radiados y se centrifugan durante 1 min a máxima velocidad en una microcentrífuga. Mediante este proceso se consigue licuar la agarosa, que pasa al eppendorf inferior.
- Se añaden 4 volúmenes (unos 70 µl si partimos de 17 µl de DNA) de NaCl 0.1M en TE (para DNA con tamaño superior a 200 pb) o de NaCl 0.2 M en TE (para DNA con tamaño inferior a 200 pb).
- Se calientan los tubos a 70 °C durante 5 min en un baño térmico para favorecer la disolución total de la agarosa.
- Se congelan durante 10 minutos depositándolos en nieve carbónica-etanol.
- Se centrifugan inmediatamente después durante 15 min en una microcentrífuga a velocidad máxima, tras lo que la agarosa queda en el fondo en forma de un precipitado blanco. Se recupera entonces todo el sobrenadante y se pasa a unos nuevos eppendorfs.
- Si se parte de mucho volumen o de poca cantidad de DNA, como es el caso, resulta recomendable concentrar la muestra. Para ello se acaban de llenar los eppendorfs con n-Butanol y se agitan fuertemente con el vórtex para mezclar las dos fases. A continuación se elimina la fase superior, que es la que contiene el butanol, dejando la inferior o acuosa, que contiene el DNA. Este proceso se repite durante 2 o 3 veces.
- Se añade aproximadamente 1 volumen de Isopropanol (2- Propanol) para precipitar el DNA.
- Se congelan los tubos con nieve carbónica-etanol.
- Se centrifugan a máxima velocidad durante 15 minutos. Tras la centrifugación el DNA forma un *pellet*, muchas veces no visible, adherido a la pared del tubo. Se elimina entonces con cuidado todo el sobrenadante con ayuda de una pipeta *Pasteur* procurando no tocar el *pellet*.
- Para eliminar los restos de Isopropanol se añaden aproximadamente 700µl de Etanol al 80%, se agita, y se elimina el sobrenadante. El lavado con etanol puede desenganchar

el *pellet* de la pared del tubo, por lo que la absorción del sobrenadante debe realizarse con especial cuidado.

Mediante este proceso puede llegar a recuperarse desde un 85% hasta un 95% de DNA. Al finalizar el protocolo es posible que quede una pequeña cantidad de agarosa en el *pellet*, sin embargo esto no resulta demasiado problemático puesto que hasta concentraciones superiores a un 1.5 % la agarosa no interfiere en las reacciones enzimáticas.

#### *11.1.2.2. Extracción mediante Kits comerciales*

##### 11.1.2.2.1. Concert Gel Extraction System (BRL-Life Technologies®)

El kit Concert Gel Extraction System de la casa comercial BRL-Life Technologies permite purificar DNA directamente del gel de agarosa. La agarosa se disuelve con un tampón que contiene Perclorato Sódico y el DNA es adsorbido y retenido en una membrana de sílica incorporada a la columna, mientras que los residuos de tampón son eliminados mediante otro tampón que contiene alcohol. La elución del DNA se produce con TE concentrado a elevada temperatura. Este método permite aislar DNA a partir de geles fabricados con agarosa de bajo o alto punto de fusión con TAE o TBE.

Los componentes del kit son:

Tampón de solubilización L1: Perclorato sódico, acetato, TBE.

Tampón de lavado: NaCl, EDTA, Tris-HCl, Etanol 95%

TE: 10mM Tris-HCl, 0'1mM EDTA

Columnas, tubos de lavado y tubos de recuperación

- Este método exige saber el peso de las bandas cortadas a partir del gel. Para calcular el peso de la banda se pesan los eppendorfs antes y después de guardar en ellos las bandas y se restan los dos valores. Si el peso de la banda de gel es superior a 400mg es recomendable repartirla en dos tubos eppendorf y procesar cada uno de ellos por separado.

- Una vez pesadas las bandas se añaden 30µl del tampón L1 por cada 10mg de gel a cada uno de los eppendorf y se incuban a 50°C durante aproximadamente 15 minutos,



agitando cada 3 minutos aproximadamente. Una vez disuelto se deja incubando 5 minutos más

- Se coloca la columna dentro del tubo de lavado (suministrado) y se le añade la mezcla del paso anterior. Se tapan los tubos y se centrifugan en una microcentrífuga durante 1 minuto a 12000×g. Tras la centrifugación el DNA queda adherido al filtro de la columna mientras que el tampón con la agarosa disuelta pasan al tubo de lavado. Se extrae entonces la columna y se tira el líquido contenido en el tubo de lavado.

- Cuando el peso de la banda cortada es superior a 250mg o cuando el extracto final vaya a ser utilizado para reacciones de secuenciación o transcripción *in vitro* se recomienda lavar la columna añadiendo 500µl del tampón L1, incubar durante 1 minuto y centrifugar a 12000×g.

- Se coloca la columna de nuevo en el tubo de lavado y se añaden 700µl del tampón L2. Se tapa y se incuba durante 5 minutos a temperatura ambiente. Inmediatamente después se centrifuga durante 1 minuto a 12000×g.

- Se tira el líquido que ha pasado al tubo de lavado y se centrifuga de nuevo durante 1 minuto a 12000×g para eliminar posibles restos del tampón.

- Se deposita la columna en el interior de un tubo de recuperación de 1.5ml y se añaden 50µl de TE precalentado a 65°C directamente sobre el centro de la membrana. Se incuba durante 1 minuto a temperatura ambiente y se centrifuga durante 2 minutos a 12000×g.

Los extractos de DNA se guardan en el congelador a -20°C hasta su uso.

#### 11.1.2.2.2. QIAquick Gel Extraction Kit (Qiagen®)

Este protocolo está diseñado para extraer y purificar DNA entre 70pb y 10kb a partir de geles de agarosa de bajo o alto punto de fusión con TAE o TBE. Tras la disolución de la agarosa mediante el tampón QG, el DNA es adsorbido y retenido en una membrana de sílica incorporada a la columna en presencia de una elevada concentración de sales y a un pH determinado ( $\leq 7.5$ ) mientras que los contaminantes atraviesan la columna y son eliminados. El DNA finalmente es eluido con tampón Tris a pH 8.5.

El protocolo experimental es el siguiente:

- Una vez cortadas las bandas del gel se pesan tal y como se describe en el apartado anterior y se depositan en el interior de microtubos eppendorf de 1.5 ml. Se añaden 3 volúmenes de tampón QG por cada volumen de gel (100mg ~ 100µl).
- Se incuban los tubos a 50°C durante 10 minutos o hasta que la agarosa se haya disuelto completamente agitando cada 2-3 minutos durante la incubación.
- Una vez disuelto se comprueba que el color de la mezcla es amarillo, similar al *stock* de tampón QG. Este tampón contiene un indicador de pH que es amarillo a  $\text{pH} \leq 7.5$  y naranja o violeta a pHs mayores. La adsorción del DNA a la membrana de la columna es eficiente únicamente a  $\text{pH} \leq 7.5$ . Si el color de la mezcla es naranja o violeta se añaden 10µl de acetato sódico 3M pH 5 y se mezcla, lo que provocará que el color vire a amarillo.
- Se añade 1 volumen de isopropanol a la muestra y se mezcla. Este paso aumenta la eficiencia de recuperación de fragmentos pequeños (<500pb) o excesivamente grandes (>4kb).
- Se coloca cada columna QIAquick en un tubo colector de 2ml (suministrado), se aplica la muestra y se centrifuga durante 1 minuto a máxima velocidad (>12000 r.p.m.). El volumen máximo del reservorio del tubo colector es de 800µl. Si el volumen de muestra es superior a esta cantidad deberá repetirse este paso hasta agotar la totalidad de la muestra.
- Se vacía el tubo colector de líquido y se vuelve a colocar la columna QIAquick
- Se añaden 500µl de tampón QG a la columna QIAquick y se centrifuga durante 1 minuto a máxima velocidad. Este paso, recomendado cuando se desea secuenciar el DNA una vez purificado, permite eliminar totalmente cualquier traza de agarosa que pudiera quedar en la solución o adherida a la membrana de la columna.
- Se vacía el tubo colector de líquido y se vuelve a colocar la columna QIAquick.
- Se añaden 750µl de tampón PE a la columna, se deja reposar a temperatura ambiente durante 5 minutos y se centrifuga a máxima velocidad durante 1 minuto.
- Se vacía el tubo colector de líquido, se vuelve a colocar la columna QIAquick y se centrifuga de nuevo durante 1 minuto también a velocidad máxima. Este paso permite eliminar totalmente el etanol añadido junto con el tampón PE.
- Se coloca la columna QIAquick en un tubo eppendorf nuevo de 1.5ml y se añaden 30µl de tampón EB (10mM Tris-HCl pH 8.5) en el centro de la columna.

- Se deja la columna durante 1 minuto a temperatura ambiente y se centrifuga a continuación durante 1 minuto a máxima velocidad.
- Se guarda el DNA purificado a  $-20^{\circ}\text{C}$  hasta su empleo.

### **11.2. Purificación a partir de la mezcla de *PCR* (QIAquick PCR Purification Kit, Qiagen®)**

Este protocolo permite purificar fragmentos de DNA de un tamaño comprendido entre 100pb y 10kb directamente a partir de la mezcla de *PCR*. El mecanismo es muy similar al descrito en el apartado anterior:

- Se añaden 5 volúmenes de tampón PB por cada volumen de mezcla de *PCR* y se agita hasta su completa disolución.
- Se coloca cada columna QIAquick en un tubo colector de 2ml (suministrado), se añade la mezcla del paso anterior y se centrifuga a máxima velocidad ( $>12000$  r.p.m.) durante 1 minuto.
- Se vacía el tubo colector de líquido y se vuelve a colocar la columna QIAquick.
- Se añaden 750 $\mu\text{l}$  de tampón PE a la columna y se centrifuga a máxima velocidad durante 1 minuto.
- Se vacía el tubo colector de líquido, se vuelve a colocar la columna QIAquick y se centrifuga de nuevo durante 1 minuto también a velocidad máxima. Este paso permite eliminar totalmente el etanol añadido junto con el tampón PE.
- Se coloca la columna QIAquick en un tubo nuevo de microcentrífuga de 1.5ml y se añaden 30 $\mu\text{l}$  de tampón EB (10mM Tris-HCl pH 8.5) en el centro de la columna. Se centrifuga durante 2 minutos a velocidad máxima.
- Se guarda el DNA purificado a  $-20^{\circ}\text{C}$  hasta su empleo.

## 12. REACCIÓN DE SECUENCIA

Los productos de amplificación purificados se secuenciaron de forma automática utilizando terminadores de reacción marcados con fluorescencia (“dye terminators”) de la casa Applied Biosystems. Para ello se emplearon 4 kits de secuenciación diferentes (ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction kit, ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit versiones 1, 2 y 3). Se adaptaron los protocolos propuestos por la casa comercial a los requerimientos del presente estudio hasta conseguir el óptimo relativo a la muestra estudiada. La secuenciación automática fue llevada a cabo por el personal de secuenciación de los Servicios Científicotécnicos (SCT) de la Universidad de Barcelona inicialmente con el secuenciador ABI PRISM 377 y posteriormente con el modelo ABI PRISM 3700.

### 12.1. Kits de secuenciación

Los kits de secuenciación contienen diferentes reactivos premezclados en un único tubo listos para su uso (**Tabla ME18**). Todas las versiones empleadas utilizaban para la amplificación la enzima Ampli *Taq* DNA Polimerasa FS, una variante de la polimerasa de Termus aquaticus que presenta dos mutaciones puntuales en su centro activo y en su dominio amino terminal. La primera de ellas disminuye la discriminación de la enzima por los dideoxinucleótidos, lo que permite trabajar con concentraciones pequeñas de terminadores y simplifica la eliminación de aquellos no incorporados. La segunda mutación provoca el reemplazamiento de un residuo de Glicina por un Aspartato que se traduce en la eliminación total de la actividad nucleásica 5'-3'. La Ampli *Taq* DNA Polimerasa FS ha sido formulada junto con una pirofosfatasa inorgánica termoestable que elimina los residuos de fósforo inorgánico generados en la reacción de extensión. Un acúmulo de estos residuos puede provocar una reversión de la reacción de extensión, en la cual un nucleósido monofosfato es eliminado de la cadena en elongación y fosforilado para formar el correspondiente nucleósido trifosfato (pirofosforólisis).

La concentración de los dideoxinucleótidos marcados y de los nucleótidos está optimizada para dar una distribución de la señal equilibrada hasta más de 700 bases. En todos los formatos del kit se emplea dITP en lugar de dGTP para minimizar fenómenos

de compresión de bases. Adicionalmente, en las tres versiones de BigDye Terminator, el dTTP ha sido sustituido por dUTP para tratar de mejorar la incorporación de los terminadores de Timina.

**Tabla ME18.** Composición de los kits de secuencia utilizados (AbiPrism™ Dye Terminator Cycle Sequencing Kit y AbiPrism™ BigDye Terminator Cycle Sequencing Ready Reaction Kit). (\*) El kit AbiPrism™ BigDye Terminator Cycle Sequencing Ready Reaction Kit utiliza dUTP en lugar de dTTP.

Componentes de los kits	Composición
<b>Terminadores</b>	A-Dye Terminator, T-Dye Terminator C-Dye Terminator, G- Dye Terminator
<b>dNTPs</b>	dATP, dTTP/dUTP*, dCTP, dITP
<b>Tampón 5x</b>	Tris-HCl pH 9.0, MgCl <sub>2</sub>
<b>AmpliTaQ DNA Polimerasa, FS + Pirofosfatasa termoestable</b>	

Los dos kits utilizados -ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction kit y ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit- difieren en el tipo de marcaje de los ddNTPs (**Tabla ME19**). Los “BigDye terminators” incorporan un donador de fluoresceína (6-carboxifluoresceína) a cada aceptor de dRhodamina (ROSENBLUM *et al.* 1997). El donador está optimizado para absorber la energía de excitación del láser de argón, de manera que el máximo de excitación de cada *dye* corresponde al donador de fluoresceína mientras que el espectro de emisión procede del aceptor de dRhodamina descrito en la **Tabla ME19**. Mediante este sistema se consigue una señal entre 2 y 3 veces mayor que la proporcionada por los terminadores de rodamina. El espectro de emisión de los BigDye terminators es más estrecho que el de los terminadores de rodamina, lo que se traduce en una disminución del solapamiento entre el espectro de emisión de cada uno de los dyes. Este tipo de terminadores proporcionan una lectura correcta más allá de la base 700, por lo que resultan especialmente adecuados para la secuenciación de fragmentos largos de DNA. El aumento de la señal de emisión hace posible la secuenciación a partir de una concentración de DNA molde muy pequeña. En las sucesivas versiones de BigDye se ha variado la proporción de DdNTPs respecto a dNTPs, provocando un aumento progresivo de la intensidad de la señal. Esta mejora queda patente sobre todo en la secuenciación de fragmentos largos de DNA.

**Tabla ME19.** Marcaje de cada tipo de terminador y máximo de emisión asociado en nanómetros (nm).

DdNTPs	Marca asociada					
	Rhodamine Dye Terminators (AbiPrism™ Dye Terminator Cycle Sequencing Kit)		dRhodamine Dye Terminators (AbiPrism™ dRhodamine Terminator Cycle Sequencing Ready Reaction Kit)		BigDye Terminators (AbiPrism™ BigDye Terminator Cycle Sequencing Ready Reaction Kit)	
<b>DdA</b>	(R6G)	560nm	Dicloro(R6G)	565nm	Dicloro (R6G)	565nm
<b>DdC</b>	(TAMRA)	580nm	Dicloro(ROX)	595nm	Dicloro (ROX)	595nm
<b>DdG</b>	(R110)	528nm	Dicloro(R110)	533nm	Dicloro (R110)	533nm
<b>DdT</b>	(ROX)	610nm	Dicloro(TAMRA)	620nm	Dicloro (TAMRA)	620nm

## 12.2. Cuantificación del DNA molde para las reacciones de secuenciación

La cantidad de DNA molde es crítica para el correcto desarrollo de la reacción de secuenciación, puesto que tanto el exceso como la escasez de DNA afectan a la calidad de la secuenciación final obtenida.

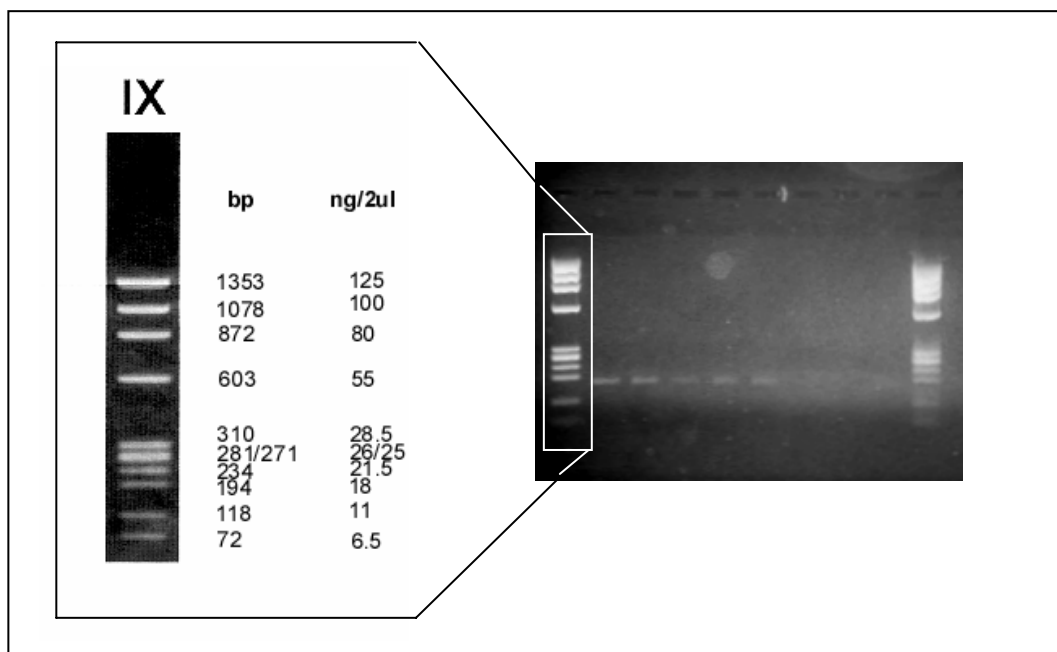
En la **Tabla ME20** se resumen las cantidades mínimas de DNA recomendadas para los dos tipos de química utilizada durante el desarrollo de este trabajo.

**Tabla ME20.** Cantidades de DNA molde recomendadas para la secuenciación automática con Rhodamine Dye Terminators (AbiPrism™ Dye Terminator Cycle Sequencing Kit) y BigDye Terminators (AbiPrism™ BigDye Terminator Cycle Sequencing Ready Reaction Kit)

DNA molde	Química de secuenciación	
	Rhodamine Dye Terminators (AbiPrism™ Dye Terminator Cycle Sequencing Kit)	BigDye Terminators (AbiPrism™ BigDye Terminator Cycle Sequencing Ready Reaction Kit)
Producto de PCR		
100-200pb	1-3ng	1-3ng
200-500pb	3-10ng	3-10ng
500-1000pb	5-20ng	5-20ng
1000-2000pb	10-40ng	10-40ng
>2000pb	40-100ng	40-100ng
Cadena única	100-250ng	50-100ng
Doble cadena	200-500ng	200-500ng
Cósmido, BAC	0.5-2ng	0.5-1ng
DNA genómico bacteriano	No recomendado	2-3µg

La cuantificación del producto de *PCR*, de 170pb (primer segmento) y 151pb (segundo segmento) se realizó comparando en un gel al 1-2% la intensidad de banda de cada producto purificado con la observada en las diferentes bandas de un marcador de peso molecular de concentración conocida. Para ello se fabrica un gel a la concentración deseada de la manera descrita y se cargan entre 2 y 4 µl de cada muestra. En el primer y último carril del gel se carga

una cantidad pequeña (entre 0'5 y 2 $\mu$ l) del marcador (Marcador IX de la casa Boëhringer-Manheim). Se deja correr a 70V durante aproximadamente tres cuartos de hora, transcurridos los cuales se examina en un transiluminador con luz UV y se fotografía con una cámara Polaroid o con un sistema de análisis de imagen. Se identifica la banda del marcador cuya intensidad corresponda con la de la muestra. La cantidad de DNA de la muestra será la misma que la de la banda en cuestión. En la **Figura FME10** se muestra un ejemplo de gel de cuantificación y se detalla la cantidad de DNA presente en 2 $\mu$ l del marcador



**Figura FME10.** Cuantificación del DNA purificado.

### 12.3. Reacción de secuencia

Las reacciones de secuenciación son asimétricas, es decir que en cada muestra se secuencia cada cadena por separado. La reacción se prepara en el interior de una campana biológica de flujo laminar siguiendo las condiciones de esterilidad descritas.

No existen unas condiciones universales para la secuenciación. Éstas dependen de diversos factores tales como el tipo de muestra, la química empleada (kit de secuenciación), la pureza y cantidad del DNA molde, las condiciones de amplificación y el protocolo de eliminación de terminadores. Durante el desarrollo de la presente tesis

doctoral fueron ensayadas diferentes concentraciones de reactivos en combinación con diferentes cebadores, condiciones de secuenciación y métodos de purificación de DNA y eliminación de terminadores hasta obtener la calidad de secuencia deseada. En la **Tabla ME21** se resumen las condiciones ensayadas.

**Tabla ME21.** Condiciones de secuenciación ensayadas en el presente trabajo.

Dye Terminators		BigDye Terminators				
Tipo de cebador	L16158- L16283/ H16394	Vs 1		Vs 1 y 2		Vs 3
		L16158- L16283/ H16394	L16158- L16283/ H16394	L16125 H16259 L16257 H16370	L16125, H16259, L16257, H16370	
Kit	4µl	4µl	1µl	3µl	3µl	0.5µl/1µl
DNA molde	1µl/2µl	1µl/2µl	1µl/2µl	6µl	6µl	6µl
Cebadores	1µl	1µl	1µl	1µl	1µl	1µl
H <sub>2</sub> O	4µl/3µl	4µl/3µl	7µl/8µl	-	-	2.5µl/2µl
Programa	L16158- L16283: SEC1 H16394: SEC2	L16158- L16283: SEC1 H16394: SEC2mod	L16158- L16283: SEC1 H16394: SEC3	L16125, H16259, L16257, H16370: SEC4		
Eliminación Terminadores	Acetato sódico-etanol con precipitación en frío	MgCl <sub>2</sub> -Etanol	Acetato sódico-etanol	Columnas qiagen	Columnas Etanol	Etanol

Las reacciones de secuencia se prepararon en el interior de una campana de flujo laminar como se describe en el apartado 9.2.2.2.

#### 12.4. Programas de secuenciación

Las amplificación de las reacciones de secuencia se realizó en dos modelos de termociclador: Perkin Elmer TC1 y Eppendorf modelo *Mastercycler personal*. Los programas de secuenciación se resumen en la **Tabla ME22**.



**Tabla ME22.** Programas de secuenciación empleados en el presente trabajo.

<b>Cebadores</b>	<b>Nombre programa</b>	<b>Programa secuenciación</b>
L16158, L16283	SEC1	96°C 30seg 40°C 15seg 25 ciclos 60°C 4min
H16394	SEC2	96°C 30seg 50°C 15seg 25 ciclos 60°C 4min
H16394	SEC2mod	96°C 30seg 50°C 15seg 30 ciclos 60°C 4min
H16394	SEC3	96°C 30seg 45°C 15seg 25 ciclos 60°C 4min
L16125, H16259, L16257, H16370	SEC4	96°C 30seg 50°C 15seg 25 ciclos 60°C 4min

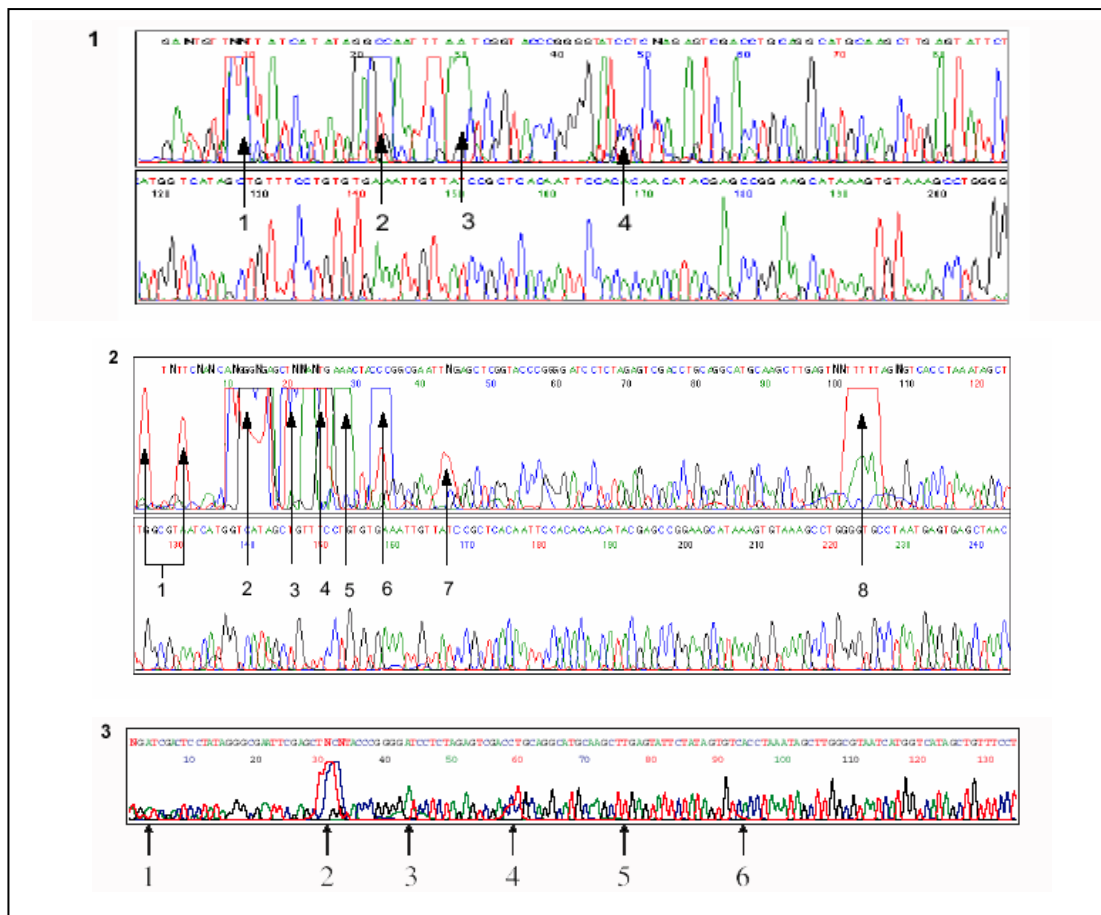
## 13. ELIMINACIÓN DE LOS TERMINADORES DE REACCIÓN NO INCORPORADOS

### 13.1. Localización de los terminadores de reacción no incorporados

Una baja eficiencia de la reacción de secuencia o una incorrecta limpieza del producto de precipitación provocan el acúmulo de los terminadores marcados que no han sido incorporados al producto de extensión. En el secuenciador, éstos migran junto con los fragmentos terminados y emiten picos de luz que pueden impedir la lectura de algunas porciones del cromatograma.

La localización de estos terminadores no incorporados depende de la naturaleza de su marcaje (tipo de química empleada) y del secuenciador empleado para su análisis. En la **Figura FME11** se muestra la situación de diferentes tipos de terminadores en el cromatograma. Como puede apreciarse en la figura, los *dye terminators* (rodamina) no incorporados aparecen generalmente al inicio de la secuencia, afectando a la lectura de las primeras 50 bases (**FME11, 1**). Los *BigDye terminators* (versiones 1 y 2) no incorporados pueden aparecer en 8 localizaciones diferentes a lo largo del cromatograma (**FME11, 2**). Los terminadores no incorporados que se sitúan en la posición 1 no interfieren en la lectura de las bases. Sin embargo la presencia de los mismos en posiciones de 2 a 7 y en posición 8 sí afectan a la calidad de la secuencia, interfiriendo en la lectura de las 40 primeras bases y entre las posiciones 90 y 100 respectivamente. Un protocolo adecuado de precipitación consigue eliminar completamente los terminadores presentes en las posiciones 1, 5, 6, 7 y 8. Aquellos presentes en las posiciones 2, 3, 4 siguen apareciendo incluso tras una correcta precipitación. A pesar de ello la presencia de terminadores en estas posiciones puede afectar únicamente a la lectura de las 10 primeras bases de la secuencia. A este efecto la versión 3 de BigDye terminator resulta mucho más problemática (**Figura FME11, 3**). Mediante el empleo de un protocolo adecuado de precipitación se consiguen eliminar los terminadores de reacción no incorporados de 4 de las 6 posibles localizaciones (1, 4, 5 y 6). Como en anteriores versiones, se mantienen los terminadores presentes en las posiciones 2 y 3. Estas posiciones corresponden a las bases 25-40 del cromatograma en un secuenciador modelo ABI PRISM 3700. En este último caso por lo tanto los

terminadores de reacción no incorporados en estas posiciones afectan a la lectura de la secuencia.



**Figura FME11.** Situación en los cromatogramas de los terminadores de reacción no incorporados. (1) Rhodamine Dye Terminators (AbiPrism™ Dye Terminator Cycle Sequencing Kit). (2) BigDye Terminators (AbiPrism™ BigDye Terminator Cycle Sequencing Ready Reaction Kit) versiones 1 y 2. (3) BigDye Terminators (AbiPrism™ BigDye Terminator Cycle Sequencing Ready Reaction Kit) versión 3.

### **13.2. Protocolos para la eliminación de los terminadores no incorporados**

Hay una gran variedad de protocolos que resultan eficientes para eliminar los terminadores de reacción no incorporados. La elección de uno u otro protocolo dependerá esencialmente del tipo de marcaje de la molécula terminadora (tipo de química empleada), de las características del segmento analizado y del tipo de secuenciador empleado para el análisis. La mayoría de estos protocolos se basan en la precipitación de los productos de extensión mediante algún solvente químico como el etanol, sólo o combinado con algunas sales, y la posterior eliminación de aquellos terminadores no incorporados mediante lavados con etanol al 70%. Estos protocolos presentan una eficiencia óptima de eliminación de los terminadores de reacción no incorporados, sin embargo algunos de ellos no son recomendables en ciertos casos. Por ejemplo para la secuenciación en un equipo ABI PRISM 3700 DNA Analyzer no resultan adecuados aquellos protocolos que precisan de sales para precipitar el producto de extensión, tales como la precipitación con etanol/acetato sódico o etanol/cloruro de magnesio. Las sales pueden obturar los capilares de inyección de muestra del secuenciador disminuyendo su vida útil.

Alternativamente a los protocolos de precipitación pueden emplearse columnas comerciales rellenas de una solución de resina para eliminar los terminadores de reacción no incorporados. La matriz de resina retiene estos terminadores y deja pasar los fragmentos de extensión, que se recogen en un tubo colector situado en su base. La ventaja de este método respecto a los métodos de precipitación es que reduce enormemente el tiempo de manipulación.

En la **Tabla ME23** se detalla la compatibilidad de los protocolos más comunes de eliminación de los dos tipos de terminadores de reacción no incorporados con los sistemas de análisis empleados en la presente tesis doctoral. Se describen a continuación los protocolos de eliminación de terminadores de reacción no incorporados empleados en el presente trabajo.

**Tabla ME23.** Protocolos más comunes de eliminación de los terminadores de reacción no incorporados y su compatibilidad con el tipo de secuenciador empleado para el análisis de secuencias (377 o 3700). ✓ recomendado, × no recomendado, - puede emplearse.

Protocolo de eliminación	Dye Terminators		BigDye Terminators	
	377	3700	377	3700
Isopropanol (no descrito)	-	-	✓	✓
Etanol/Acetato sódico	✓	×	✓	×
Etanol/Cloruro de Magnesio	✓	×	×	×
Etanol	×	×	✓	✓
Columnas	✓	✓	✓	✓

### 13.2.1. Precipitación con etanol/acetato sódico

La precipitación con etanol/acetato sódico proporciona una señal de secuencia clara y consistente, minimizando la presencia de terminadores no incorporados. Sin embargo puede provocar la pérdida de fragmentos de bajo peso molecular.

La metodología es la siguiente:

- Se prepara un tubo de microcentrifuga de 1.5 ml por cada reacción (10µl) y se añaden diferentes cantidades de acetato sódico, etanol 95% y agua dependiendo del tipo de kit empleado:

	Dye Terminator/ BigDyeTerminator vs 1,2		BigDyeTerminator vs 3	
	Volumen	Concentración final	Volumen	Concentración final
Acetato sódico 3M pH 4.6	1µl	83mM	1.5µl	90mM
Etanol 95%	25 µl	65.97%	31.25µl	59.375%
H <sub>2</sub> O	-		7.25µl	

En lugar de acetato sódico 3M pH 4.6 puede emplearse también acetato sódico 3M pH 5.2 o acetato potásico 3M pH 5.6. La combinación de estas concentraciones minimiza la incorporación de terminadores y la pérdida de aquellos fragmentos de extensión más pequeños manteniendo unos niveles aceptables de señal. Una disminución de la concentración de etanol resulta en la pérdida de fragmentos cortos de extensión mientras que un aumento favorece la coprecipitación de los terminadores no incorporados. El mismo efecto se aprecia aumentando la concentración de acetato sódico.

- Se añaden a la solución de etanol las reacciones de secuencia tras retirar el aceite mineral y se mezcla en un agitador vórtex. El aceite puede eliminarse directamente del eppendorf con ayuda de una pipeta o bien cargando la muestra sobre parafilm y aspirando la fase acuosa.

- Existen dos opciones para precipitar los productos de extensión:

- a. Precipitación en frío manteniendo los eppendorf en hielo (4°C) durante 10 minutos (AbiPrism™ Dye Terminator Cycle Sequencing Kit).
- b. Precipitación a temperatura ambiente durante 15 minutos (AbiPrism™ BigDye Terminator Cycle Sequencing Ready Reaction Kit).

Mediante una precipitación en frío con un elevado tiempo de precipitación la eficiencia de recuperación de fragmentos cortos es mayor. Sin embargo este método favorece también la co-precipitación junto con el producto de extensión de los terminadores no incorporados en la reacción de secuencia. La precipitación a temperatura ambiente minimiza la precipitación de los terminadores de reacción no incorporados pero la eficiencia de recuperación de fragmentos de DNA cortos es también menor. En el presente trabajo se emplearon ambos procedimientos con un resultado similar.

- Se centrifugan los tubos en una microcentrífuga a máxima velocidad (12000g) durante 20 minutos. Es conveniente centrifugar los tubos manteniendo siempre la misma orientación. Tras la centrifugación el DNA forma un *pellet* que queda adherido a la pared del tubo.

- Se aspira cuidadosamente todo el sobrenadante con una pipeta Pasteur o capilar de vidrio procurando no tocar el *pellet*. Si los tubos están orientados de igual manera, tras la centrifugación el *pellet* quedará en el extremo del tubo opuesto al eje de la centrífuga.

- Se resuspende el *pellet* añadiendo 250 µl de 70% etanol.

Los dos pasos siguientes no son necesarios para AbiPrism™ Dye Terminator Cycle Sequencing Kit, pero sí para AbiPrism™ BigDye Terminator Cycle Sequencing Ready Reaction Kit.

- Se agita la solución en un vórtex.

- Se centrifuga en una microcentrífuga a máxima velocidad (12000g) durante 5 minutos procurando de nuevo mantener para todos ellos la misma orientación.

- Se aspira cuidadosamente toda la solución de alcohol con una pipeta Pasteur o capilar de vidrio procurando no tocar el *pellet*.

- Para secar el *pellet*, se dejan los eppendorfs a temperatura ambiente con el tapón abierto el tiempo que sea necesario.

### 13.2.2. Precipitación con etanol/cloruro de magnesio

Este protocolo no resulta recomendable para la precipitación de fragmentos de extensión marcados con BigDye Terminators debido a que favorece la precipitación de estos terminadores.

La metodología es la siguiente:

- Se transfieren las reacciones de secuencia a tubos de microcentrífuga de 1.5 ml tras retirar el aceite mineral. Aquellas reacciones de 10µl se llevan a un volumen de 20µl añadiendo agua estéril para inyectables (Braun).
- Se añade a cada una de las reacciones (20µl) lo siguiente:

	Dye Terminator	
	Volumen	Concentración final
MgCl <sub>2</sub> 2mM	20µl	0.42mM
Etanol 95%	55µl	55%

- Se cierran los tubos y se mezcla la solución en un vórtex.
- Se dejan a temperatura ambiente durante 15 minutos para precipitar los productos de extensión. Tiempos de precipitación inferiores producen la pérdida de aquellos fragmentos de extensión más pequeños, mientras que tiempos largos (superiores a 24 horas) aumentan la precipitación de terminadores no incorporados.
- Se centrifugan los tubos en una microcentrífuga a máxima velocidad (12000g) durante 20 minutos. Al colocar los tubos en la centrífuga es recomendable orientarlos todos de la misma manera.
- Se aspira cuidadosamente todo el sobrenadante con una pipeta Pasteur o capilar de vidrio procurando no tocar el *pellet*. Si los tubos están orientados de igual manera tras la centrifugación éste quedará en el extremo opuesto al eje de la centrífuga.
- Se resuspende el *pellet* añadiendo 250 µl de 70% etanol.
- Se agita la solución en un vórtex.

- Se centrifugan los tubos en una microcentrífuga a máxima velocidad (12000g) durante 5 minutos procurando de nuevo mantener para todos ellos la misma orientación.
- Se aspira cuidadosamente toda la solución de alcohol con una pipeta Pasteur o capilar de vidrio procurando no tocar el *pellet*.
- Para secar el *pellet*, se dejan los eppendorfs a temperatura ambiente con el tapón abierto el tiempo que sea necesario.

### 13.2.3. Precipitación con etanol

Este tipo de precipitación resulta especialmente recomendable para la eliminación de BigDye Terminators. La concentración en reacción de etanol óptima es de 60%  $\pm$  3%. Por encima de este valor se produce un aumento en la cantidad de terminadores residuales, mientras que por debajo disminuye la eficiencia de precipitación de los fragmentos de extensión, especialmente de aquellos más cortos. Por ello resulta aconsejable trabajar con una solución *stock* de etanol al 95% en lugar de con etanol absoluto, ya que éste último absorbe fácilmente agua de la atmósfera provocando una disminución paulatina de su concentración. Trabajando con las concentraciones estipuladas para este método pueden quedar residuos de terminadores no incorporados a lo largo de las 40 primeras bases de la secuencia, aunque suelen ser mínimos. Análogamente puede producirse una pérdida de los fragmentos de extensión más pequeños.

La metodología es la siguiente:

- Se transfieren las reacciones de secuencia a tubos de microcentrífuga de 1'5 ml tras retirar el aceite mineral. Aquellas reacciones de 10 $\mu$ l se llevan a un volumen de 20 $\mu$ l añadiendo agua mili-Q.
- Se añade a cada una de las reacciones (20 $\mu$ l) lo siguiente:

	Dye Terminator/ BigDye Terminator vs 1,2,3	
	Volumen	Concentración final
Etanol 95%	64 $\mu$ l	60.8%
H <sub>2</sub> O	16 $\mu$ l	-

- Se cierran los tubos y se mezcla la solución en un vórtex.



- Se dejan a temperatura ambiente durante 15 minutos para precipitar los productos de extensión. Tiempos de precipitación inferiores producen la pérdida de aquellos fragmentos de extensión más pequeños, mientras que tiempos largos (superiores a 24 horas) aumentan la precipitación de terminadores no incorporados.
- Se centrifugan los tubos en una microcentrífuga a máxima velocidad (12000g) durante 20 minutos procurando orientarlos todos de la misma manera.
- Se aspira cuidadosamente todo el sobrenadante con una pipeta Pasteur o capilar de vidrio procurando no tocar el *pellet*. Si los tubos están orientados de igual manera tras la centrifugación éste quedará en el extremo opuesto del eje de la centrífuga.
- Se resuspende el *pellet* añadiendo 250  $\mu$ l de 70% etanol.
- Se agita la solución en un vórtex.
- Se centrifugan los tubos en una microcentrífuga a máxima velocidad (12000g) durante 5 minutos procurando de nuevo mantener para todos ellos la misma orientación.
- Se aspira cuidadosamente toda la solución de alcohol con una pipeta Pasteur o capilar de vidrio procurando no tocar el *pellet*.
- Para secar el *pellet*, se dejan los eppendorfs a temperatura ambiente con el tapón abierto el tiempo que sea necesario.

#### **13.2.4. Precipitación con columnas (DyeEx Spin columns, Qiagen®)**

- Para resuspender la resina se agitan las columnas en un vórtex.
- Se desenroscan levemente los tapones para evitar que al centrifugar se haga el vacío en el interior y se colocan en el interior del tubo colector (suministrado).
- Para compactar la matriz de resina se centrifugan durante 3 minutos a 750g.
- Se transfiere la columna a un nuevo tubo de centrifuga de 1.5ml y se aplica cuidadosamente la reacción de secuencia sobre el centro de la columna con una micropipeta procurando no tocar la matriz de resina ni los bordes del tubo. Para aquellas reacciones de secuencia de 10 $\mu$ l se recomienda ajustar el volumen a 20 $\mu$ l añadiendo agua mili-Q. Resulta también aconsejable eliminar el aceite mineral de la reacción antes de su aplicación sobre la columna debido a que el aceite puede formar una película sobre la superficie de la resina impidiendo su correcta absorción.
- Se enroscan levemente los tapones de las columnas y se centrifugan durante 3 minutos a 750g.

- Se extraen las columnas de los tubos y se desechan. La solución acuosa eluida remanente en el tubo contiene el DNA libre de terminadores.
- Para secar el DNA se dejan los tubos abiertos a temperatura ambiente durante toda la noche o sobre un bloque térmico a 60°C durante 2 o 3 horas.

## 14. CRITERIOS DE AUTENTICIDAD

Para autenticar las secuencias de DNA mitocondrial obtenidas se emplearon un conjunto de criterios. La mayoría de ellos coinciden con aquellos propuestos en la literatura, que aparecen expuestos detalladamente en el apartado 3.7 del capítulo “Introducción”. Otros criterios, como la comparación de la secuencia obtenida con otras secuencias de su misma serie de amplificación y reamplificación o la eliminación de aquellas secuencias con más de una posición ambigua han sido introducidos y aplicados por primera vez en la presente tesis doctoral.

### 14.1. Infraestructura

#### *14.1.1. Análisis en un laboratorio exclusivo de DNA antiguo*

El proceso de limpieza, trituración y extracción del DNA de la muestra se realizó en el laboratorio de DNA antiguo de la Sección de Antropología del Departamento de Biología Animal de la Facultad de Biología de la Universidad de Barcelona.

La amplificación y procesado *post-PCR* de la mayor parte de las muestras se realizó también en este laboratorio. Sin embargo la amplificación de mtDNA de las muestras del yacimiento de Mari y de parte de las muestras del yacimiento de Tell Halula se llevó a cabo en el laboratorio de Genética Forense y Genética de Poblaciones del Departamento de Toxicología y Legislación Sanitaria de la Facultad de Medicina de la Universidad Complutense de Madrid.

En ambos laboratorios las áreas en las que se realizaron los análisis eran de uso exclusivo de DNA antiguo.

#### *14.1.2. Separación de las áreas de trabajo pre y post-PCR*

En ninguno de los dos laboratorios en los que se realizó el procesado experimental de las muestras existen áreas separadas para la limpieza, trituración y extracción del material genético y la amplificación por *PCR*. De esta forma únicamente

se puede hablar de *separación física* en aquellos casos en los que la limpieza, trituración y extracción de la muestra fue realizada en la Universidad de Barcelona y la amplificación en la Universidad Complutense de Madrid. Esto es válido para las muestras procedentes del yacimiento de Mari y buena parte de las de Tell Halula.

En los restantes casos se utilizó una *separación temporal* de las fases *pre-PCR* (limpieza, trituración y extracción) y *PCR*, concentrando las extracciones en un período determinado y las amplificaciones en otro.

Los procesos de limpieza, trituración, extracción y preparación de las reacciones de amplificación se efectuaron en una campana de flujo laminar situada en una habitación separada mientras que la amplificación en el termociclador y todos los procesos *post-PCR* (electroforesis en gel de agarosa, purificación del DNA a partir de agarosa...) se realizaron en una habitación separada. En este caso sí se puede hablar de *separación física* entre las fases *pre-PCR* y *post-PCR*. Para evitar el flujo de DNA entre ambos compartimentos se llevaron a cabo un conjunto de precauciones adicionales, que incluyen el uso de un juego de pipetas diferentes en cada fase (extracción, amplificación y tratamiento *post-PCR*), el uso de guantes y material de plástico estéril de un solo uso en cada proceso y el empleo de batas de laboratorio diferentes. Las muestras se guardaron en congelador a  $-20^{\circ}\text{C}$  dentro de la habitación aislada utilizando diferentes compartimentos para el hueso sin procesar, el hueso pulverizado, los extractos de DNA sin amplificar, los productos de amplificación y los productos de *PCR* purificados.

#### **14.1.3. Equipamiento de los laboratorios de DNA antiguo.**

El laboratorio de DNA antiguo de la Universidad de Barcelona cuenta con un cuarto aislado que contiene la arenadora (Sand-Blaster Base 1 Plus), el molino refrigerado (Freezer Mill), una campana de flujo laminar vertical Telstar Micro-V, un congelador y una centrífuga de tubos falcon.

El laboratorio de Genética Forense y Genética de Poblaciones de la Universidad Complutense de Madrid cuenta también con una habitación aislada que contiene un congelador en el que se guardan las muestras y los reactivos necesarios para su análisis y una campana de flujo laminar vertical Heraeus modelo Gemini. Para la esterilización

del interior de este cuarto se ha dispuesto en una de sus paredes un potente fluorescente de luz ultravioleta que se enciende desde el exterior.

## **14.2. Metodología**

### ***14.2.1. Empleo de instrumental y equipamiento exclusivo para el análisis de DNA antiguo.***

#### *14.2.1.1. Limpieza de las áreas de trabajo*

Las campanas de flujo laminar se limpiaron antes de comenzar cualquier proceso. La limpieza se realizó humedeciendo un trozo de papel con lejía (hipoclorito sódico) al 70% en el caso de la campana modelo Telstar Micro-V (Universidad de Barcelona) y con etanol al 70% en la campana modelo Heraeus Gemini (Universidad Complutense de Madrid) atendiendo a las recomendaciones de cada fabricante.

Tras la limpieza se depositó en el interior de la campana el material necesario para la realización de cada fase experimental: pipetas, cajas de puntas, gradilla, un par de guantes, etc. En la preparación de las reacciones de amplificación se dispusieron sobre la gradilla los eppendorfs necesarios con el tapón abierto. Una vez así dispuesto el material fue irradiado con el fluorescente de luz ultravioleta de la propia campana. En aquellas amplificaciones llevadas a cabo en la Universidad Complutense se irradió también la habitación con luz UVA durante 15-20 minutos.

#### *14.2.1.2. Material empleado*

##### 14.2.1.2.1. Material de un solo uso

El material de un solo uso utilizado en el presente trabajo consiste en:

- Material de plástico:
  - Tubos falcon de polipropileno de 50ml estériles en bolsitas de plástico individuales (Lab-Center).

- Puntas de micropipeta con filtro resistente a aerosoles estériles en rack de 96 puntas o en bolsitas individuales (Iberlabo, Deltalab).
  - Tubos eppendorf de 1.5ml y de *PCR* de 0.5ml (Iberlabo, Lab-Center). Estos tubos se guardan en el interior de tarros de vidrio, se autoclavan durante 15 minutos a 210°C y posteriormente se secan en una estufa a 60°C.
  - Guantes de látex o de cloruro de polivinilo (Rubilabor, Lab-Center).
- Material de vidrio:
    - Pipetas Pasteur de vidrio. Se lavan tal y como se describe en el apartado 7.1 de este capítulo, se secan en estufa a 60°C, se envuelven en papel de aluminio y se autoclavan 15min a 210°C de la misma forma que los tubos eppendorf.

#### 14.2.1.2.2. Material no desechable

El material no desechable utilizado incluye:

- Pipetas de vidrio graduadas de 10 y 5ml.
- Viales y tapones de trituración.
- Pinzas metálicas.
- Dispensador de volumen de pipetas de vidrio de 10 y 5 ml.
- Micropipetas automáticas (GILSON).
- Gradillas para eppendorf y tubos falcon.
- Recipiente para desechar las puntas.
- Peritas de goma para las pipetas *Pasteur*.

El material de plástico no autoclavable (micropipetas automáticas, gradillas, recipiente para desechar las puntas) se limpió externamente con papel secante impregnado en una dilución de lejía al 70% y se introdujo en la campana de seguridad biológica previamente a su uso. Una vez dispuesto en su interior se irradió con luz ultravioleta durante 15 minutos.

El material de vidrio, metal y plástico autoclavable se limpió y autoclavó antes de cada uso. La limpieza se realizó de la siguiente manera:

- Un primer lavado con detergente, preferiblemente aniónico tras el cual se aclaró con agua del grifo.
- Las piezas de vidrio y metal se limpiaron externamente con un trozo de papel secante impregnado primero en etanol 70% y luego en lejía 70%. Tras cada limpieza se secaron bien para evitar que quedaran residuos de ambos compuestos. Tras la última limpieza con lejía se aclararon con agua del grifo. Las piezas de plástico, como los viales de trituración, no pueden limpiarse de esta manera porque se estropean.
- Se enjuagaron a continuación abundantemente con con agua mili-Q y se secaron con papel.
- Se enjuagaron abundantemente con agua estéril para inyectables (Braun).
- Se secaron en estufa a 60°C.

Tras la limpieza se autoclavó el material durante 15 minutos a 210°C y se depositó posteriormente en una cámara con luz UV, donde permaneció durante al menos 3 horas.

#### 14.2.1.3. Soluciones

La mayor parte de los reactivos y soluciones *stock* empleados en aquellas fases del proceso susceptibles de contaminación se compraron ya preparados de fábrica con garantías de esterilidad.

Este es el caso de los reactivos de *PCR* (Tampón de *PCR*, Cloruro de Magnesio, dNTPs, cebadores, *Taq* polimerasa) y de algunos reactivos utilizados en las fases de limpieza y extracción (EDTA 0.5M, Tris-HCl 1M, Fenol, Cloroformo). El EDTA 0.5M utilizado en el lavado del polvo de hueso se sirve preparado en botellas de 500ml. Cada botella se abrió antes de comenzar un grupo de extracción y el contenido sobrante se alicuotó en tubos falcon de 50ml para su uso posterior. La cantidad utilizada en cada lavado se añadió directamente de la botella o tubo falcon para no introducir pipetas en su interior. El fenol y el cloroformo vienen preparados en botellas de vidrio de 1l. Las cantidades necesarias para la extracción se añadieron mediante pipetas de vidrio de

10ml autoclavadas y esterilizadas con luz UV tal y como se ha descrito en la sección anterior. Los reactivos de *PCR* se alicuotaron en cantidades pequeñas, normalmente para cuatro reacciones, en el interior de la campana de seguridad biológica. Se numeró cada serie de alícuotas para facilitar la detección de las alícuotas contaminadas. Estas alícuotas se guardaron en congelador a -20°C en una caja separada de las soluciones *stock*.

Para la preparación de soluciones *stock* en el laboratorio (SDS 5%) o de *stocks* diluidos a partir de uno concentrado (EDTA 50mM, Tris-HCl 100mM) se utilizó siempre agua estéril para inyectables (Braun). Las soluciones así preparadas se filtraron con filtros MILLEX-GS de 0.22µm, se irradiaron con luz UV y se alicuotaron en tubos falcon de 50ml estériles de fábrica. En aquellos casos en los que se preparó un *stock* de solución de lisis, éste se filtró e irradió con luz UV antes de añadir la proteinasa K. La proteinasa K se compró liofilizada y se diluyó en el interior de la campana de flujo laminar a una concentración de 20mg/ml con agua estéril para inyectables. A partir de esta solución *stock* se fabricaron alícuotas de 250µl que se guardaron en congelador a -20°C.

#### ***14.2.2. Eliminación de la capa superficial del hueso/diente previamente a la extracción del material genético.***

La capa superficial de los huesos o dientes empleados en el presente trabajo fue eliminada por abrasión mediante una arenadora (*Sandblaster*) modelo Base 1 Plus que inyecta óxido de aluminio a presión. Una vez limpios los dientes o huesos se irradiaron durante 15 minutos con luz UV en el interior de una campana de flujo laminar. Ambos procesos se describen detalladamente en el apartado 3 de este capítulo.

#### ***14.2.3. Análisis de cada espécimen por un único investigador.***

El análisis de cada muestra por un único investigador permite minimizar e identificar más fácilmente la contaminación introducida durante el procedimiento experimental. Se ha intentado que la autora de la presente tesis doctoral realizara todo el proceso experimental desde la limpieza superficial de la muestra hasta su amplificación por *PCR*. Este propósito se cumple en 140 de los 169 extractos de DNA



obtenidos y analizados. Sin embargo se analizaron también 44 extractos obtenidos anteriormente por otro investigador del laboratorio, todos aquellos procedentes de los yacimientos de Açıklı (8 extractos), Atxuri (13 extractos), Garai (11 extractos) y Zafarraya (2 extractos) y las muestras CA1, CA2, CA3 y CA4 del yacimiento de Caldeirao, así como los extractos TA1, TA2, y TA3 de Muge. Estos extractos fueron incorporados a esta tesis dada su importancia y la dificultad para obtener otras muestras de los mismos yacimientos. Las secuencias obtenidas a partir de estos extractos se compararon con las del investigador que realizó la extracción de su material genético y se descartaron aquellas coincidentes.

#### **14.2.4. Procesado de controles o blancos**

Durante todo el proceso de obtención de información genética a partir de la muestra de hueso o diente se llevaron a cabo una serie de controles para tratar de detectar la introducción de DNA contaminante.

##### *14.2.4.1. Controles negativos de extracción*

En cada grupo de extracción formado por siete muestras se incluyó un control o blanco de extracción con el objeto de tratar de detectar la contaminación introducida durante esta fase. Este control fue procesado de forma paralela a las muestras desde la fase de lavado con EDTA 0.5M hasta la concentración con Centriplus. Se obtuvo así un “extracto” acuoso que se amplificó posteriormente mediante *PCR*. La presencia de DNA en este extracto responde necesariamente a la introducción de DNA exógeno durante esta fase del proceso. Sin embargo la presencia de bandas de amplificación en los controles de extracción no siempre es indicativa de contaminación de extracción, tal y como detallaremos en los capítulos “Resultados” y “Discusión” de la presente tesis doctoral.

En los 21 grupos de extracción realizados se procesaron 23 de estos controles de extracción. Se introdujo un control por cada siete muestras en todos los casos excepto en el grupo de extracción número 3, en el que se incluyeron 3. Los diferentes controles se numeraron de forma consecutiva. Los diferentes grupos de extracción con sus

correspondientes controles se detallan en la **Tabla R1** del apartado 1 del capítulo “Resultados”.

#### 14.2.4.1.1. Amplificación conjunta de controles de extracción y muestras

En una fase inicial todos los controles excepto CE20 y CE21 se amplificaron junto con las muestras de su mismo grupo de extracción por lo menos una vez, y la mayor parte de ellos se reamplificaron junto con aquellas muestras que dieron resultado positivo en la primera amplificación. En la **Tabla ME24** se detallan el número de amplificaciones y reamplificaciones llevadas a cabo con cada control de extracción.

Para tratar de determinar el origen del DNA contaminante se secuenciaron tres controles de extracción (CE11, CE12 y CE13) que habían producido banda de amplificación en el primer y segundo *PCR* y dos que habían producido banda de amplificación únicamente en el segundo (CE1 y CE7).

**Tabla ME24.** Número de amplificaciones y reamplificaciones llevadas a cabo con cada control de extracción.

<b>Grupo extracción</b>	<b>Control extracción</b>	<b>Nº amplificaciones</b>	<b>Nº reamplificaciones</b>
1	CE1	13	13
2	CE2	8	8
3	CE3	3	3
3	CE4	3	3
3	CE5	3	3
4	CE6	11	9
5	CE7	8	8
6	CE8	4	4
7	CE9	1	0
8	CE11	3	1
9	CE12	3	1
10	CE13	3	1
11	CE14	2	0
12	CE15	2	0
13	CE16	1	0
14	CE17	1	0
15	CE18	1	0
16	CE19	1	0
17	CE20	0	0
18	CE21	0	0
19	CE22	1	0
20	CE23	2	1
21	CE24	2	1

#### 14.2.4.1.2. Amplificación de los controles de extracción aisladamente

Se amplificaron todos los controles de extracción en grupos de amplificación formados por seis de estos controles, tres blancos de amplificación y un control positivo de amplificación. Después de visualizar el resultado de la primera *PCR* en un gel de agarosa al 2%, se sometieron las muestras a una segunda amplificación.

#### *14.2.4.2. Controles negativos de amplificación*

En cada grupo de amplificación, formado por entre 10 y 12 reacciones, se incluyeron un mínimo de dos controles o blancos de amplificación. Estos controles permiten detectar la introducción de DNA exógeno a la muestra durante la *PCR*. El DNA contaminante puede proceder tanto del investigador como de anteriores amplicones que en forma de aerosoles permanecen en el ambiente o sobre las superficies de trabajo (contaminación por *carry-over*). No se admitieron los resultados procedentes de aquellas reacciones de amplificación con alguno de los controles de amplificación contaminado. En algunos casos se secuenciaron estos controles contaminados con el objeto de obtener información acerca de la fuente de origen de la contaminación.

### **14.3. Reproducibilidad**

#### *14.3.1. Replicación del experimento dentro del mismo laboratorio*

La replicación de los resultados obtenidos dentro del mismo laboratorio, a través de la amplificación de diferentes extractos de un mismo individuo o mediante la realización de diferentes amplificaciones de una misma muestra, permite detectar la contaminación del recinto y ofrece información acerca del estado de preservación del material genético en la muestra. Este criterio sin embargo está sujeto a la disponibilidad de la muestra, por lo que su aplicación no siempre resulta factible.

En la presente tesis doctoral únicamente se pudieron realizar amplificaciones independientes de 10 de las muestras estudiadas (**Tabla ME25**). De la misma manera se analizaron diferentes extractos de un mismo individuo en 6 casos (**Tabla ME26**). La

dificultad en la obtención de otras piezas dentales u óseas de un mismo individuo y el escaso volumen de extracto remanente tras la realización de los análisis hizo imposible la replicación de un número mayor de muestras.

**Tabla ME25.** Amplificaciones de diferentes extractos de un mismo individuo. Primer fragmento: posiciones 16126-16258. Segundo fragmento: posiciones 16258-16369.

Individuo	Muestra	Primer fragmento	Segundo fragmento
H7	TH5		X
H7	TH6		X
H70	2H11	X	X
H70	3H11	X	X
R65-4 II	TR9	X	X
R65-4 II	TR11	X	X
SK-R3	1DJ6		X
SK-R3	2DJ6		X
MA4	1MA4	X	X
MA4	2MA4	X	X
AH 95 2c-2	SK107-3	X	
AH 95 2c-2	SK107-3	X	

**Tabla ME26.** Amplificaciones de un mismo extracto de un mismo individuo. Primer fragmento: posiciones 16126-16258. Segundo fragmento: posiciones 16258-16369.

Individuo	Muestra	Primer fragmento	Segundo fragmento
	AB17	X	
	AB17	X	
	2TM4	X	
	2TM4	X	
	4NE		X
	4NE		X
	TH4		X
	TH4		X
	1H18		X
	1H18		X
	1H21		X
	1H21		X
	TR7		X
	TR7		X
	TR7		X
	DJ3	X	
	DJ3	X	
	2MA4	X	
	2MA4	X	
	1MA8	X	
	1MA8	X	

Aquellas secuencias procedentes de diferentes amplificaciones de un mismo extracto o de diferentes extractos de un mismo individuo que no proporcionaron un resultado idéntico no fueron consideradas en el análisis poblacional y estadístico.

#### 14.4. Interpretación de los resultados

##### 14.4.1. Sentido filogenético

Cuando se trabaja con poblaciones humanas el sentido filogenético de los resultados obtenidos resulta difícil de establecer, especialmente cuando el investigador pertenece a la misma población que las muestras analizadas. En este sentido conviene demostrar, por una parte, que los resultados obtenidos no son fruto de contaminaciones

exógenas y, por otra, que existe cierta relación entre los resultados obtenidos y lo que se sabe acerca de la población analizada.

*14.4.1.1. Comparación de las secuencias obtenidas con las del personal arqueológico y/o antropológico y con el personal de laboratorio*

La contaminación introducida por el personal de laboratorio puede ser fácilmente detectada comparando la secuencia obtenida con las secuencias de los investigadores. Para detectar la contaminación introducida por el equipo de arqueólogos y/o antropólogos que han manipulado la muestra previamente a su estudio genético puede y debe seguirse el mismo criterio. Sin embargo, en algunos casos resulta muy difícil e incluso imposible seguir la pista de este conjunto de personas, especialmente si la muestra ha estado depositada durante mucho tiempo en un museo.

En el presente estudio se compararon las secuencias obtenidas con las de tres investigadores (I1, I2 e I3) que manipularon las muestras en los dos laboratorios en los que se realizaron tareas de extracción, amplificación y secuenciación (Laboratorio de DNA antiguo, Sección de Antropología, Universitat de Barcelona y Laboratorio de Genética Forense y Genética de Poblaciones, Departamento de Toxicología y Legislación Sanitaria, Universidad Complutense de Madrid). Se compararon también las secuencias de un arqueólogo y un antropólogo (A1, A2) que participaron en la excavación y análisis antropológico de los restos de los yacimientos Neolíticos de Tell Halula, Tell Ramad, Dja'de (Siria) y Sant Pau del Camp (Barcelona).

Tanto el personal arqueológico/antropológico como los investigadores son europeos al igual que las muestras antiguas estudiadas, por lo que una coincidencia entre los haplotipos de unos y otros no tiene por qué significar contaminación necesariamente. Teniendo esto en cuenta únicamente se eliminaron aquellas secuencias en las que existían otras razones para pensar que pudiera haber tenido lugar dicha contaminación.

*14.4.1.2. Comparación de la secuencia obtenida con una base de datos de secuencias actuales de la misma región geográfica o de regiones próximas*

Las secuencias antiguas obtenidas se buscaron en una base de datos de secuencias actuales construida para la presente tesis doctoral. Puesto que mediante el diseño experimental propuesto se obtenían dos fragmentos parciales de la HVRI del mtDNA de aproximadamente 150pb cada uno, se buscó la coincidencia entre dichos fragmentos parciales y en aquellos casos en que pudieron obtenerse ambos, del haplotipo completo de 300pb.

Que las secuencias antiguas no estén descritas actualmente no implica su falta de validez. Podrían estar representadas de forma minoritaria en individuos no muestreados. Por otra parte la composición genética de las poblaciones puede haber cambiado con el tiempo y ciertos linajes pueden haberse perdido por deriva genética. Teniendo esto en cuenta, se procedió como en el criterio anterior y se eliminaron sólo aquellas secuencias no registradas en la base de datos de cuya autenticidad cabía dudar razonablemente.

***14.4.2. Comparación de las secuencias obtenidas con las de su mismo grupo de amplificación y reamplificación***

Este criterio tiene como objetivo detectar la contaminación por *carry over*. Este tipo de contaminación es la más difícil de detectar y pasa habitualmente inadvertida. Sin embargo la comparación de la secuencia sospechosa con el resto de secuencias obtenidas del mismo grupo de amplificación y/o reamplificación puede ofrecer importantes pistas al respecto.

Para evaluar este tipo de contaminación las secuencias obtenidas fueron ordenadas en grupos de amplificación y reamplificación y se compararon entre sí las secuencias de cada grupo.

Aquellas muestras sospechosas de haber sufrido una contaminación por *carry over* no fueron tenidas en cuenta para el análisis poblacional y estadístico posterior.

### ***14.4.3. Eliminación de aquellas secuencias con más de una posición ambigua***

La presencia de posiciones con más de un nucleótido (habitualmente dos) evidencia más de un tipo de DNA en el producto final de amplificación. Una explicación para ello es la presencia de moléculas de DNA de diferente naturaleza en el extracto original. Esto a su vez puede deberse a heteroplasmia, a la presencia de DNA contaminante en el extracto o a la existencia de cadenas “dañadas” molecularmente en algunas posiciones (*Miscoding lesions*). Sin embargo tal situación puede responder además a la introducción de una contaminación puntual durante el proceso de amplificación no detectada en los blancos. Mediante la clonación bacteriana del producto de amplificación y la secuenciación de múltiples clones pueden separarse los diferentes productos de amplificación, lo que facilita la determinación del origen de cada uno de ellos.

Aquellas secuencias con más de una de éstas posiciones ambiguas que no pudieron ser clonadas se descartaron directamente. Las secuencias con una única posición con doble nucleótido no clonadas fueron discutidas una por una prescindiendo de aquellas en las que existía alguna duda acerca de su autenticidad.

## **14.5. Pruebas adicionales**

### ***14.5.1. Ensayos bioquímicos de preservación de otras macromoléculas de la misma muestra***

Para determinar si el estado de preservación de las muestras de diente o hueso analizadas era compatible con el estado de preservación de las moléculas de DNA y con la consecuente obtención de secuencias de mtDNA auténticas a partir de las mismas, se estimaron el grado de racemización del ácido aspártico y la composición de aminoácidos en estas mismas muestras. Dado que ambos procedimientos pueden ser empleados con una finalidad prospectiva los hemos descrito con detalle al inicio de este capítulo en el subapartado 2.1. del apartado “Prospección molecular”.

Este criterio no pudo ser aplicado a todas las muestras de la presente tesis doctoral puesto que en algunos casos no quedaba muestra de hueso o diente disponible tras la extracción del material genético.

#### **14.5.2. Clonación bacteriana de los productos de amplificación y secuenciación de múltiples clones**

##### *14.5.2.1. Muestras clonadas*

Se clonaron un total de 25 productos de amplificación, 13 correspondientes al primer segmento (posiciones 16106-16276) y 12 al segundo (16238-16389), de 13 muestras de diferentes yacimientos y épocas (**Tabla ME27**).

**Tabla ME27.** Número de amplificaciones clonadas de cada muestra.

Procedencia	Individuo	Muestra	Amplificaciones clonadas	
			Primer segmento	Segundo segmento
Dja'de	SK-A2	1DJ1	1	1
Nerja	NE-NT79	3NE	1	1
Nerja	NE-NM82.2	2NE	1	1
Mari	12	1MA12	1	1
Muge	TAS 216	TA3	1	0
El Pirulejo	1PI	1PI	1	1
Tell Halula	H-68	H12	1	1
Tell Halula	H-37	2H20	1	1
Tell Ramad	R65-14	TR12	1	1
Tell Ramad	R69(2)	TR14	1	1
Tell Ramad	R66-N4-N°400	TR19	1	1
Toledo	K-13	TO1	1	1
Tres Montes	TM-11	1TM6	1	1
<b>Total</b>			<b>13</b>	<b>12</b>

No pudieron ser clonados todos los productos de amplificación generados durante el presente trabajo puesto que, en la mayoría de los casos, no quedaba producto de amplificación suficiente después de realizar las reacciones de secuencia. Se clonaron sin embargo algunas de las muestras de mayor antigüedad para las que se habían obtenido amplificaciones, como 2NE, 3NE, TA3 o TO1; algunas cuya secuencia directa presentaba posiciones con doble nucleótido, como 1DJ1, 2H20 o TR14 y un conjunto de muestras al azar de diferentes yacimientos como 1MA12, H12, TR12, TR19 y 1TM6



para las que se habían obtenido secuencias sin ambigüedades con el fin de evaluar la fiabilidad de la secuenciación directa.

#### *14.5.2.2. Protocolo experimental*

La clonación de las amplificaciones se llevó a cabo en el Centro de Biología Molecular del C.S.I.C. en Barcelona, con la supervisión de la Dra. Eva Prats, y en el laboratorio de Genética Forense y Genética de Poblaciones del Departamento de Toxicología y Legislación Sanitaria de la Facultad de Medicina de la Universidad Complutense de Madrid.

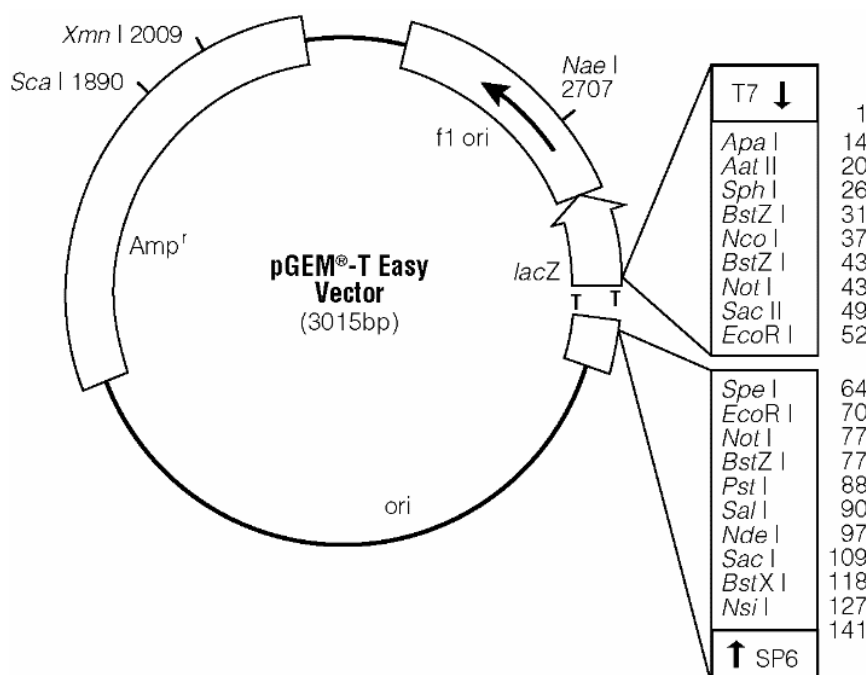
##### 14.5.2.2.1. Sistema de clonación pGEM<sup>®</sup>-T Easy Vector System II (Promega)

Los plásmidos bacterianos son moléculas circulares de DNA bicatenario de pequeño tamaño (entre 1 y 200 kb) que tienen la capacidad de auto-replicarse mediante los enzimas y proteínas codificados por su huésped bacteriano. Muchas veces, los plásmidos codifican algún enzima que puede conferirle alguna ventaja sobre el huésped (por ejemplo, la resistencia a un determinado antibiótico). La capacidad replicativa está regulada por el propio plásmido, de manera que algunos pueden multiplicarse hasta llegar a las 700 copias dentro de una misma célula, mientras que otros sólo presentan una copia.

Los plásmidos se utilizan en el laboratorio como vectores de clonación. Se dispone de varios tipos ampliamente establecidos desde hace algunas décadas para su uso en el laboratorio. En nuestro caso se ha empleado el sistema de clonación pGEM<sup>®</sup>-T Easy Vector System II de Promega. Este sistema consiste en el plásmido pGEM<sup>®</sup>-T Easy (**Figura FME12**) digerido con EcoRV al que se le ha añadido una T 3' terminal. La incorporación de esta T incrementa la eficiencia de ligación de un producto de PCR en este vector puesto que la mayor parte de las polimerasas incorporan una A en ambos extremos 3' del producto de amplificación.

El plásmido pGEM<sup>®</sup>-T Easy presenta una zona “polylinker” (*multiple cloning region*) flanqueada por los promotores SP6 y T7 de la RNA polimerasa. Esta región contiene una secuencia que puede ser reconocida por diferentes enzimas de restricción lo

que permite cortarla en puntos determinados e insertar el fragmento de DNA exógeno que se desea estudiar. Esta región está además flanqueada por dianas de restricción para los enzimas EcoRI, BstZ y NotI, lo que permite liberar el inserto mediante una digestión simple o doble.



**Figura FME12.** Mapa circular del plásmido pGEM<sup>®</sup>-T Easy (Promega) (Technical manual N° 042, pGEM<sup>®</sup>-T and pGEM<sup>®</sup>-T Easy Vector Systems, Promega).

#### 14.5.2.2.2. Selección de las colonias con plásmido

El DNA plasmídico se introduce en el interior de una célula huésped mediante un proceso denominado transformación que requiere de la permeabilización de su membrana (bacteria competente) para que sea posible. La eficiencia de la transformación es reducida, por lo que es necesario que el plásmido tenga algún marcador selectivo para identificar los transformantes. Este marcador suele ser un gen que codifica para la resistencia a un antibiótico determinado, la ampicilina en el caso de pGEM<sup>®</sup>-T Easy. Este antibiótico inhibe una serie de enzimas implicados en la síntesis de la pared bacteriana, lo cual conduce a la lisis de la célula (SAMBROOK *et al.* 1989). La incorporación del plásmido confiere a las células huésped la capacidad de crecer en un medio que contiene este antibiótico (resistencia a la ampicilina o Amp<sup>R</sup>), posibilitando su selección.

#### 14.5.2.2.3. Selección de las colonias con inserto

La región *polylinker* del plásmido está contenida dentro del gen lac Z que codifica para un fragmento de la enzima  $\beta$ -galactosidasa. La célula huésped codifica la fracción complementaria de esta enzima, necesaria para que sea activa (ULLMANN *et al.* 1967).

La incorporación de un inserto en un plásmido rompe la pauta de lectura del gen lacZ de manera que la bacteria que lo presente no producirá  $\beta$ -galactosidasa activa (bacteria lac-). En presencia de un activador del gen lacZ, el IPTG (isopropiltio- $\beta$ -D-galactosidasa), y de un análogo cromogénico de la galactosa, el X-Gal (HORWITZ *et al.* 1964), sólo aquellas bacterias con actividad  $\beta$ -galactosidasa (lac+) serán capaces de hidrolizar el X-Gal generando un sustrato azulado apreciable a simple vista. Sin embargo, aquellas bacterias que hayan incorporado el plásmido, carentes por tanto de  $\beta$ -galactosidasa activa, no serán capaces de hidrolizar el X-Gal y se multiplicarán dando lugar a colonias blancas.

#### 14.5.2.2.4. A-Tailing

Para un correcto ligado del inserto al vector pGEM<sup>®</sup>-T Easy es necesario que el primero presente dos A en posición 3' terminal.

Las polimerasas termoestables con actividad correctora de errores, como la empleada en la mayor parte de las amplificaciones de la presente tesis doctoral (*Taq* Expand High Fidelity, Roche), generan productos de amplificación con extremos romos. Para la clonación de un producto de *PCR* de estas características es necesario incorporar un paso previo a la reacción de ligado, consistente en la introducción *in vitro* de adeninas en sus extremos mediante una reacción especial de *PCR* denominada “A-tailing”. En esta reacción el sustrato es dATP y se emplea una *Taq* polimerasa sin actividad correctora de errores. En la **Tabla ME28** se detallan las cantidades de cada componente de la reacción de *A-tailing* empleadas en la presente tesis doctoral.

**Tabla ME28.** Componentes de la mezcla de reacción de A-tailing.

Reactivo	Stock	Reacción	Volumen (1 reacción)
PCRB	10x	1x	1µl
MgCl <sub>2</sub>	50mM	2.5mM	0.2µl
DATP	10mM	0.2mM	0.5µl
<i>Taq</i> pol	1u/µl	0.1u/µl	1µl
Producto <i>PCR</i>	-	-	5µl
H <sub>2</sub> O	-	-	2.3µl

Una vez preparadas las reacciones, se amplificaron en un termociclador eppendorf a 70°C durante 30 minutos.

#### 14.5.2.2.5. Reacción de ligado

La reacción de ligado consiste en la unión del DNA exógeno al vector plasmídico mediante unas enzimas denominadas “ligasas”. Éstas enlazan los grupos fosfato de los extremos 5' y los grupos hidroxilo 3' de las moléculas de DNA bicatenario en presencia de ATP. Para que se formen 2 enlaces fosfodiéster nuevos es necesario que el plásmido haya sido desfosforilado previamente por sus extremos 5'. El sistema pGEM<sup>®</sup>-T Easy incorpora la DNA ligasa T4, que es capaz de unir fragmentos de DNA con extremos romos, junto con un tampón de ligado de esta enzima a concentración 2x.

En las reacciones de ligación, la proporción vector-inserto es un factor importante para obtener una eficiencia máxima en el número de moléculas recombinantes. En nuestro caso siguiendo las recomendaciones del proveedor la relación empleada fue de 3 a 1. Para el cálculo de la cantidad de inserto necesaria para dicha relación se empleó la siguiente fórmula (Technical manual N° 042, pGEM<sup>®</sup>-T and pGEM<sup>®</sup>-T Easy Vector Systems, Promega):

$$\frac{\text{ng de vector} \times \text{tamaño del inserto (kb)}}{\text{tamaño del vector (kb)}} \times \text{ratio molar inserto:vector} = \text{ng de inserto}$$

Teniendo en cuenta que el tamaño del plásmido pGEM<sup>®</sup>-T Easy es de aproximadamente 3kb y que se sirve a una concentración de 50ng/µl:

$$\frac{50\text{ng de vector} \times 0.15\text{kb de inserto}}{3\text{kb de vector}} \times \frac{3}{1} = 7.5\text{ng de inserto}$$

La concentración del producto de A-tailing se estimó por comparación con un marcador de peso molecular y concentración conocidas en un gel de electroforesis, tal y como se describe en el apartado 12.2 de este mismo capítulo.

Junto con los productos de A-tailing, se procesaron dos controles siguiendo las recomendaciones del kit de clonación (Technical manual N° 042, pGEM<sup>®</sup>-T and pGEM<sup>®</sup>-T Easy Vector Systems, Promega). Éstos son necesarios para evaluar el rendimiento del sistema de clonación.

El *control positivo*, incluido en el kit, consiste en un fragmento de 542pb del plásmido pGEM<sup>®</sup>-luc. Este inserto ha sido diseñado específicamente para producir colonias blancas puesto que contiene mutaciones que producen múltiples codones stop en las seis pautas de lectura del gen lacZ. Cualquier colonia azul que crezca en este control procederá por tanto de aquellos plásmidos no digeridos o que hayan perdido las T terminales. Este control permite evaluar por tanto la eficiencia de la reacción de ligado. Si ésta es correcta, el porcentaje esperado de colonias blancas debe ser superior al 60%. Un porcentaje de colonias blancas inferior al 50% es indicador de unas condiciones sub-óptimas de ligado.

El *control de background* consiste en una reacción de ligación en ausencia de inserto. Este control permite cuantificar el número de colonias azules resultantes de plásmidos no digeridos o que han perdido las T terminales. Habitualmente el número de colonias obtenidas si las eficiencias de transformación y ligación son adecuadas oscila entre 10 y 30. Este número debe ser similar al número de colonias azules obtenidas con un producto de PCR. Si el número de colonias azules obtenidas con un producto de PCR es superior al del control de background es probable que la eficiencia de ligación sea sub-óptima aunque también puede suceder que los recombinantes se encuentren dentro de estas colonias azules.

Para llevar a cabo la reacción de ligado se siguieron los siguientes pasos:

- Se centrifugaron brevemente los tubos de pGEM<sup>®</sup>-T Easy Vector y el inserto control.
- Se fabricaron las reacciones de ligación tal y como se describe en la **Tabla ME29**. Antes de añadir el tampón de ligación 2x se agitó el tubo vigorosamente.

- Se mezclaron las reacciones suavemente con la pipeta.
- Se incubaron a 16°C toda la noche en el termociclador.

**Tabla ME29.** Componentes de las reacciones de ligado.

Reactivo	Stock	Reacción	Volumen 1 reacción		
			Productos <i>A-tailing</i>	Control positivo	Control de <i>background</i>
Tampón de ligado	2x	1x	5µl	5µl	5µl
pGEM easy-vector	50ng	5ng	1µl	1µl	1µl
Producto de <i>PCR</i>	-	-	3µl	-	-
Inserto de DNA control	4ng/µl	-	-	2µl	-
DNA ligasa T4			1µl	1µl	1µl
H <sub>2</sub> O			-	1µl	3µl
<b>Total</b>			<b>10µl</b>	<b>10µl</b>	<b>10µl</b>

#### 14.5.2.2.6. Transformación

La transformación consiste en la introducción de un plásmido en una bacteria previamente tratada (bacteria competente). Existen diversos métodos para obtener bacterias competentes, aunque el más común consiste en la permeabilización de la membrana bacteriana mediante CaCl<sub>2</sub>. El sistema de clonación pGEM<sup>®</sup>-T Easy Vector System II (Promega) incluye células competentes JM109 de elevada eficiencia de transformación ( $1 \times 10^8$  cfu/µg DNA). La ligación de fragmentos que contienen una única A terminal suele ser ineficiente, por lo que resulta esencial emplear células con una elevada eficiencia de transformación para obtener un número óptimo de colonias.

- Se prepararon tantas placas LB/ampicilina/IPTG/X-Gal como reacciones de ligación más dos placas para determinar la eficiencia de transformación. Una vez preparadas se equilibraron a temperatura ambiente antes de plaquear.
- Se descongelan las células competentes (mantenidas a -70°C) depositándolas en un recipiente con hielo picado durante aproximadamente 20 minutos. Una vez descongeladas se mezclaron agitando el tubo suavemente.
- Se prepararon tantos eppendorf estériles de 1.5ml como mezclas de ligado más dos para los controles. Una vez descongeladas las células se transfirieron 100µl a cada eppendorf y se depositaron en hielo picado.

- Se centrifugaron los tubos que contenían las reacciones de ligado y se añadieron 5µl de cada reacción a cada eppendorf depositado en hielo que contenía las células competentes. Se dejó un eppendorf únicamente con células competentes. Éste, una vez plaqueado, sirve para testar la presencia de plásmidos en el ambiente o en el interior de las propias células competentes (***control negativo de transformación***). Para evaluar la eficiencia de transformación de las células competentes se añadió a otro eppendorf 0.1ng de un plásmido circularizado y sin inserto (***control positivo de transformación***). Este plásmido ha de ser necesariamente diferente a pGEM® -T Easy, puesto que este vector se suministra linearizado.

Teniendo en cuenta que la eficiencia de transformación se calcula:

$$\frac{\text{n}^\circ \text{ c.f.u.}}{\mu\text{g DNA}/\mu\text{l célula competente}} = \text{eficiencia de transformación (c.f.u./}\mu\text{g)}$$

si la eficiencia de las células competentes es la esperada ( $1 \times 10^8$  cfu/µg) el número de unidades formadoras de colonia (*colony forming units* ó *c.f.u.*) del control positivo de transformación deberá ser superior a 100.

- Se agitaron suavemente los tubos y se dejaron en hielo durante al menos 30 minutos.
- Se sometieron las mezclas de células y plásmido y los dos controles a un choque térmico manteniéndolas durante 2 minutos exactamente a 42°C en un bloque térmico o estufa. A continuación se depositaron rápidamente de nuevo en hielo.
- Se plaqueó el contenido de los diferentes eppendorf en las placas LB/ampicilina/IPTG/X-Gal
- Se incubaron las placas durante 16-24 horas a 37°C. Cuando se plaquearon 100µl de células competentes con una eficiencia de  $1 \times 10^8$  cfu/µg DNA, se observaron aproximadamente 100 colonias por placa. Después de esta incubación las células pueden mantenerse a 37°C durante unas horas más o guardarse a 4°C para facilitar el desarrollo de la coloración azul. Aquellas colonias con actividad β-galactosidasa suelen crecer menos que aquellas que carecen de dicha actividad, por lo que las colonias azules deberán ser más pequeñas que las blancas, que suelen medir aproximadamente un milímetro de diámetro.

14.5.2.2.7. Selección de colonias con inserto mediante *colony-PCR*

Como ya se ha mencionado en el apartado 14.5.2.2.3. la introducción de un inserto en el vector pGEM<sup>®</sup>-T Easy interrumpe la secuencia del gen de la  $\beta$ -galactosidasa por lo que los clones recombinantes y no recombinantes pueden ser distinguidos por su coloración. Sin embargo las características de los productos de *PCR* clonados pueden afectar significativamente al ratio de colonias azules y blancas obtenidas tras la transformación de células competentes. En la mayoría de los casos los clones que contienen productos de *PCR* producen colonias blancas. Sin embargo algunas colonias azules pueden contener también plásmidos con inserto. Esto sucede por ejemplo cuando los fragmentos de *PCR* se clonan en pauta de lectura con el gen *lacZ*. Estos fragmentos son normalmente múltiplos de longitud de 3 pares de bases, incluyendo las A terminales, y no contienen codones stop en su secuencia. En otros casos el proceso de amplificación puede introducir ciertas modificaciones como mutaciones o deleciones que al ser leídas dan lugar a una proteína activa que resulta en colonias de color azul. Por eso, de manera previa a la purificación del DNA plasmídico y su secuenciación, ambos procesos caros y metodológicamente muy laboriosos, resulta conveniente comprobar que las colonias a purificar realmente contienen el inserto. En el presente estudio las colonias con inserto fueron seleccionadas mediante “*colony PCR*”.

- Se prepararon reacciones de *PCR* de 25  $\mu$ l que contenían todos los reactivos excepto el DNA con los mismos cebadores con los que se había amplificado el inserto (**Tabla ME30**).

**Tabla ME30.** Componentes de la mezcla de reacción de “*colony PCR*”.

Componentes mezcla reacción	Concentración stock	Concentración final en reacción	Volumen (1 reacción)
Agua	-	-	19.55 $\mu$ l
PCRB	10x	1x	2.5 $\mu$ l
MgCl <sub>2</sub>	50 mM	2 mM	1 $\mu$ l
dNTPs (Biotools)	10 mM	0.2 mM	0.5 $\mu$ l
Cebadores	100 $\mu$ M	0.4 $\mu$ M	0.1 $\mu$ l/cada
<i>Taq</i> (Biotools)	1 u/ $\mu$ l	0.05 u/ $\mu$ l	1.25 $\mu$ l
Total	-	-	25 $\mu$ l



- Se picaron las colonias con una punta de pipeta estéril y se mojó cada una de ellas en una mezcla diferente de *PCR*.
- Se tiró cada punta en un tubo con medio LB líquido y se dejaron crecer toda la noche a 37°C. Alternativamente puede hacerse una pequeña estría en una placa de LB agar. Ambas formas permiten conservar las colonias bacterianas vivas para la posterior purificación de su DNA.
- Se amplificaron las mezclas de reacción en un termociclador con el siguiente programa de *PCR*:

**Tabla ME31.** Programa de amplificación para “*colony PCR*”.

Proceso	Nº ciclos	Temperatura	Tiempo
Lisis celular	1	94°C	10 min
Desnaturalización		94°C	1 min
Annealing	30	52°C	1 min
Extensión		72°C	1 min
Extensión final	1	72°C	10 min

- Se visualizaron los resultados en electroforesis en gel de agarosa al 2%. Únicamente aquellas colonias que habían incorporado el plásmido con el inserto dieron resultado positivo de amplificación. El DNA de estas mismas cepas puede ser purificado a partir de su cultivo líquido o en placa.

#### 14.5.2.2.8. Minipreparaciones de DNA plasmídico

Se pusieron a crecer en un cultivo líquido aquellas colonias con plásmido con inserto, seleccionadas mediante “*colony PCR*”, para obtener un número mayor de ellas. A continuación se separó el DNA plasmídico del bacteriano mediante purificación por mini-preparaciones o “mini-preps”.

Se empleó un kit comercial para la obtención de “mini-preps” plasmídicas (Jetquick Plasmid Miniprep Spin Kit, Genycell®). Este procedimiento se basa en la incubación de las células en una solución alcalina con NaOH y un agente tensoactivo (SDS) que provocan la lisis bacteriana y la desnaturalización del DNA cromosómico. Tras la neutralización el lisado limpio se aplica directamente en una columna Jetquick Micro-Spin que permite separar por centrifugación el DNA plasmídico, que queda pegado a la

membrana, del sobrenadante, que se deposita en el fondo del tubo. Este DNA es lavado para eliminar el RNA y otras impurezas y posteriormente se eluye de la membrana mediante TE buffer o agua.

- Se inoculó cada colonia previamente aisladas de la cepa transformada en 1.5ml de LB con 50 µg/ml de ampicilina. Se incubaron a 37° C durante una noche, en agitación.

- Se centrifugó cada cultivo durante 5min a 14000 r.p.m. Tras la centrifugación se genera un precipitado blanquecino que queda adherido a la pared del tubo. Se eliminó todo el sobrenadante primero con una pipeta y luego con un capilar.

- Se añadieron 250µl de solución G1 al *pellet* y se resuspendieron las células con una pipeta hasta conseguir una mezcla homogénea. La solución G1 contiene 50mM Tris/HCl pH 8, 10mM EDTA y 100µg/ml de RNAsa A.

- Se añadieron 250µl de solución G2 y se mezcló cuidadosamente invirtiendo el tubo hasta obtener una solución homogénea. Se incubaron durante 5 min. La solución G2 contiene 200mM de NaOH y 1% SDS.

- Se añadieron 350µl de solución G3 y se mezcló cuidadosamente invirtiendo el tubo hasta obtener una solución homogénea. A continuación se centrifugó la mezcla a temperatura ambiente a máxima velocidad durante 10 minutos. La solución G3 contiene acetato y hidrocloreuro de guanidina.

- Se colocó cada columna en el interior de un tubo colector (suministrado) y se transfirió el sobrenadante del paso anterior al interior de la columna. Se centrifugó a máxima velocidad (>12000 g) durante 1 minuto y se eliminó a continuación el sobrenadante depositado en el tubo colector.

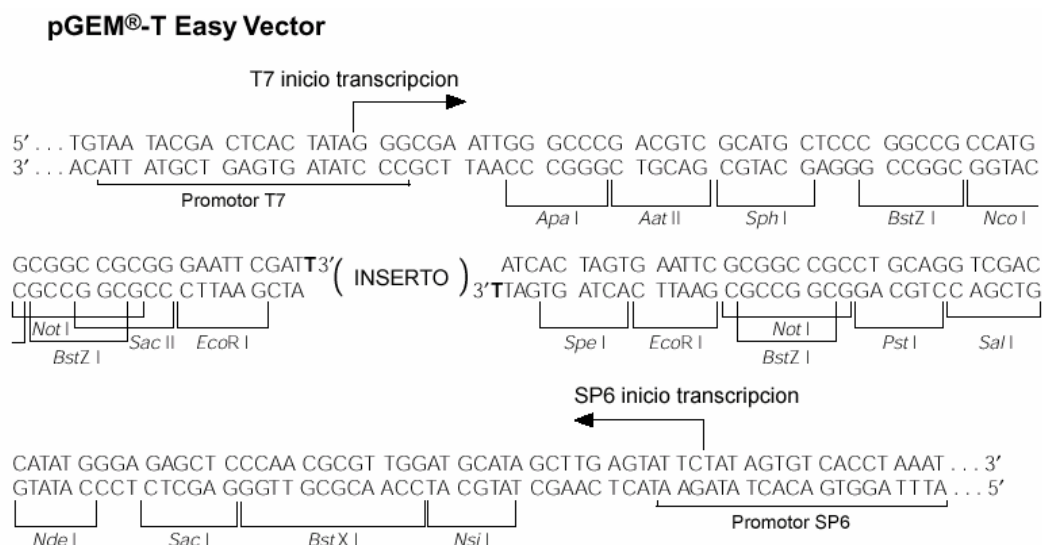
- Una vez vaciado el tubo colector, se volvió a colocar en la columna y se añadieron 500µl de la solución G4 que contiene etanol, NaCl, EDTA y Tris-HCl. Se centrifugó a máxima velocidad durante 1 minuto. Se vació cada tubo colector, se encajó en la columna y se volvió a centrifugar a velocidad máxima durante 1 minuto. Mediante esta segunda centrifugación se consigue eliminar totalmente la solución G4.

- Se depositó cada columna en el interior de un nuevo tubo de microcentrífuga de 1.5ml y se añadieron 50µl de agua estéril directamente en el centro de la matriz de sílica de la columna. Se centrifugó la columna a máxima velocidad durante 2 minutos.

- Se separaron las columna de los microtubo y se desecharon. Los microtubos con la solución plasmídica se guardaron a -20°C.

## 14.5.2.2.9. Secuenciación de los productos de clonación

Se secuenció el DNA plasmídico purificado empleando los cebadores T7 y SP6. Estos cebadores hibridan con la región correspondiente a dos promotores del plásmido, que flanquean la zona del inserto (**Figura FME13**).



**Figura FME13.** Secuencia de los promotores y sitios de clonación múltiple del plásmido **pGEM®-T Easy**. La secuencia de la cadena superior corresponde al RNA sintetizado por la T7 RNA polimerasa. La cadena inferior corresponde al RNA sintetizado por la SP6 RNA polimerasa.

En las **Tablas ME32 y ME33** se detallan respectivamente la composición de la mezcla de secuencia y el programa de amplificación empleado. El DNA plasmídico fue cuantificado tal y como se describe en el apartado 12.2. de este mismo capítulo. Las reacciones de amplificaciones se realizaron en un termociclador TC1 (Perkin Elmer).

**Tabla ME32.** Composición de la mezcla de secuencia empleada para secuenciar el DNA plasmídico

Componentes	Cantidad	Volumen
Terminator Ready Reaction Kit vs.1	-	3µl
DNA molde	200-500µg	5µl
Cebador	1.6pmol	1µl
H <sub>2</sub> O	-	1µl
<b>Total</b>	-	<b>10µl</b>

**Tabla ME33.** Programa de secuenciación del DNA plasmídico.

<b>Temperatura</b>	<b>Tiempo</b>	<b>Número ciclos</b>
96°C	30seg	
50°C	15seg	25
60°C	4 min	

#### ***14.5.3. Cuantificación del número de moléculas de DNA template amplificable presentes en un extracto.***

La estima del número de moléculas de DNA molde en los extractos de DNA antiguo fue realizada mediante *Real Time PCR*. Tal y como sucede con la racemización del ácido aspártico y el contenido de aminoácidos de las muestras, este procedimiento también puede ser empleado como método de prospección molecular, por lo que ha sido incluido y descrito detalladamente al inicio del presente capítulo en el sub-apartado 2.2 del apartado Prospección molecular. No pudieron cuantificarse todos los extractos mediante este método dada la cantidad limitada de los mismos.

Aquellas secuencias procedentes de aquellos extractos cuantificados que proporcionaron un número de moléculas de mtDNA inferior al millar fueron descartadas directamente puesto que no se podía demostrar que sus mutaciones eran endógenas y no fruto de daño molecular en el DNA template.

## 15. ANÁLISIS ESTADÍSTICO

### 15.1. Eficiencias técnicas

#### 15.1.1. Cálculo de eficiencias

Se calcularon las eficiencias de las reacciones de amplificación, reamplificación y secuenciación sobre las muestras estudiadas dividiendo el número de resultados positivos por el número de intentos totales de cada una de las reacciones. Para dicho cálculo únicamente se tuvieron en cuenta aquellas reacciones cuyo resultado había sido visualizado en un gel de electroforesis y que no habían dado resultado de contaminación en los blancos.

Además de las eficiencias globales de amplificación, reamplificación y secuenciación, para evaluar la influencia de diversas variables sobre las mismas se calcularon también un conjunto de eficiencias parciales. De esta forma se estimaron las eficiencias en función de:

- Tipo de tejido analizado: hueso o diente.
- Tipo de pieza dental:
  - Incisivo, Canino, Premolar, Molar.
  - I1, I2, C, P1, P2, M1, M2, M3.
- Longitud del segmento amplificado: 300pb o 150pb.
- Tipo de cebador empleado en la amplificación y reamplificación.
- Diseño experimental.
- Tipo de polimerasa empleada en la amplificación.
- Procedencia de las muestras:
  - Yacimiento arqueológico.
  - Localización geográfica I: Próximo Oriente, Península Ibérica Norte, Península Ibérica Sur, Península Ibérica Oeste, Península Ibérica Este.
  - Localización geográfica II: Península Ibérica, Próximo Oriente.
  - Antigüedad: Calcolítico-Bronce, Medieval, Meroítico, Neolítico 1, Neolítico 2,
  - Paleolítico 1, Paleolítico 2.

Para el cálculo de las eficiencias en función del tipo de pieza dental empleada, únicamente se tuvieron en cuenta aquellas muestras para las que esta información era conocida. Las características de las categorías de edad empleadas para determinar la influencia de la antigüedad en las tres eficiencias técnicas se describen en la **Tabla ME34**.

**Tabla ME34.** Clasificación de las muestras en base a su cronología en siete categorías de edad.

Categoría de edad	Intervalo temporal	Yacimientos-muestras	Datación
Edad Media	S. XI-XII	Garai	S. XI-XIII
Meroítico	2130-2320 A.P	Amir Abdallah	2320-2130 A.P.
Calcolítico-Bronce	4900-3200 A.P.	Abauntz	4240 A.P.
		Tres Montes	4130 A.P.
		Mari	2900-2700 A.C.
		4NE	4260±70 A.P.
		Atxuri	4400-3500/3200 A.P.
Neolítico 2	9200-5875 A.P.	Caldeirao	6800-6200 A.P.
		Muge	7300-6300 A.P.
		Toledo	9200-7800 A.P.
		1NE-5NE	5875±80 A.P.
Neolítico 1	9400-7750 A.P.	Tell Halula	8800-8000 A.P.
		Tell Ramad	8000-7750 A.P.
		Dja'de	9400-8700 A.P.
		Açikli	8900-8200 A.P.
Paleolítico 2	20000-12500	2NE	20000-17000 A.P.
		3NE	20000-17000 A.P.
		El Pirulejo	13500-12500 A.P.
Paleolítico 1	33400 -27000 A.P.	Zafarraya	33400 -27000 A.P.

Para el cómputo total del número de reacciones totales y positivas de cada tipo, primero se codificaron los datos en una hoja de cálculo Excel (Microsoft Office 2000) en la que se especificaba: el nombre de la muestra, el yacimiento del que procedía, el tipo de tejido estudiado, el tipo de pieza dental, el tipo de cebador empleado para la amplificación, el tipo de polimerasa con la que se había realizado la amplificación y/o reamplificación, la fecha en la que se había llevado a cabo cada una de las reacciones (amplificación, reamplificación y secuenciación) y el resultado obtenido (positivo, negativo o contaminado). El cálculo fue realizado con la opción “Tablas dinámicas” del programa Excel seleccionando los campos deseados y los criterios “Contar de fecha de amplificación”, “Contar de fecha de reamplificación” o “Contar de fecha de secuenciación”.

### ***15.1.2. Comparación de las eficiencias***

Para determinar si existían diferencias significativas en las eficiencias de amplificación, reamplificación y secuenciación entre las diferentes clases de las diferentes variables se aplicaron dos tipos de pruebas estadísticas. En aquellos casos en los que existían dos clases que comparar se empleó la prueba exacta de Fisher mientras que cuando existían más de dos clases fue aplicado el Test de la Ji cuadrado ( $\chi^2$ ) con la corrección de Yates. Ambos tipos de análisis se realizaron con el programa GraphPad Prism versión 4.00 para Windows (GraphPad Software, San Diego California USA).

## 16. ANÁLISIS POBLACIONAL

### 16.1. Construcción de una base de datos de secuencias actuales

Para la interpretación de las secuencias de DNA mitocondrial es necesario un marco de referencia, que en este caso viene proporcionado por la diversidad mitocondrial de las poblaciones actuales. Con esta finalidad se procedió a la construcción de una base de datos de secuencias de DNA mitocondrial de diversas poblaciones europeas, asiáticas y africanas a partir de otras bases de datos disponibles en internet.

#### 16.1.1. Bases de datos utilizadas

Para la construcción de nuestra base de datos de secuencias de DNA mitocondrial se utilizaron las siguientes bases de datos:

- HVRbase (HANDT *et al.* 1998).

<http://www.HVRbase.de>

- FHG Base (Ferrara Human Genetics Database).

<http://web.unife.it/progetti/genetica/pdata.htm>

##### 16.1.1.1. HVRbase

La base de datos HVRbase es un compendio de secuencias de la región control del mtDNA de humanos actuales, primates y neandertales procedentes de diferentes publicaciones o recuperadas directamente del GenBank. Contiene 4079 secuencias alineadas de la Región Hipervariable I (HVRI) (posiciones 16001-16408), 969 de la Región Hipervariable II (HVRII) (posiciones 1-418) así como información acerca de los individuos a los que pertenecen dichas secuencias.

En esta base de datos pueden realizarse búsquedas de secuencias por continentes, referencias originales, especies, países, orígenes geográficos, poblaciones y *phyla* lingüísticos.



Una vez definidas las opciones de búsqueda, pueden escogerse las características del “output” de secuencias obtenidas, como las posiciones inicial y final, el tipo de alineamiento (phylip o texto), el nombre de la secuencia y su comparación o no con la secuencia consenso (*CRS*) o con cualquier otra secuencia.

La búsqueda proporciona los resultados organizados en una doble ventana: en la parte superior se detalla la información relativa a cada individuo (continente, especie, origen geográfico, referencia original, *phylum* lingüístico, población de origen y nombre original de la secuencia) mientras que la parte inferior contiene las secuencias alineadas en el formato escogido. Los resultados pueden bajarse seleccionando por separado cada una de las dos ventanas, copiándolos y pegándolos en un editor de texto ASCII del tipo Notepad o Wordpad.

#### 16.1.1.2. FHG Base

La base de datos de la Universidad de Ferrara contiene cuatro bases de datos diferentes, una de secuencias de DNA mitocondrial (“MtDNA database”; con las secuencias de SIMONI *et al.* 2000) y tres de SNPs de cromosoma Y procedentes de diferentes publicaciones (ROSSER *et al.* 2000; SEMINO *et al.* 2000; UNDERHILL *et al.* 2000).

La “MtDNA database” contiene 2566 secuencias de la Región Hipervariable I de 35 poblaciones europeas y de Próximo Oriente. La información de cada una de las poblaciones (código, número de secuencias, procedencia, *phylum* lingüístico, subfamilia lingüística y referencia original) se encuentra resumida en un fichero *excel*. La base de datos consiste en un fichero ASCII que contiene todas las secuencias alineadas y separadas por poblaciones. Ambos ficheros pueden descargarse directa y gratuitamente del sitio web.

#### 16.1.2. Fiabilidad de las bases de datos

Para comprobar si existían errores en las bases de datos se eligieron al azar 336 secuencias de un total de 13 poblaciones de HVRbase y se cotejaron con las referencias bibliográficas correspondientes (**Tabla ME35**). En 9 de ellas no se detectó

ningún error. En 3 de las 4 poblaciones restantes fueron detectados errores menores al 12%, y sólo en una de ellas (Alemania del Sur) se detectaron errores superiores al 20%. Las secuencias que presentaban diferencias entre la publicación y la base de datos fueron eliminadas de la base de datos. Los datos de la población de Alemania del Sur se eliminaron totalmente y no se emplearon en la construcción de la base de datos de poblaciones actuales.

**Tabla ME35.** Comparaciones realizadas entre las secuencias de las bases de datos y las publicaciones originales. N: número de secuencias.

Población	N		Comparaciones		Eliminadas	Referencia
	publicación	Base datos	N	Coincidencias		
Alemania Rhin	50	50	15	15	0	Baasner <i>et al.</i> 1998
Alemania Sur	200	199	40	32	8	Lutz <i>et al.</i> 1998
Austria	101	100	20	20	0	Parson <i>et al.</i> 1998
Bulgaria	30	29	10	10	0	Calafell <i>et al.</i> 1996
Canarias	54	52	20	20	0	Pinto <i>et al.</i> 1996
Francia	50	49	12	12	0	Rousselet <i>et al.</i> 1998
Galicia	92	91	91	90	1	Salas <i>et al.</i> 1998
Cerdeña	46	68	33	31	2	DiRienzo y Wilson 991
Italia Toscana	52	48	15	15	0	Françalacci <i>et al.</i> 1996
País Vasco	45	44	15	15	0	Bertranpetit <i>et al.</i> 1995
Suecia	28	27	27	24	3	Kittles <i>et al.</i> 1999
Turquía	29	26	8	8	0	Calafell <i>et al.</i> 1996
Kurdistán	30	30	30	30	0	Comas <i>et al.</i> 2000
<b>Total</b>	<b>807</b>	<b>813</b>	<b>336</b>	<b>322</b>	<b>14</b>	

### 16.1.3. Composición de la base de datos

Se seleccionaron un total de 44 poblaciones actuales, 35 procedentes de la base de datos HVRbase y 9 de la FHG Base. A partir de estas poblaciones originales se generó una nueva base de datos formada por 38 poblaciones de Europa, Próximo Oriente y África. En la Tabla **ME36** se detalla la composición y características de estas poblaciones. En las figuras **FME14** y **FME15** se localizan estas poblaciones en el mapa.

Tabla ME36. Composición de la base de datos de secuencias actuales diseñada en la presente tesis doctoral. N: Número de secuencias. P.c. Comunicación personal.

Población	Procedencia	Grupo	Lengua	N	Identificador	Referencia	Base datos	
África Este	Etiopía	Somali	Somali (Afro-asiático)	9	AFRE	Watson <i>et al.</i> 1996	HVRbase	
	Etiopía	Borana	Oromo (Afro-asiático)	1	AFRE	Watson <i>et al.</i> 1996	HVRbase	
	Kenya	Somali	Somali (Afro-asiático)	15	AFRE	Watson <i>et al.</i> 1996	HVRbase	
	Kenya	Turkana	Turkana (Nilo-Sahariano)	37	AFRE	Watson <i>et al.</i> 1996	HVRbase	
	Kenya	Kikuyu	Kikuyu (Niger-Kordofaniano)	25	AFRE	Watson <i>et al.</i> 1996	HVRbase	
	Somalia	Somali	Somali (Afro-asiático)	3	AFRE	Watson <i>et al.</i> 1996	HVRbase	
				<b>90</b>				
		Benin	Fulbe	Fulfulde (Niger-Kordofaniano)	1	AFRO	Watson <i>et al.</i> 1996	HVRbase
		Burkina Faso	Fulbe	Fulfulde (Niger-Kordofaniano)	2	AFRO	Watson <i>et al.</i> 1996	HVRbase
		Camerún	Fulbe	Fulfulde (Niger-Kordofaniano)	1	AFRO	Watson <i>et al.</i> 1996	HVRbase
África Oeste	Mali	Tuarég	Tamasheq (Afro-asiático)	1	AFRO	Watson <i>et al.</i> 1996	HVRbase	
	Mali	Songhai	Songhai (Nilo-Sahariano)	1	AFRO	Watson <i>et al.</i> 1996	HVRbase	
	Niger	Fulbe	Fulfulde (Niger-Kordofaniano)	10	AFRO	Watson <i>et al.</i> 1996	HVRbase	
	Niger	Songhai	Songhai (Nilo-Sahariano)	9	AFRO	Watson <i>et al.</i> 1996	HVRbase	
	Niger	Tuarég	Tamasheq (Afro-asiático)	25	AFRO	Watson <i>et al.</i> 1996	HVRbase	
	Niger	Hausa	Hausa (Afro-asiático)	3	AFRO	Watson <i>et al.</i> 1996	HVRbase	
	Niger	Kanuri	Kanuri (Nilo-Sahariano)	1	AFRO	Watson <i>et al.</i> 1996	HVRbase	
	Nigeria	Fulbe	Fulfulde (Niger-Kordofaniano)	47	AFRO	Watson <i>et al.</i> 1996	HVRbase	
	Nigeria	Hausa	Hausa (Afro-asiático)	17	AFRO	Watson <i>et al.</i> 1996	HVRbase	
	Nigeria	Kanuri	Kanuri (Nilo-Sahariano)	13	AFRO	Watson <i>et al.</i> 1996	HVRbase	
	Nigeria	Yoruba	Yoruba (Niger-Kordofaniano)	21	AFRO	Watson <i>et al.</i> 1996	HVRbase	
	Mauritania	Mauritania		30	AFRO	Rando <i>et al.</i> 1998	HVRbase	
	Sahara Oeste	Sahara Oeste		25	AFRO	Rando <i>et al.</i> 1998	HVRbase	
	Senegal	Senegal		121	AFRO	Rando <i>et al.</i> 1998	HVRbase	
	Senegal E	Mandenka		105	AFRO	Graven <i>et al.</i> 1995	HVRbase	
				<b>433</b>				
	África Centro	República Centroafricana	Pigmeo Biaka	Bantú (Niger-Kordofaniano)	17	AFRC	Vigilant <i>et al.</i> 1991	HVRbase
Zaire		Pigmeo Mbuti	Bantú (Niger-Kordofaniano)	13	AFRC	Vigilant <i>et al.</i> 1991	HVRbase	
			<b>30</b>					
África Sur	Bostwana	!Kung	Zhu/Twasi (Khoisanido)	19	AFRS	Vigilant <i>et al.</i> 1991	HVRbase	
Alemania	Alemania Rhin		Alemán (Indo-hitita)	50	AL	Baasner <i>et al.</i> 1998	HVRbase	
Andalucía	Andalucía, España		Castellano (Indo-hitita)	15	AND	Corte-Real <i>et al.</i> 1996	HVRbase	
Albania	Albania		Albano (Afro-asiático)	42	ALB	Belledi <i>et al.</i> 2000	HVRbase	
Austria	Austria Oeste		Alemán (Indo-hitita)	100	AUS	Parson <i>et al.</i> 1998	HVRbase	

Tabla ME36. (Continuación).

Procedencia	Grupo	Lengua	N	Identificador	Referencia	Base datos
Bélgica	Bruselas, Bélgica	Alemán (Indo-hitita)	25	BEL	Decorie <i>et al.</i> 1996	FHG Base
Bulgaria	Bulgaria	Báltico-Eslavo (Indo-hitita)	30	BUL	Calafell <i>et al.</i> 1996	HVRbase
Canarias	Tenerife, España	Castellano (Indo-hitita)	54	CAN	Pinto <i>et al.</i> 1996	HVRbase
Cataluña	Cataluña, España	Catalán (Indo-hitita)	15	CAT	Corte-Real <i>et al.</i> 1996	HVRbase
Dinamarca	Copenaghe, Dinamarca	Danés (Indo-hitita)	32	DEN	Richards <i>et al.</i> 1996	FHG Base
Drusos Israel	Israel	Semítico (Afro-asiático)	45	DRUZ	Macaulay <i>et al.</i> 1999a	FHG Base
Estonia	Estonia	Estonio/Finno-Ugric (Uralic-Yukaghir)	26	EST	Sajantila <i>et al.</i> 1995	HVRbase
Finlandia	Finlandia	Finlandés/Finno-Ugric (Uralic-Yukaghir)	147	FIN	Sajantila <i>et al.</i> 1995	HVRbase
Francia	Francia	Francés (Indo-hitita)	50	FR	Rousselet <i>et al.</i> 1998	HVRbase
Galicia	Galicia, España	Gallego (Indo-hitita)	90	GAL	Salas <i>et al.</i> 1998	HVRbase
Georgia	Georgia	Georgio (Sur-Caucásico)	54	GEO	Nasidze y Stoneking 2001	HVRbase
Gran Bretaña	Gran Bretaña	Inglés (Indo-hitita)	97	GB	Piercy <i>et al.</i> 1993	HVRbase
Grecia	Atenas, Grecia	Griego (Indo-hitita)	45	GRE	FHG Base, P.c.	FHG Base
Cerdeña	Cerdeña, Italia	Sardo (Indo-hitita)	67	SAR	DiRienzo y Wilson 1991	HVRbase
Toscana	Toscana	Italiano (Indo-hitita)	49	TUS	Francalacci <i>et al.</i> 1996	HVRbase
Italia Sur	Basilicata, Italia	Italiano (Indo-hitita)	37	ITSU	Richards O. (P.c.)	FHG Base
Sicilia	Sicilia, Italia	Italiano (Indo-hitita)	63	SIC	Richards O., Nigro L. (P.c.)	FHG Base
Kurdistán	Tbilisi, Georgia	Irani (Indo-hitita)	29	KUR	Comas <i>et al.</i> 2000	FHG Base
Noruega	Este		216	NOR	Opdal <i>et al.</i> 1998	HVRbase
	Mansoura, Egipto	Arábigo (Afro-asiático)	40	NIL	Krings <i>et al.</i> 1999a	HVRbase
	El Cairo, Egipto	Arábigo (Afro-asiático)	10	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Sohag, Egipto	Arábigo/Cóptico (Afro-asiático)	3	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Kena, Egipto	Arábigo/Cóptico (Afro-asiático)	2	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Tarna, Egipto	Arábigo (Afro-asiático)	1	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Chephen, Egipto	Arábigo (Afro-asiático)	1	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Monofia, Egipto	Arábigo (Afro-asiático)	1	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Assiut, Egipto	Arábigo/Cóptico (Afro-asiático)	21	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Luxor, Egipto	Arábigo (Afro-asiático)	1	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Mimia, Egipto	Arábigo/Cóptico (Afro-asiático)	2	NIL	Krings <i>et al.</i> 1999a	HVRbase
	El-Sawat, Egipto	Arábigo (Afro-asiático)	1	NIL	Krings <i>et al.</i> 1999a	HVRbase
	El-Sharkia, Egipto	Arábigo (Afro-asiático)	1	NIL	Krings <i>et al.</i> 1999a	HVRbase
	New Valley, Egipto	Arábigo (Afro-asiático)	1	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Assuan, Egipto	Nubio (Nilo-Saharaniano)	11	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Halfa, Sudán	Nubio (Nilo-Saharaniano)	1	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Dongola, Sudán	Nubio (Nilo-Saharaniano)	14	NIL	Krings <i>et al.</i> 1999a	HVRbase

Tabla ME36. (Continuación).

Población	Procedencia	Grupo	Lengua	N	Identificador	Referencia	Base datos
Valle del Nilo	Kerma, Sudán	Nubio	Nubio (Nilo-Sahariano)	39	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Hadandawa, Sudán	Nubio	Nubio (Nilo-Sahariano)	1	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Falata, Sudán	Nubio	Nubio (Nilo-Sahariano)	1	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Sikut, Sudán	Nubio	Nubio (Nilo-Sahariano)	1	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Mahas, Sudán	Nubio	Nubio (Nilo-Sahariano)	3	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Sudán	Nubio	Nubio (Nilo-Sahariano)/Niger-Kordofanian	7	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Sudan Sur	Nuer	Nilótico (Nilo-Sahariano)	11	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Sudan Sur	Schilluk	Nilótico (Nilo-Sahariano)	7	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Sudan Sur	Dinka	Nilótico (Nilo-Sahariano)	40	NIL	Krings <i>et al.</i> 1999a	HVRbase
	Sudan Sur	Nuba	Niger-Kordofanian	14	NIL	Krings <i>et al.</i> 1999a	HVRbase
			<b>235</b>				
Jordania	Amman, Jordania		Semítico (Afro-asiático)	42	JOR	DiRienzo y Wilson 1991	FHG Base
País Vasco	País Vasco, España		Vasco	45	PV	Bertranpetit <i>et al.</i> 1995	HVRbase
	Portugal		Portugués (Indo-hitita)	54	POR	Corte-Real <i>et al.</i> 1996	HVRbase
Rusia Karelian	Petrozavodsk, Rusia	Karelians	Karelian/Finno-Ugric (Uralic-Yukaghir)	83	RKAR	Sajantila <i>et al.</i> 1995	HVRbase
Rusia Norte Cáucaso	Sochynskiy, Rusia	Abazinian	Abazin (Norte-Caucásico)	23	RNC	Nasidze y Stoneking 2001	HVRbase
	Chechenia, Rusia	Checheno	Checheno (Norte-Caucásico)	23	RNC	Nasidze y Stoneking 2001	HVRbase
	Karachay-Cherkessia, Rusia	Cherkessia	Norte-Caucásico	44	RNC	Nasidze y Stoneking 2001	HVRbase
	Dagestán, Rusia	Darginia	Dargin (Norte-Caucásico)	37	RNC	Nasidze y Stoneking 2001	HVRbase
	Kabardino-Balkaria, Rusia	Kabardmia	Norte-Caucásico	51	RNC	Nasidze y Stoneking 2001	HVRbase
					<b>178</b>		
Saami	Inari Lake, Finlandia	Saami	Saamic/ Finno-Ugric (Uralico-Yukaghir)	22	SAA	Sajantila <i>et al.</i> 1995	HVRbase
	Skolt, Finlandia	Saami	Saamic/ Finno-Ugric (Uralico-Yukaghir)	47	SAA	Sajantila <i>et al.</i> 1995	HVRbase
	Karajjok, Noruega	Saami	Saamic/ Finno-Ugric (Uralico-Yukaghir)	21	SAA	Sajantila <i>et al.</i> 1995	HVRbase
				<b>90</b>			
Siria	Siria		Semítico (Afro-asiático)	49	SIR		FHG Base
Suecia	Lund, Suecia		Sueco (Uralico-Yukaghir)	24	SUEC	Kittles <i>et al.</i> 1999	HVRbase
	Alpes Norte, Suiza		Alemán (Indo-hitita)	7	SUIZ	Pult <i>et al.</i> 1994	HVRbase
Suiza	Alpes Sur, Suiza		Suizo-Italiano (Indo-hitita)	5	SUIZ	Pult <i>et al.</i> 1994	HVRbase
	Tessin, Suiza		Italiano (Indo-hitita)	5	SUIZ	Pult <i>et al.</i> 1994	HVRbase
				<b>17</b>			
Turquia	Peninsula Anatolia		Turco (Altaico)	44	TUR	Comas <i>et al.</i> 1996	HVRbase
	Turquia		Turco (Altaico)	28	TUR	Calafell <i>et al.</i> 1996	HVRbase
				<b>72</b>			
<b>Total</b>				<b>2839</b>			

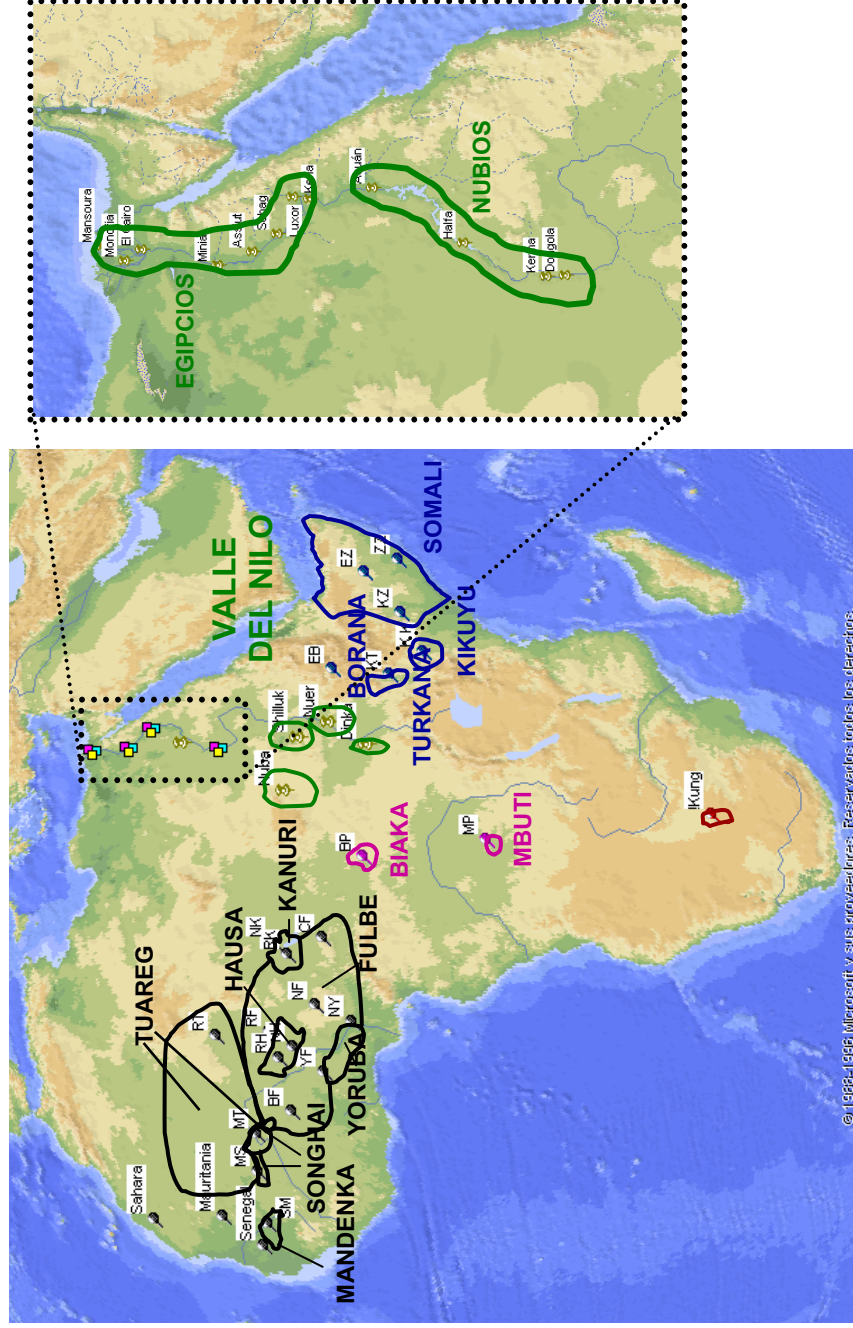


Figura FME14. Localización de las poblaciones y grupos étnicos de la base de datos de secuencias actuales utilizada en el presente trabajo.



**Figura FME15.** Localización de las poblaciones y grupos étnicos de la base de datos de secuencias actuales utilizada en el presente trabajo.

## 16.2. Subconjuntos de población antigua

Para poder establecer una comparación entre el conjunto de poblaciones actuales de la base de datos y las secuencias antiguas obtenidas éstas fueron agrupadas en cuatro subconjuntos atendiendo a criterios geográficos y cronológicos (**Tabla ME37**).

**Tabla ME37.** Subconjuntos de población antigua creados para la comparación con poblaciones actuales.  
(\*) Muestras de yacimientos españoles.

Población	Yacimientos	Procedencia	Datación	Nº secuencias		
				244 pb	133pb	112pb
Paleolítico	El Pirulejo	Córdoba*	13500-12500 A.P.	2	2	2
	Nerja	Málaga*	20000-17000 A.P.	2	2	2
	Toledo	Portugal	9200-7800 A.P.	1	1	1
	<b>Total</b>			<b>5</b>	<b>5</b>	<b>5</b>
Neolítico	Abauntz	Navarra*	4240 A.P.	3	5	5
	Tres Montes	Navarra*	4130 A.P.	5	6	5
	Atxuri	Vizcaya*	4400-3500/3200 A.P.	-	-	1
	Nerja	Málaga*	4260±70 A.P. 5.875±80 A.P.	2	2	2
	Caldeirao	Portugal	6800-6200 A.P.	-	-	1
	Tell Halula	Siria	10000 A.P.	7	12	16
	Mari	Siria	10000 A.P.	6	8	8
	Tell Ramad	Siria	10000 A.P.	6	7	7
	Dja'de	Siria	10000 A.P.	1	1	3
	<b>Total</b>			<b>30</b>	<b>41</b>	<b>48</b>
	Neolítico	Tell Halula	Siria	8800-8000 A.P.	7	12
Próximo Oriente	Tell Ramad	Siria	8000-7750 A.P.	6	7	7
	Dja'de	Siria	9400-8700 A.P.	1	1	3
<b>Total</b>			<b>14</b>	<b>20</b>	<b>26</b>	
Neolítico Península Ibérica	Abauntz	Navarra*	4240 A.P.	3	5	5
	Tres Montes	Navarra*	4130 A.P.	5	6	5
	Atxuri	Vizcaya*	4400-3500/3200 A.P.	-	-	1
	Nerja	Málaga*	4260±70 A.P. 5.875±80 A.P.	2	2	2
	Caldeirão	Portugal	6800-6200 A.P.	-	-	1
<b>Total</b>			<b>10</b>	<b>13</b>	<b>14</b>	

Los diferentes subconjuntos comprenden la mayor parte de secuencias autenticadas correspondientes al fragmento completo de 244pb así como de los fragmentos parciales de 133pb y 112pb detalladas en la **Tabla R62** del capítulo “Resultados”. No se incluyeron las secuencias de AT19 y AT27 por no comprender la totalidad de ninguno de los fragmentos analizados. La secuencia de GA22, aunque completa para el segmento de 112pb, tampoco fue incluida por proceder de un período posterior a las muestras restantes.



La población “Paleolítico” está constituida por secuencias de época Paleolítica de los yacimientos de El Pirulejo y Nerja y Mesolítica del yacimiento de Toledo. Este conjunto pretende reflejar la variabilidad existente en la Península Ibérica antes del Neolítico.

La población “Neolítico” incluye a individuos de época Neolítica (yacimientos de Tell Halula, Tell Ramad, Dja’de, Nerja y Caldeirão) y post-Neolítica (yacimientos de Abauntz, Tres Montes, Atxuri y Mari). Una vez subdividida en base a su procedencia, el conjunto de Oriente Próximo comprende a individuos estrictamente Neolíticos (Tell Halula, Tell Ramad y Dja’de) mientras que el de la Península Ibérica reúne a individuos de época Neolítica (Nerja, Caldeirão), del Calcolítico (Abauntz, Tres Montes) y de la Edad del Bronce (Atxuri).

### **16.3. Índices de diversidad**

Se calcularon un conjunto de índices de diversidad separadamente para los fragmentos de 244pb, 133pb y 112pb dentro de cada yacimiento y en los diferentes subconjuntos de población antigua.

Todos los índices se computaron con el programa Arlequin Ver. 2.000 (SCHNEIDER *et al.* 2000) con la opción *Standard diversity indices* dentro del apartado *Diversity indices*.

#### **16.3.1. Número y porcentaje de sitios polimórficos**

El número de sitios variables o polimórficos ( $v$ ) se define como el número de *loci* útiles que presentan más de un alelo por *locus* (SCHNEIDER *et al.* 2000). En el caso de secuencias refleja el número de posiciones que varían entre el conjunto de secuencias alineadas. Este índice es claramente dependiente del tamaño de la secuencia estudiada, lo que puede corregirse dividiendo el número de posiciones variables entre el número de nucleótidos de la secuencia (porcentaje de sitios polimórficos o  $\%v$ ).

#### **16.3.2. Haplotipos diferentes**

Cada variante mitocondrial definida respecto a la secuencia consenso (ANDERSON *et al.* 1981) constituye un haplotipo mitocondrial ( $k$ ). El porcentaje de haplotipos diferentes ( $\% k$ ) se estimó en relación al número de individuos de cada yacimiento y población.

#### **16.3.3. Diversidad haplotípica**

La diversidad genética de Nei (NEI 1987) se define como la probabilidad de que dos alelos de una muestra escogidos al azar sean diferentes. Cuando se trabaja con secuencias cada variante (haplotipo) equivale a un alelo, por lo que puede utilizarse este índice para evaluar la diversidad haplotípica del conjunto. Este estimador y su varianza asociada se calculan de la siguiente forma:

$$H = \frac{n}{(n-1)} \left( 1 - \sum_{i=1}^k p_i^2 \right)$$

$$V(H) = \frac{2}{n(n-1)} \left\{ 2(n-2) \left[ \sum_{i=1}^k p_i^3 - \left( \sum_{i=1}^k p_i^2 \right)^2 \right] + \sum_{i=1}^k p_i^2 - \left( \sum_{i=1}^k p_i^2 \right)^2 \right\}$$

dónde  $n$  es el número de copias del gen,  $k$  es el número de alelos (haplotipos) y  $p_i$  es la frecuencia en la muestra del alelo (o haplotipo)  $i$ .

#### 16.3.4. Número medio de diferencias por parejas y diversidad nucleotídica

El número medio de diferencias por parejas de secuencias se calcula:

$$\pi = \sum_{i=1}^k \sum_{j<i} p_i p_j d_{ij}$$

dónde  $p_i$  y  $p_j$  son las frecuencias en la muestra de los haplotipos  $i$  y  $j$  respectivamente y  $d_{ij}$  un estimador del número de mutaciones acontecidas desde la divergencia de dichos haplotipos.

La varianza total asociada a este índice, asumiendo neutralidad y ausencia de recombinación, es:

$$V(\pi) = \frac{3n(n+1)\pi + 2(n^2 + n + 3)\pi^2}{11(n^2 - 7n + 6)}$$

La diversidad nucleotídica es análoga a la diversidad genética de Nei pero al nivel de nucleótido. Este índice describe la probabilidad de que dos nucleótidos homólogos escogidos al azar de entre un conjunto de secuencias sean diferentes. La diversidad nucleotídica puede calcularse como el número medio de diferencias entre parejas de secuencias ( $\pi$ ) dividido por la longitud en pares de bases de la secuencia estudiada (TAJIMA 1993; NEI 1987). Las fórmulas correspondientes para este estimador y su varianza asociada, asumiendo neutralidad y ausencia de recombinación, son por lo tanto:

$$\pi_n = \frac{\sum_{i=1}^k \sum_{j<i} p_i p_j d_{ij}}{L}$$

$$V(\pi_n) = \frac{n+1}{3(n-1)L} \pi_n + \frac{2(n^2+n+3)}{9n(n-1)} \pi_n^2$$

dónde  $p_i$  y  $p_j$  son las frecuencias en la muestra de los haplotipos  $i$  y  $j$ ,  $d_{ij}$  un estimador del número de mutaciones acontecidas desde la divergencia de dichos haplotipos y  $L$  la longitud de la secuencia analizada.

#### 16.4. Distribución de las diferencias por parejas (*Mismatch distribution*)

La representación de la distribución de las diferencias por parejas en un histograma además de resumir la diversidad genética de una población puede reflejar su historia demográfica. Según Rogers y Harpending (1992) los episodios de crecimiento o declive de una población dejan huellas discernibles en la distribución de las diferencias de nucleótidos entre pares de individuos.

La distribución de las diferencias por parejas de los diferentes fragmentos de mtDNA en los diferentes subconjuntos de población antigua y en algunas poblaciones actuales se calculó con el programa DnaSP 3.5 (Rozas y Rozas 1999) con la opción *Population Size Changes* comprendida dentro del menú *Analysis*. La distribución esperada según un modelo de expansión súbita (ROGERS y HARPENDING 1992) fue calculada con este mismo programa.

El modelo de expansión súbita (ROGERS y HARPENDING 1992) se basa en tres parámetros:

- $\theta_0$ : Theta inicial (anterior al episodio de crecimiento o declive de la población)
- $\theta_1$ : Theta Final (posterior al episodio de crecimiento o declive)
- $\tau$ : Fecha del episodio expresada en unidades de tiempo mutacional.  $\tau = 2\mu t$  siendo  $t$  el tiempo en generaciones y  $\mu$  la tasa de mutación por secuencia y generación (ROGERS y HARPENDING 1992).

Considerando  $\theta_1$  igual a infinito es posible estimar  $\theta_0$  y  $\tau$  a partir de los datos (ROGERS 1995):

$$\theta_0 = \sqrt{v - m}$$

donde  $m$  y  $v$  son respectivamente la media y la varianza de la distribución observada de las diferencias por parejas.

Estos dos parámetros se calcularon con el programa DnaSP a partir de los datos y se usaron para calcular la distribución esperada. Una vez computadas ambas

distribuciones se representaron gráficamente con el programa EXCEL del paquete MICROSOFT OFFICE 2000.

Para evaluar el ajuste de la distribución observada y la esperada según este modelo se realizó una prueba de bondad de ajuste. El estadístico  $\chi^2$  y su probabilidad asociada para un nivel de confianza del 95% se computaron con el programa EXCEL.

Para los diferentes subconjuntos de población antigua y todas las poblaciones actuales estudiadas se calculó también el índice de irregularidad de Harpending (r) (HARPENDING 1994) con el programa DnaSP 3.5. Este índice es simplemente la suma de los cuadrados de las diferencias entre clases contiguas de frecuencia:

$$r = \sum_{i=1}^{d+1} (x_i - x_{i-1})^2$$

$d$  es el número máximo de diferencias observadas entre haplotipos y  $x_i$  es la frecuencia relativa observada del  $i$ -ésimo número de diferencias.

El índice de irregularidad de Harpending estima la irregularidad de la distribución de las diferencias por parejas. Generalmente aquellas distribuciones acampanadas presentan valores más bajos de  $r$ , inferiores a 0.03 en el caso de secuencias, que aquellas que muestran un patrón multimodal (JOBBLING *et al.* 2003).

## 16.5. Análisis filogenético interpoblacional

### 16.5.1. Distancias genéticas

A partir de la frecuencia de los haplotipos presentes en las poblaciones actuales y en los diferentes subconjuntos de población antigua se calculó la distancia genética de Reynolds entre pares de poblaciones (REYNOLDS *et al.* 1983). Esta distancia asume que no existe mutación y que la deriva genética es el único factor responsable del cambio en las frecuencias de los haplotipos. A diferencia de otras distancias como la distancia genética de Nei (NEI 1972) esta medida no asume un tamaño efectivo de población constante e igual para todas las poblaciones. De hecho, las fluctuaciones en el tamaño poblacional juegan un importante papel en la magnitud de la distancia puesto que pueden modificar el efecto de la deriva genética.

La distancia genética de Reynolds (REYNOLDS *et al.* 1983) se calcula:

$$D = \frac{\sum_m \sum_i (p_{1mi} - p_{2mi})^2}{2 \sum_m \left( 1 - \sum_i p_{1mi} - p_{2mi} \right)}$$

donde  $m$  es el número de loci,  $i$  el número total de alelos del conjunto de loci y  $p_{1mi}$  y  $p_{2mi}$  las frecuencias del alelo  $i$  del locus  $m$  en las poblaciones 1 y 2 respectivamente. En el caso de secuencias de DNA mitocondrial el número de loci es igual a 1 y el número de alelos equivale al número de haplotipos de la muestra.

El cálculo de la matriz de distancias de Reynolds entre los pares de poblaciones se realizó con el programa GENDIST del paquete PHYLIP 3.5c (FELSENSTEIN 1993). La matriz de frecuencias de haplotipos fue calculada con el programa ARLEQUIN a partir de las secuencias en bruto, con la opción *Haplotype frequencies* contenida en el apartado *Diversity indices*. Posteriormente esta matriz se editó con el programa TEXTPAD 4.7.1. (HELIOS SOFTWARE SOLUTIONS 1992-2003) para adecuarla al formato de entrada de datos de PHYLIP.

### 16.5.1. Reconstrucción filogenética a partir de distancias genéticas

A partir de la matriz de distancias de Reynolds (REYNOLDS *et al.* 1983) entre todas las poblaciones estudiadas generada con el programa GENDIST del paquete PHYLIP se construyeron dos árboles sin raíz utilizando los métodos de reconstrucción filogenética de Neighbor-Joining (SAITOU y NEI 1987) y de Fitch y Margoliash (1967) y un árbol con raíz utilizando el método de UPGMA (SNEATH y SOKAL 1973).

El método de Neighbor-Joining (SAITOU y NEI 1987) se basa en el agrupamiento sucesivo (*clustering*) de las diferentes líneas para la construcción del árbol, asignando la longitud de las ramas a medida que éstas se van uniendo. Este método de reconstrucción no asume reloj molecular, por lo que los árboles generados con él no presentan raíz. Para la construcción de un árbol de este tipo a partir de la matriz de distancias de Reynolds entre las diferentes poblaciones actuales y los subconjuntos de población antigua se empleó el programa NEIGHBOR del paquete PHYLIP 3.5c (FELSENSTEIN 1993).

En la reconstrucción de UPGMA (*Unweighted Pair-Group Method with Arithmetic Mean*) (SNEATH y SOKAL 1973) se produce una agrupación iterativa de aquellas líneas que presentan una menor distancia entre sí. Este método asume la existencia de reloj molecular y produce un árbol aditivo conocido como *árbol ultramétrico*, en el que la distancia entre cada taxón y la raíz es la misma. Este método asume una misma tasa de evolución para todos los taxones. Para la construcción de este árbol se incorporó al conjunto de poblaciones la población “Neandertal”, compuesta por tres secuencias de tres individuos Neandertales (KRINGS *et al.* 1997; OVCHINNIKOV *et al.* 2000 y KRINGS *et al.* 2000).

Este tipo de reconstrucción se realizó también con el programa NEIGHBOR del paquete PHYLIP 3.5c (FELSENSTEIN 1993) cambiando la opción N de Neighbor-joining a UPGMA.

El método de Fitch- Margoliash es otro método de agrupación iterativa similar al método de Neighbor-Joining en el que tampoco se asume la existencia de un reloj



molecular, por lo que los árboles generados no presentan raíz. La diferencia entre este método y los dos anteriores es el peso que otorga a la desviación entre lo esperado y lo observado. En el método de Fitch-Margoliash la varianza de la distancia observada es proporcional al cuadrado del valor esperado. Para la construcción de un árbol de este tipo a partir de la matriz de distancias de Reynolds se empleó el programa FITCH del paquete PHYLIP 3.5c (FELSENSTEIN 1993). Fueron seleccionadas las opciones G (Global rearrangements) y J (seed: 245, 10 times). La primera de estas opciones permite eliminar los sub-árboles del árbol para después recolocarlos de diferentes maneras, lo que aumenta la probabilidad de encontrar el mejor árbol posible. La segunda permite variar el orden de entrada de las poblaciones para la construcción del árbol a través de la introducción de un número aleatorio (*seed*). Puede seleccionarse también el número de veces que se desea repetir este proceso.

Una vez generados los árboles se visualizaron y editaron con el programa TREEVIEW (Win32) 1.6.6 (PAGE 2001).

## 16.6. Clusters

### 16.6.1. Poblaciones estudiadas

A partir de las secuencias de mtDNA antiguas obtenidas y las de nuestra base de datos de poblaciones actuales se realizó una clasificación de las poblaciones mediante la técnica de las  $\varepsilon$ -clases.

Los conjuntos de población actual empleados en esta clasificación se corresponden con los descritos en la **Tabla ME36**. Para constituir las poblaciones antiguas se agruparon 30 de las secuencias obtenidas, procedentes de 6 yacimientos arqueológicos, en cuatro grupos en base a su origen geográfico y a su cronología (**Tabla ME38**).

**Tabla ME38.** Composición de las poblaciones antiguas diseñadas para el análisis de clusters.

Yacimiento	Muestras	Individuos	Población	Identificador	N
Abauntz	AB5	ABb1.1			
Abauntz	AB9	AB23C			
Abauntz	AB17	AB35E			
Tres Montes	1TM2	TM-3	Calcolítico	CALC	8
Tres Montes	2TM4	TM-6			
Tres Montes	1TM5	TM-8			
Tres Montes	1TM6	TM-11			
Tres Montes	TM11	TM-18			
Nerja	2NE	NE-NM82.2	Paleolítico-Neolítico Península Ibérica	PNEPI	6
Nerja	3NE	NE-NT79			
Nerja	4NE	NE-NAP			
Nerja	5NE	NE-1829			
El Pirulejo	1PI	1PI			
El Pirulejo	2PI	2PI			
Tell Halula	TH2	H-3	Neolítico Próximo Oriente	NEOP	10
Tell Halula	TH5	H-7			
Tell Halula	H12	H-68			
Tell Halula	1H17	H-45			
Tell Halula	2H20	H-37			
Tell Ramad	TR8	R65-3I			
Tell Ramad	TR12	R65-14			
Tell Ramad	TR16	R65-C8-SEB			
Tell Ramad	TR18	R65-1S			
Tell Ramad	TR19	R65-N4-N°400			
Mari	1MA1	1	Mari	MARI	6
Mari	2MA2	2			
Mari	1MA5	5			
Mari	2MA10	10			
Mari	3MA11	11			
Mari	1MA12	12			
<b>Total</b>					<b>30</b>

### 16.6.2. Descripción del método

La construcción de éstas  $\varepsilon$ -clases se realiza en base a una desemejanza entre pares de poblaciones. La función de desemejanza utilizada es la distancia media de Hadamard medida entre secuencias individuales para todos los pares de poblaciones a comparar según la expresión:

$$D(P_1, P_2) = \frac{1}{\text{Car}(P_1)\text{Car}(P_2)} \sum_{i \in P_1} \sum_{j \in P_2} \sum_{l=1}^{243} \delta |X_{il} - X_{jl}|$$

dónde  $l$  es la posición  $l$ -ésima de la secuencia,  $i$  es el individuo  $i$ -ésimo de la población  $P_1$ ,  $j$  es el individuo  $j$ -ésimo de la población  $P_2$  y  $\delta$  es la  $\delta$  de Dirac.

En función de la desemejanza media de Hadamard definida entre las poblaciones se construyen los “clusters” a nivel  $\varepsilon$ , formados por todos los subconjuntos del conjunto de poblaciones, de tal forma que la desemejanza entre dos poblaciones cualesquiera de cada subconjunto o “cluster” es menor que  $\varepsilon$ . Los “cluster” obtenidos proporcionan una clasificación, hasta cierto punto difusa, de las poblaciones en función de la desemejanza. Así, para una desemejanza dada, se construyen “clusters” distintos para cada nivel de  $\varepsilon$ . Cuando  $\varepsilon$  es suficientemente grande, se forma un único “cluster” con la totalidad de poblaciones. A fin de reducir el número de “clusters”, se impuso que, dados dos “clusters” A y B, donde  $\text{Car}(A) < \text{Car}(B)$  si se cumple que:

$$F = \frac{\text{Car}(A) \cap \text{Car}(B)}{\text{Car}(A)} \geq 0.85$$

y además que

$$\text{Car}(A) \cap \text{Car}(B) > 6$$

entonces los “clusters” A y B son fusionados en una misma clase.

Todos los cálculos se realizaron en un programa en FORTRAN basado en el método de las  $\varepsilon$ -clases (SÁNCHEZ, 1978) especialmente diseñado para el análisis de secuencias de mtDNA por el Catedrático de Estadística de la Facultad de Medicina de la Universidad Complutense de Madrid, Dr. Miguel Sánchez.

## 16.7. Haplotipos

### 16.7.1. Reconstrucción filogenética a partir de haplotipos

Para la reconstrucción filogenética a partir de los haplotipos de mtDNA de las poblaciones actuales y los obtenidos de las muestras antiguas se emplearon dos procedimientos, uno basado en las distancias genéticas entre los haplotipos y otro en la construcción de Median Joining Networks.

#### 16.7.1.1. Distancias genéticas entre los haplotipos

Para el cálculo de la matriz de distancias entre los haplotipos se utilizó el método de Tamura y Nei (1993). Este método es un modelo de evolución molecular desarrollado a partir de datos de la región control del mtDNA que según ciertos autores resulta el más verosímil para el segmento hipervariable I (WEISS y VON HAESELER 1998). Se utilizó la corrección de la distribución gamma de la tasa de mutación mediante el parámetro  $\alpha=0.47$  estimado por WAKELEY 1993 para este mismo segmento. Para clarificar la representación filogenética y evitar en la medida de lo posible la distorsión de la estructura ancestral del árbol por aquellas variantes más recientes (RICHARDS *et al.* 1996), de los 1118 haplotipos diferentes presentes en la muestra de población actual, únicamente se incluyeron aquellos con una frecuencia absoluta superior a dos, asumiendo que en una población que ha sufrido un proceso reciente de expansión aquellos haplotipos más frecuentes suelen ser los más antiguos (DONNELLY y TAVARÉ 1986).

Para la construcción del árbol UPGMA se incluyeron también dos secuencias de dos individuos de Neandertal (KRINGS *et al.* 1997; OVCHINNIKOV *et al.* 2000) a modo de *outgroup*.

A partir de esta matriz de distancias se construyeron dos árboles filogenéticos, uno sin raíz mediante la reconstrucción de Neighbor-Joining y otro con raíz mediante el método de UPGMA. El cálculo de las distancias, la construcción de los dos árboles filogenéticos y su visualización fueron realizadas íntegramente con el programa MEGA Versión 2.1. (KUMAR *et al.* 1993).

### 16.7.1.2. Median Joining Networks

Se construyeron dos árboles filogenéticos con el algoritmo de Median Joining Network a partir de la mayor parte de los haplotipos de población actual de frecuencia mayor a dos y todos los recuperados de las muestras antiguas. En la primera de estas reconstrucciones se agruparon aquellos haplotipos de poblaciones europeas mientras que en la segunda se representaron aquellos pertenecientes a haplogrupos africanos y asiático-americanos.

Ambas reconstrucciones se realizaron con el programa Network Ver. 4.000 empleando la opción *Median Joining* (BANDELT *et al.* 1999) del menú *Calculation*. La asignación de pesos a las diferentes posiciones se realizó de acuerdo con RICHARDS *et al.* 1998 con algunas modificaciones. Se otorgó un peso igual a 1 a aquellas posiciones de elevada tasa de mutación (16129, 16189, 16192, 16311 y 16362), igual a 2 a aquellas posiciones de tasa de mutación moderada (16209, 16223, 16224, 16278, 16293, 16294, 16304, 16325 y 16356), e igual a 4 para el resto de posiciones estudiadas.

La relación entre los diferentes haplogrupos y sub-haplogrupos europeos y africanos ha sido establecida en poblaciones actuales a partir del estudio de un gran número de secuencias (RICHARDS *et al.* 1998, 2000; SALAS *et al.* 2002). La construcción de un árbol Median Joining con secuencias antiguas y actuales persigue un doble objetivo: confirmar la asignación previa de haplogrupo a las secuencias obtenidas en el presente trabajo y evaluar las relaciones entre las variantes antiguas y las actuales. Dado que el conjunto de secuencias antiguas recuperadas es demasiado pequeño como para permitir una reconstrucción filogenética de este tipo *per se*, se consideró a efectos prácticos que la relación filogenética deducida de los estudios de población actual era correcta, procediéndose a la simple inclusión de las secuencias antiguas en el entramado de secuencias actuales. Este supuesto no tiene que ser necesariamente cierto como ya desarrollaremos más ampliamente en el apartado “Discusión”, pero se ha considerado el más acertado dada la resolución de las secuencias antiguas disponibles.

Para reproducir una relación entre los haplotipos de población actual lo más clara posible y acorde a la obtenida por otros investigadores (RICHARDS *et al.* 1998,

2000; SALAS *et al.* 2002) se hizo necesario eliminar algunos de ellos. Así pues, para la construcción del árbol de secuencias europeas se eliminaron los haplotipos 271, 94, 99, 65, 201, 199, 48, 446, 91, 277, 204, 168, 764, 448, 69 y 17 (haplogrupo H); 62 y 564 (H/HV/HV1); 211, 205 (H/U); 587 (H/W); 14 (H/HV\*/pre-HV/R\*/U\*); 44 (H/HV1/U\*); 268 (R1); 566 (U1a); 852, 745 y 453 (U3); 123, 456, 739 y 193 (U5); 571(X); 107, 1004, 337, 445, 912, 104, 261 y 538 (haplogrupo desconocido). En este mismo árbol para una correcta resolución del cuadrado 16126C-16294T; 16126C-16294T-16296T; 16126C-16294T-16296T-16304C y 16126C-16294T-16296T-16304C se asignó peso 0 a la posición 16296 tal y como se sugiere en (RICHARDS *et al.* 2000). En la construcción del árbol de secuencias africanas y asiático-americanas únicamente se eliminó el 462 perteneciente al haplogrupo L3.

Una vez calculado el Median Joining Network, se generó el Network que contenía aquellos árboles más cortos seleccionando la primera opción dentro del submenú *MP calculation* del menú *Calculation*. Esta opción emplea como *input* el fichero *output* resultante del cálculo de Median Joining (fichero \*.out) generando a su vez un segundo fichero de salida (fichero \*.sto). La apertura de este fichero con el menú *Draw Network* produce un árbol “resumen” cuyo torso puede ser modificado por el usuario. Para los dos Median Joining Network generados se utilizó la información de estudios previos realizados con poblaciones actuales europeas (RICHARDS *et al.* 1998, 2000) y africanas (SALAS *et al.* 2002) para la elección del árbol más adecuado.

### **16.7.2. Haplotipos compartidos entre poblaciones antiguas y actuales**

Otra forma sencilla de comparar la información obtenida de las secuencias antiguas y el conjunto actual de poblaciones es el estudio de la frecuencia y distribución de los haplotipos de la muestra antigua en las diferentes poblaciones actuales de la base de datos.

La búsqueda de haplotipos comunes entre ambas muestras de población se realizó mediante la opción *Search for shared haplotypes between populations* contenida en el submenú *Haplotype frequencies* del apartado *Diversity indices* del programa Arlequin Ver. 2.000 (SCHNEIDER *et al.* 2000).

## 16.8. Haplogrupos mitocondriales

### 16.8.1. Asignación de haplogrupos a las muestras

Las secuencias obtenidas de las muestras antiguas se clasificaron en haplogrupos en base a la clasificación de RICHARDS *et al.* 2000 (**Tabla II** de “Introducción”). Esta clasificación resulta precisa siempre y cuando existan sustituciones características de haplogrupo en el segmento analizado. En algunos casos sin embargo se hace necesaria la información de regiones codificantes del DNA mitocondrial y de otras posiciones de la región Hipervariable I y II no analizadas aquí para una correcta clasificación de las secuencias en haplogrupos. Aquellos individuos que no presentaban posiciones diagnóstico en el segmento analizado de la Región Hipervariable I se clasificaron de acuerdo a la asignación de haplogrupos a secuencias hecha por RICHARDS *et al.* 2000 (haplogrupos europeos) y WATSON *et al.* 1997 (haplogrupos africanos) y a la información obtenida de los árboles filogenéticos de Neighbor-Joining y Median Joining Network (apartados 16.5.1. y 16.7.1.2. del presente capítulo).

En aquellos casos en los que la región analizada resultaba insuficiente para una asignación única se señalaron todos los haplotipos posibles. Las muestras que no pudieron ser clasificadas mediante ninguno de los criterios anteriores fueron designadas con un interrogante.

### 16.8.2. Frecuencias de haplogrupos

Los diferentes haplotipos mitocondriales definidos respecto a la secuencia consenso o CRS (ANDERSON *et al.* 1981) presentes en la muestra de poblaciones actuales y en los subconjuntos de población antigua fueron establecidos mediante un programa en FORTRAN diseñado especialmente para el análisis de secuencias de mtDNA por el Catedrático de Estadística de la Facultad de Medicina de la Universidad Complutense de Madrid, Dr. Miguel Sánchez. Una vez definidos, los haplotipos fueron clasificados manualmente en haplogrupos de acuerdo con RICHARDS *et al.* 2000. A aquellos no clasificados de esta forma se le asignó el mismo haplogrupo con el que aparecían en la publicación original. Los linajes no clasificados bajo ninguno de estos criterios fueron designados “sin clasificar”. Las frecuencias de los diferentes

haplogrupos dentro de cada población se calcularon con el programa EXCEL del paquete Microsoft Office 2000. Las diferentes distribuciones se representaron en diagramas circulares dentro de dos mapas, uno de Europa y Próximo Oriente y otro de África. Se calculó también la frecuencia de los diferentes haplogrupos en el conjunto de poblaciones europeas, de Próximo Oriente y de África.