

Fig. R-E1 Espectro de emisión de fluorescencia ($\lambda_{exc}= 310\text{nm}$) de 3 residuos de extractos de hueso después del “almacenamiento frío” y del sobrenadante obtenido después de precipitar el DNA con espermidina según el protocolo de Persson (1992). En este espectro y todos los siguientes, los picos armónicos, que son reflejo de la luz de excitación, han sido eliminados del gráfico para observar los picos de emisión con mayor claridad. Todas las muestras provienen de individuos de la Plaça Vella. El patrón es consistente a pesar de que las muestras han sido obtenidas con 3 protocolos distintos (ver recuadro).

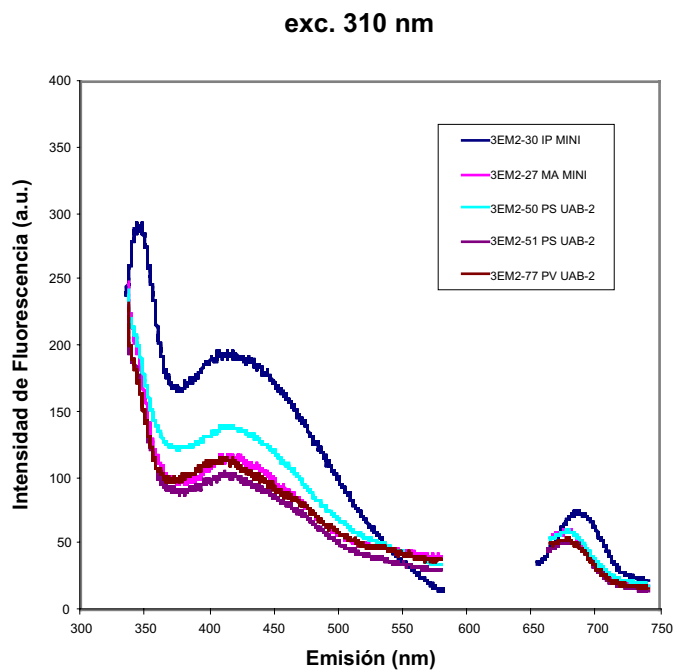


Figura R-E2 Espectro de emisión de fluorescencia ($\lambda_{exc}= 310\text{nm}$) de los residuos de extractos de dientes después del “almacenamiento frío”. Las muestras proceden de 4 poblaciones distintas (ver recuadro). El resultado es consistente sin importar la procedencia de la muestra. En algunas muestras, el primer pico resultó desplazado y casi absorbido por el pico armónico

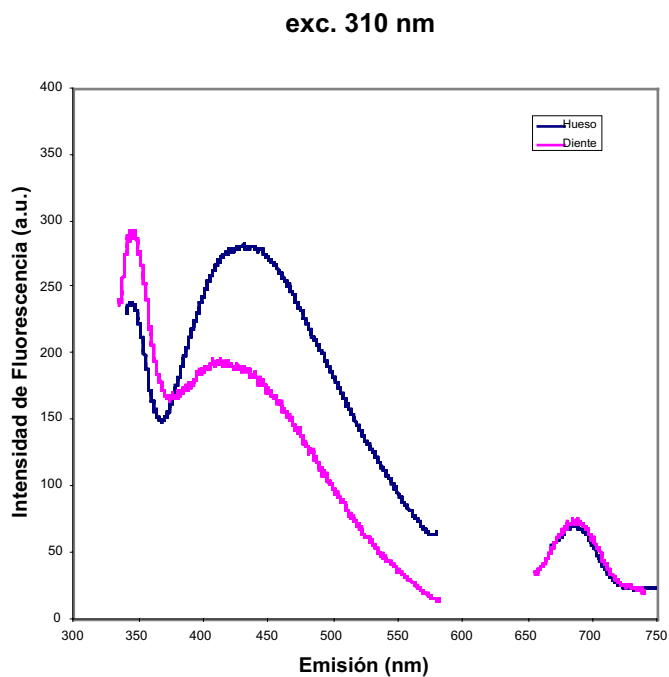


Figura R-E3 Comparación de los espectros de emisión de fluorescencia ($\lambda_{exc}= 310\text{nm}$) de residuos de extractos de hueso y diente. La relación de la intensidad entre el primer y segundo pico está invertida en las muestras de diente respecto a las muestras de hueso.

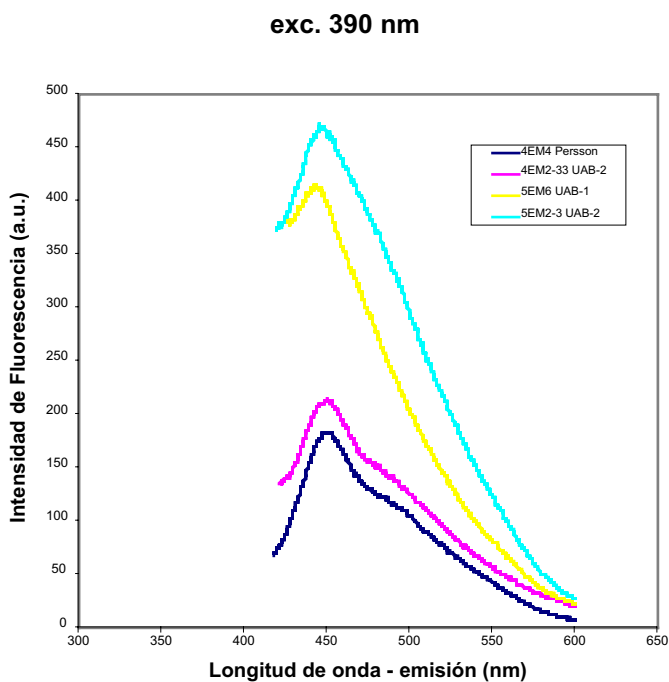


Figura R-E4 Espectro de emisión de fluorescencia ($\lambda_{exc}= 390\text{nm}$) de los residuos de extractos de hueso después del “almacenamiento frío”. El abultamiento que se observa alrededor de los 500nm podría representar un pico oculto.

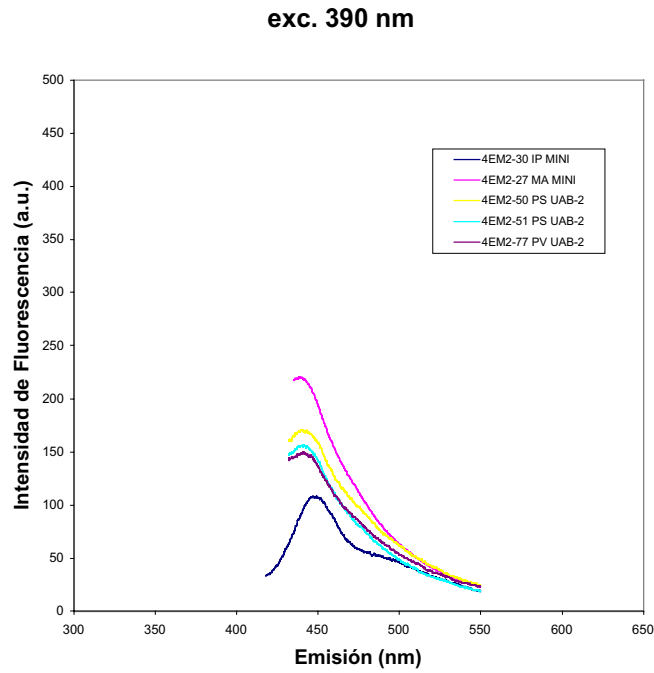


Figura R-E5 Espectro de emisión de fluorescencia ($\lambda_{exc}= 390nm$) de los residuos de extractos de dientes después del “almacenamiento frío”.

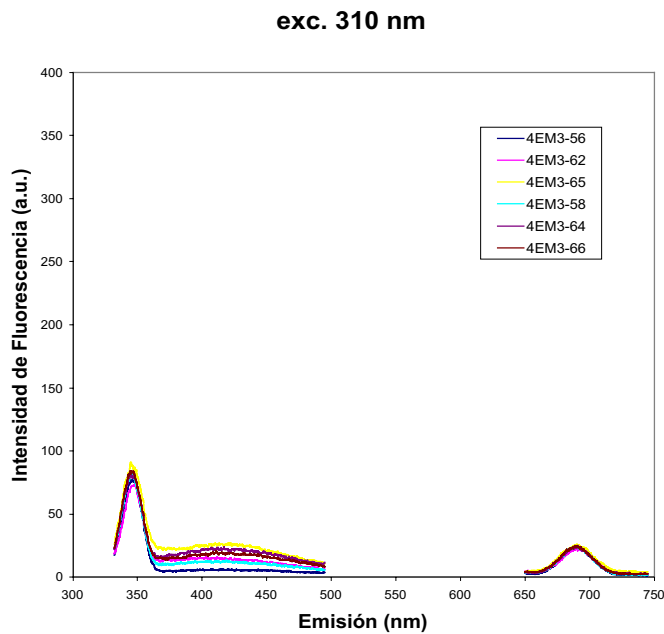


Figura R-E6 Espectro de emisión de fluorescencia ($\lambda_{exc}= 310nm$) de los residuos de extractos de dientes antes de someter los tubos al “almacenamiento frío”.

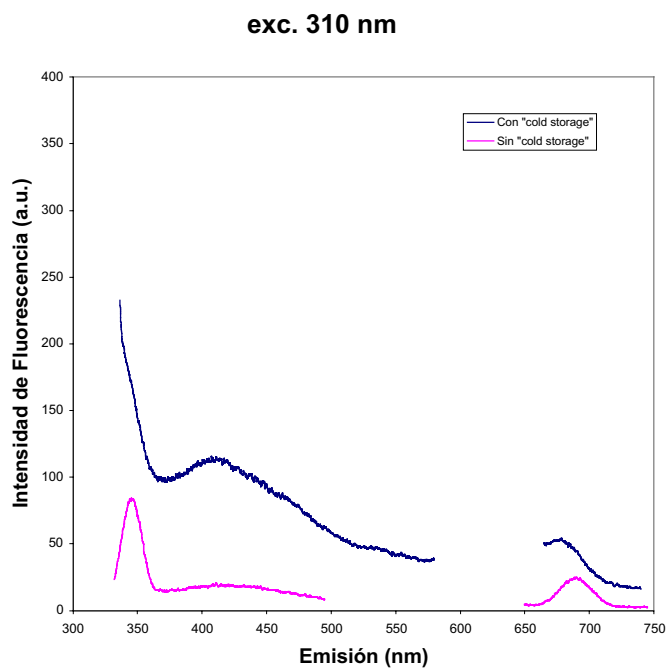


Figura R-E7 Comparación de los espectros de emisión de fluorescencia ($\lambda_{exc}=310\text{nm}$) de una muestra obtenida antes del “almacenamiento frío” y otra obtenida después de éste. La intensidad de fluorescencia es marcadamente menor cuando no se aplica el almacenamiento en nevera

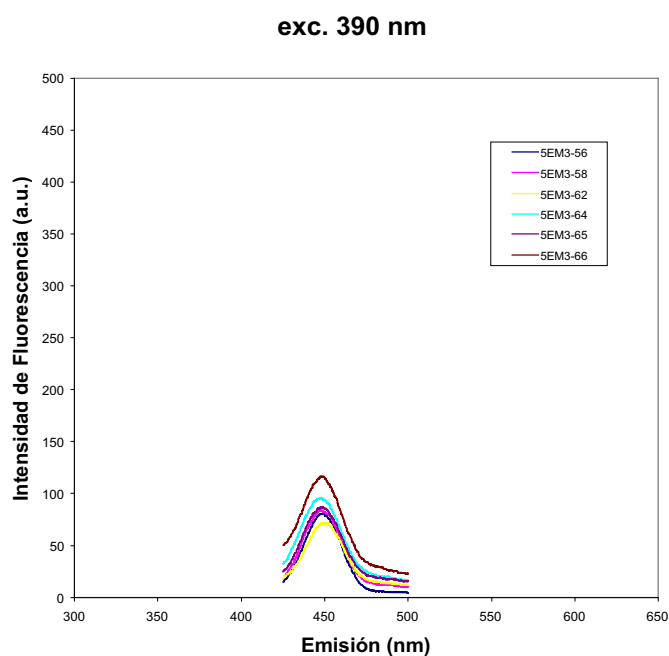


Figura R-E8 Espectro de emisión de fluorescencia ($\lambda_{exc}=390\text{nm}$) de los residuos de extractos de dientes antes de someter los tubos al “almacenamiento frío”.

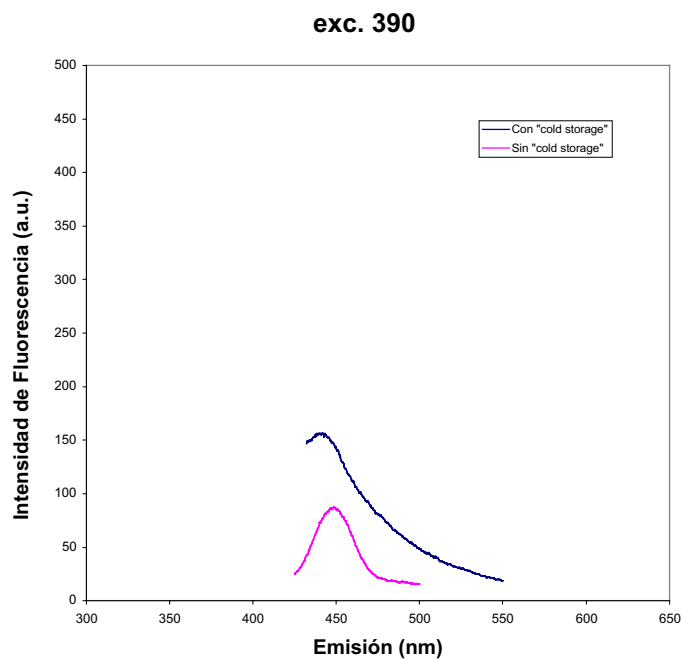


Figura R-E9 Comparación de los espectros de emisión de fluorescencia ($\lambda_{exc}=390\text{nm}$) de una muestra obtenida antes del “almacenamiento frío” y otra obtenida después de éste. La diferencia en la intensidad de fluorescencia no es tan notoria, sin embargo, la media de la intensidad de fluorescencia de las muestras antes del “almacenamiento frío” es significativamente menor de la que presentan los residuos después de este paso (ver texto).

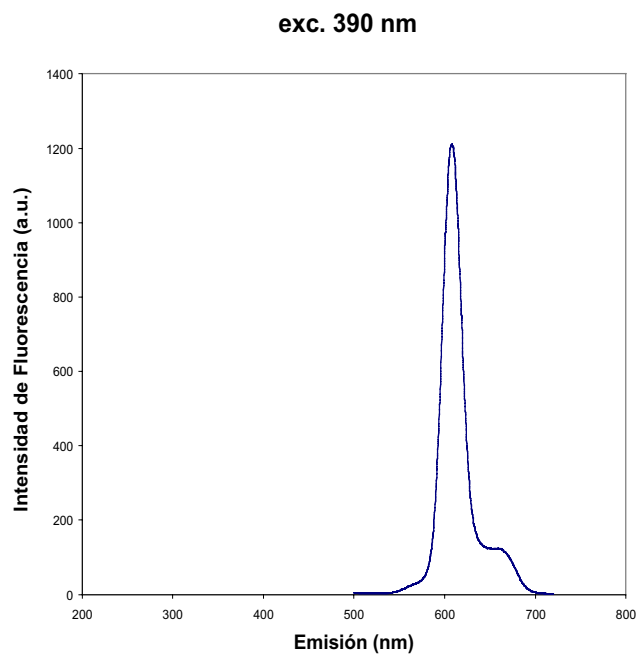


Figura R-E10 Espectro de emisión de fluorescencia ($\lambda_{exc}=390\text{nm}$) de la solución en medio básico de la mesoporfirina IX dihidroclorada.

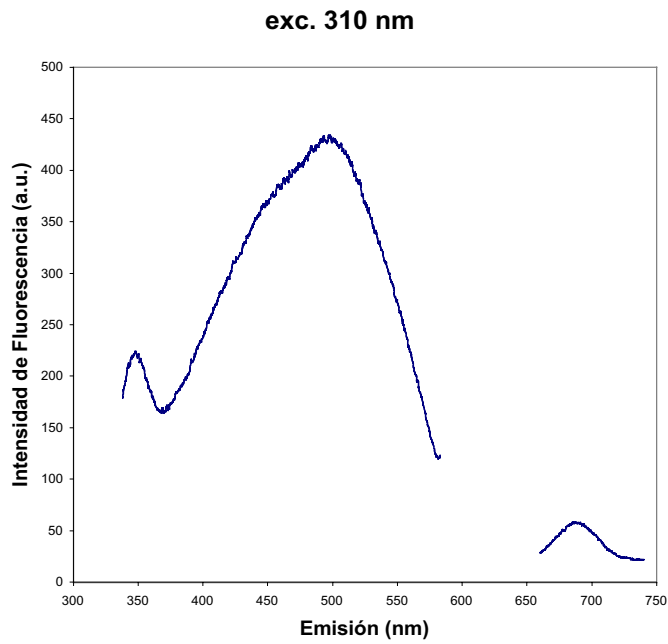


Figura R-E11 Espectro de emisión de fluorescencia ($\lambda_{exc}= 310\text{nm}$) de los ácidos húmicos. Los ácidos fúlvicos mostraron un espectro prácticamente idéntico. Obsérvese la espalda alrededor de los 445nm.

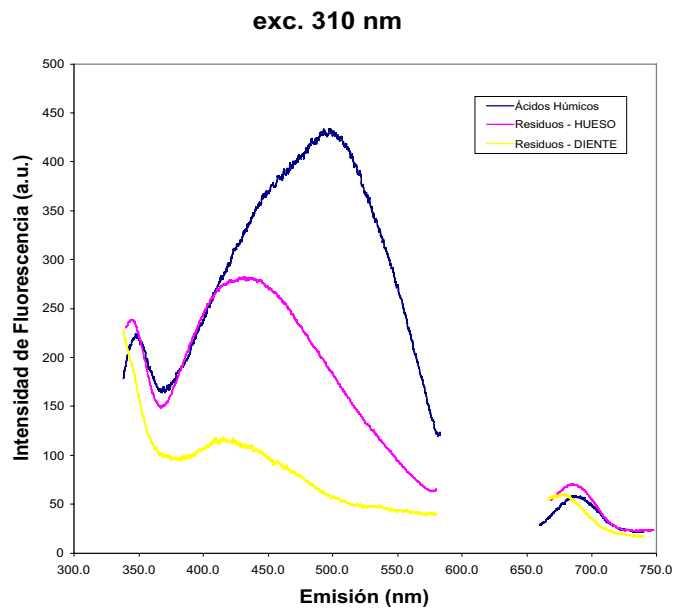


Figura R-E12 Comparación de los espectros de emisión de fluorescencia ($\lambda_{exc}= 310\text{nm}$) de los ácidos húmicos, los residuos de hueso y los residuos de diente. Es notable que el segundo pico de los residuos coincide con la espalda que se observa en el segundo pico de los ácidos húmicos.

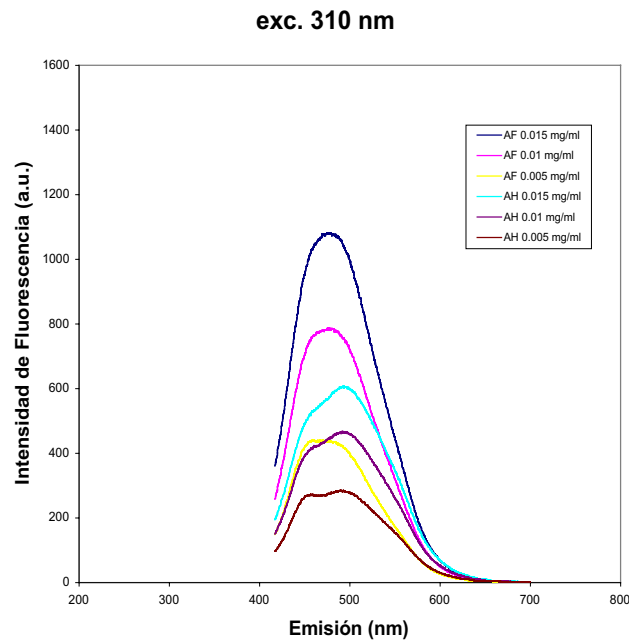


Figura R-E13 Espectro de emisión de fluorescencia ($\lambda_{exc}=310\text{nm}$) de distintas concentraciones de ácidos húmicos y fúlvicos. A bajas concentraciones se observa la presencia de dos picos solapados.

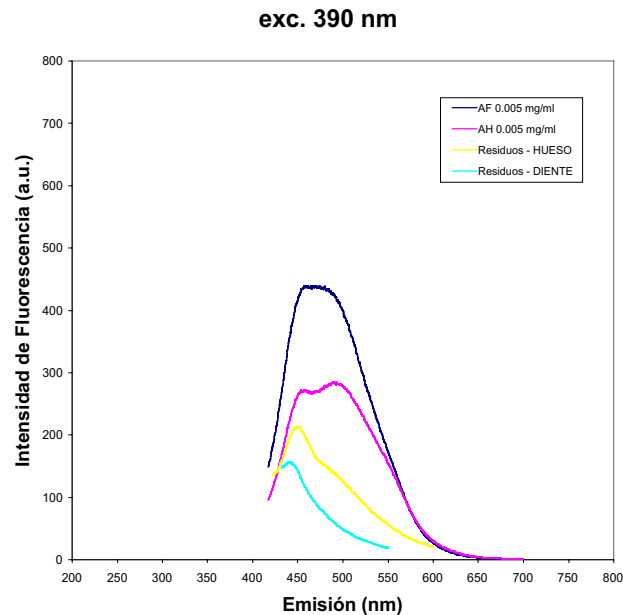


Figura R-E14 Comparación de los espectros de emisión de fluorescencia ($\lambda_{exc}=390\text{nm}$) de las soluciones de baja concentración (0.005 mg/ml) de ácidos húmicos y fúlvicos y de los residuos de hueso y diente. El pico de los residuos podría corresponder al pico de menor longitud de onda de los dos que presenta la solución de ácidos húmicos.

DNA ACTUAL

En total, fueron analizados 90 individuos de la población actual. De 50 de ellos se obtuvo el DNA a partir de muestras de sangre, en tanto que de los 40 restantes el DNA fue obtenido a partir de muestras de pelo.

Haplogrupos

En los 90 individuos analizados se encontraron 40 que pertenecen al haplogrupo H, 12 al J, 5 al K, 12 al T, 10 al U, 3 al V, 2 al X, 1 al M y 5 a “otro”. No se encontraron individuos del haplogrupo I ni del W (Tabla R-30). El haplogrupo H fue el más representado, tal y como se espera para poblaciones europeas. Sorprendentemente, la frecuencia de los haplogrupos V y K resultó mucho más baja que la encontrada en la población antigua, sin embargo las diferencias no son significativas (prueba exacta de Fisher bilateral $P=0.0915$ y $P=0.1066$ respectivamente).

Haplogrupo	N	%
H	40	44.4
I	0	0.0
J	12	13.3
K	5	5.6
T	12	13.3
U	10	11.1
V	3	3.3
W	0	0.0
X	2	2.2
M	1	1.1
Otro	5	5.6
Total	90	

Tabla R-30 Frecuencia de los haplogrupos en la población catalana actual

Llama la atención la presencia de un individuo perteneciente al haplogrupo M, ya que este haplogrupo es característico del continente asiático. Sin embargo en otras poblaciones europeas se ha encontrado este haplogrupo, aunque en frecuencias muy bajas. Torroni y col. (1996) encontraron en población Finlandesa un individuo del haplogrupo M en los 49 analizados. Más aún, Richards y col. (1998) han detectado un 1% de linajes asiáticos y africanos en el acervo mitocondrial europeo que podrían deberse a migraciones recientes. Por estos motivos, la información de estos individuos fue incluida dentro del grupo “otros” para el cálculo de frecuencias en los análisis estadísticos y filogenéticos.

Secuencias de la región de control

Se amplificó el fragmento de 446pb de la región de control de 26 individuos, sin embargo, el resultado de la secuenciación comprende únicamente 391pb (de la posición 16001 a la 16391). Se encontraron 23 secuencias diferentes definidas por 47 sitios variables, compuestos a su vez por 42 transiciones, 4 transversiones y 1 inserción. De las transiciones, 21 son del tipo C->T, 16 son T->C, 4 G->A y 1 A->G, en tanto que dos de las transversiones son cambios C->G, uno G->C y uno T->A. De las secuencias encontradas, 1 es idéntica en dos individuos, 1 la presentan 3 individuos y cada una de las 21 restantes la presenta un único individuo (Tabla R-31).

Correspondencia secuencia – haplogrupo

En la Tabla R-31 puede observarse el haplogrupo al que pertenece cada una de las secuencias obtenidas. En general se ha verificado la correspondencia encontrada por otros autores. En algunos casos, la información de la secuencia ha servido para identificar el subgrupo al que pertenecen, siguiendo el criterio de Francalacci y col. (1999). Tal es el caso de las muestras incluidas en los subgrupos J2a, T1a1, T2, T2a, U4 y U5.

El individuo C26, que fue caracterizado dentro del grupo “otro” presenta dos de las substituciones del haplogrupo I, sin embargo carece de una tercera, la 16391A. Cabe preguntarse si este individuo representa un linaje derivado del haplogrupo I que haya perdido la substitución 16391A, a la vez que ha adquirido el sitio de restricción 1715 *DdeI*.

Por otra parte, el individuo C6 clasificado también dentro del grupo “otro” no presenta un patrón correlacionado con ningún haplogrupo y ninguna de estas dos

secuencias coincide con la del individuo definido también como “otro” por Torroni y col. (1996), que tampoco presenta un patrón de substituciones específico de ningún haplogrupo. Esto descarta la posibilidad de que los tipos mitocondriales no definidos dentro de algún haplogrupo puedan estar relacionados entre ellos filogenéticamente, por lo que la hipótesis de que sean linajes derivados de haplogrupos existentes es más plausible.

La secuencia del haplogrupo M es la única que presenta una inserción (16085b). Como es de esperar, este individuo presenta una substitución relacionada con el haplogrupo M (16223T), aunque dicha substitución no es exclusiva, ya que además de los subgrupos derivados del M (A, C, D y G), la presentan los linajes africanos del grupo L y los haplogrupos europeos X, W e I.

En los individuos del haplogrupo H, no se observa ningún patrón de substituciones específico, resultado que concuerda con lo descrito por otros autores.

En los demás casos, además de las substituciones particulares de cada haplogrupo (enmarcadas en la Tabla R-31), se puede observar una variedad de substituciones que definen la historia evolutiva de cada linaje. En este sentido, sería interesante investigar si las substituciones que presentan dos de los individuos del grupo T2 (16218T y 16324C) pueden definir un subgrupo de menor jerarquía y por tanto de historia más reciente.

ANÁLISIS ESTADÍSTICO Y RECONSTRUCCIÓN FILOGENÉTICA

Haplogrupos

En la Tabla R-32 se muestra la distribución de los haplogrupos que presentan las poblaciones incluidas en el análisis. Además de las poblaciones antigua (PV) y actual (CAT) de Cataluña, se incluyeron las 8 poblaciones descritas en la sección de Material y Métodos. Como puede verse, las poblaciones de Galicia (GALI) y Barbagia (BAR) destacan por la alta frecuencia que presentan del haplogrupo H. Las poblaciones de Galura (GALU) y del País Vasco (VAS) presentan una frecuencia elevada del haplogrupo U. Es interesante destacar que las poblaciones de Galicia y del País Vasco presentan características diferentes entre sí, lo que descarta la posibilidad de que exista un sesgo importante en la distribución de los haplogrupos debido a la forma en que han sido inferidos.

Haplogrupo	GALI ^a (n=92)	VAS ^a (n=45)	PV (n=24)	CAT (n=90)	BAR (n=45)	GALU (n=51)	COR (n=56)	TOS (n=48)	FIN (n=49)	SUE (n=37)
H	63.0	48.8	33.3	44.4	64.4	41.2	42.9	41.7	40.8	40.5
I	1.0	0.0	4.2	0.0	0.0	3.9	0.0	4.2	2.0	0.0
J	8.7	4.4	12.5	13.3	2.2	9.8	10.7	14.6	14.3	2.7
K	4.3	2.2	16.7	5.6	2.2	5.9	12.5	6.3	4.1	13.5
T	2.1	4.44	8.3	13.3	8.9	9.8	8.9	10.4	6.1	21.6
U	11.9	24.44	8.3	11.1	4.5	27.5	10.7	10.4	16.3	16.2
V	3.2	11.1	12.5	3.3	8.9	0.0	0.0	0.0	4.1	5.4
W	2.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	4.1	0.0
X	1.1	4.4	0.0	2.2	0.0	0.0	8.9	8.3	4.1	0.0
Otro	2.1	0.0	4.2	6.7	8.9	2.0	5.4	2.1	4.1	0.0

Tabla R-32 Distribución de los haplogrupos (%) en distintas muestras de poblaciones europeas. GALI= Galicia, VAS= Vascos, PV= Plaça Vella, CAT= Catalana actual, BAR= Barbagia, GAL= Galura, COR= Córcega, TOS= Toscana, FIN= Finlandia, SUE= Suecia. La descripción de las poblaciones y las referencias respectivas pueden verse en la sección de Material y Métodos. ^a Los haplogrupos fueron inferidos a partir de la secuencia.

Índice de diversidad

La diversidad genética de las distintas poblaciones analizadas se presenta en la Tabla R-33. La población de la Plaça Vella mostró la mayor diversidad genética. Este

resultado puede deberse por un lado al bajo número de individuos analizados y por otro a la baja frecuencia del haplogrupo H, propiciada por la eliminación de dos individuos pertenecientes a este haplogrupo que mostraron resultados contradictorios durante la caracterización de los segundos extractos analizados, como ya se ha explicado en secciones anteriores.

Población	N	Diversidad genética	d.e.
BARBAGIA	45	0.5707	± 0.0820
GALICIA	92	0.5824	± 0.0567
PAÍS VASCO	45	0.6980	± 0.0550
GALURA	51	0.7451	± 0.0411
CATALUÑA	90	0.7538	± 0.0367
SUECIA	37	0.7613	± 0.0449
CÓRCEGA	56	0.7727	± 0.0445
TOSCANA	48	0.7863	± 0.0475
FINLANDIA	49	0.7900	± 0.0461
PLAÇA VELLA	24	0.8478	± 0.0492

Tabla R-33 Diversidad del mtDNA en distintas poblaciones europeas según la distribución de los haplogrupos que presentan, estimada con la fórmula de Nei (1987); d.e., desviación estándar.

Las poblaciones con menor diversidad genética resultaron ser las poblaciones de Barbagia (BAR) y Galicia (GALI), debido probablemente a la alta frecuencia que presentan del haplogrupo H. La población de Barbagia es una población aislada cultural y genéticamente, en la que los niveles de endogamia pueden ser elevados, lo que explicaría la baja diversidad genética. De igual forma, es posible que la baja diversidad encontrada en la población de Galicia esté relacionada con un aislamiento cultural, que propicie niveles elevados de endogamia.

Asimismo, se debe tener en cuenta que dentro de cada haplogrupo existen tipos mitocondriales distintos, por lo que en algunas poblaciones la diversidad genética puede ser en realidad mayor. Un panorama más claro se obtendrá al analizar la diversidad genética estimada a partir de las secuencias de la región de control.

El índice de diversidad en el resto de poblaciones cae en un rango que va de 0.6980 de la población vasca, hasta 0.7900 de la población finlandesa.

Prueba exacta de diferenciación poblacional

La prueba exacta de diferenciación poblacional (Raymond y Rousset, 1995) permite valorar el nivel de asociación entre la frecuencia de los haplogrupos y la población de procedencia de las muestras. Puede ser aplicada tanto a grupos de poblaciones como a parejas de poblaciones.

Según esta prueba, existe asociación significativa cuando se consideran las diez poblaciones estudiadas ($P= 0.00019$), lo que puede interpretarse como que existe diferencia significativa entre las poblaciones en cuanto a la distribución de los haplogrupos mitocondriales que presentan. Con el fin de valorar con más detalle esta diferencia, se analizaron diversos subgrupos de poblaciones y los resultados se resumen en la Tabla R-34.

Grupo	Poblaciones	valor de P	Error estándar	Dif.
1	Todas (10)	0.000190	± 0.00013	+
2	Con Haplogrupo (8)	0.005355	± 0.00152	+
3	Sin GALI, BAR (8)	0.027615	± 0.00493	+
4	Mediterráneas (6)	0.004495	± 0.00144	+
5	Ibéricas (4)	0.003630	± 0.00101	+
6	Ibéricas (3)	0.025965	± 0.00310	+
7	Italia-COR (4)	0.001015	± 0.00040	+
8	Italia-COR (3)	0.241569	± 0.00827	-
9	Islas (3)	0.000760	± 0.00032	+

Tabla R-34 Prueba exacta de diferenciación poblacional dentro de distintos grupos de poblaciones. 1= las 10 poblaciones de la Tabla R-32; 2= excluyendo GALI y VAS; 3= excluyendo GALI y BAR; 4= PV, CAT, BAR, GALU, COR, TOS; 5= GALI, VAS, PV, CAT; 6= VAS, PV, CAT; 7= BAR, GALU, COR, TOS; 8= GALU, COR, TOS; 9= BAR, GALU, COR; Dif.= indica si existe diferencia significativa ($P < 0.05$) entre las poblaciones del grupo.

En todos los grupos analizados se observó diferencia significativa, con la excepción del grupo 8. El grupo 8 consiste de tres poblaciones (GALU, TOS, COR) y es interesante observar que cuando se añade la población de Barbagia (grupo 7) la diferencia se torna significativa. Dado que las poblaciones de Galicia y Barbagia son las que presentan una diversidad considerablemente menor que el resto de poblaciones, podría pensarse que son las responsables de la diferencia en todos los grupos analizados, no obstante, cuando se excluyen estas dos poblaciones y se analiza el grupo de las 8 poblaciones restantes (grupo 3), también se observa diferencia significativa ($P=0.027615$), que podría ser explicada por la diferencia entre las poblaciones ibéricas y las poblaciones italianas, o bien por la presencia de la población del País Vasco, que también presenta un índice de diversidad relativamente bajo.

Para esclarecer más esta situación, se realizó la prueba de diferenciación entre todas las parejas posibles de las 10 poblaciones analizadas (Tabla R-35). Las poblaciones que presentaron más diferencia fueron la de Barbagia que se diferencia de 7 poblaciones, la de Galicia, que es significativamente diferente de 6 poblaciones y la del País Vasco con 5 poblaciones diferentes.

La población de Barbagia no presenta diferencias significativas con la población de Galicia ni con la población actual de Cataluña (CAT); la población de Galicia no se diferencia de las poblaciones de Finlandia y del País Vasco ni, obviamente, de la de Barbagia; en tanto que la población del País Vasco (VAS) no es diferente de las poblaciones de Finlandia, Galicia, Galura y Suecia.

En el otro extremo, la población de Finlandia resultó ser la menos diferenciada, ya que únicamente presenta diferencias significativas con la población de Barbagia. En la Figura R-5 se observa un diagrama que refleja el número de poblaciones de las que se diferencia cada una de las poblaciones analizadas.

Por lo que respecta a la población de la Plaça Vella (PV), ésta únicamente se diferencia de las poblaciones de Galicia, Barbagia y del País Vasco, en tanto que la población actual de Cataluña (CAT) únicamente presenta diferencias significativas con las poblaciones de Galicia y del País Vasco. No existe diferencia significativa entre las poblaciones antigua y actual de Cataluña en cuanto a la distribución de haplogrupos que presentan.

En general, estos resultados concuerdan con observaciones previas que indican que las poblaciones Europeas son bastante homogéneas en cuanto a la variabilidad del mtDNA. Asimismo, indican que la distribución de los haplogrupos en una población es bastante sensible a la deriva genética en un entorno de aislamiento, ya sea cultural o geográfico, como lo demuestran las poblaciones de Barbagia y Galicia y en menor medida la del País Vasco.

Prueba de Diferenciación

	BAR	CAT	COR	FIN	GALI	GALU	P.V.	SUE	TOS	VAS
BAR	0.0000	0.0050	0.0000	0.0020	0.0040	0.0000	0.0020	0.0010	0.0000	0.0013
CAT	0.1104	0.0000	0.0080	0.0080	0.0020	0.0030	0.0060	0.0070	0.0080	0.0014
COR	0.0068	0.3864	0.0000	0.0060	0.0010	0.0030	0.0050	0.0030	0.0060	0.0009
FIN	0.0408	0.4903	0.3984	0.0000	0.0050	0.0050	0.0060	0.0030	0.0030	0.0066
GALI	0.0987	0.0342	0.0175	0.3641	0.0000	0.0040	0.0020	0.0010	0.0020	0.0067
GALU	0.0013	0.1139	0.0641	0.4943	0.0428	0.0000	0.0030	0.0040	0.0070	0.0036
PV	0.0367	0.2243	0.1682	0.5581	0.0314	0.0699	0.0000	0.0060	0.0060	0.0034
SUE	0.0176	0.1674	0.0716	0.0859	0.0039	0.1342	0.2705	0.0000	0.0020	0.0023
TOS	0.0038	0.3056	0.7874	0.8910	0.0492	0.2484	0.2882	0.0681	0.0000	0.0022
VAS	0.0192	0.0258	0.0097	0.3883	0.1556	0.0755	0.0466	0.0508	0.0290	0.0000

Tabla de Diferencias Significativas

	BAR	CAT	COR	FIN	GALI	GALU	PV	SUE	TOS
CAT	-								
COR	+	-							
FIN	+	-	-						
GALI	-	+	+	-					
GALU	+	-	-	-	+				
PV	+	-	-	-	+	-			
SUE	+	-	-	-	+	-	-		
TOS	+	-	-	-	+	-	-	-	
VAS	+	+	+	-	-	-	+	-	+

Tabla R-35 Prueba exacta de diferenciación poblacional entre todas las parejas posibles de las poblaciones analizadas, basada en la frecuencia de los haplogrupos. En la tabla superior se muestra el valor de P (debajo de la diagonal) y su error estándar (encima de la diagonal). En la tabla inferior el signo + indica que entre las poblaciones analizadas existe diferencia significativa ($P < 0.05$). Las poblaciones más diferenciadas resultaron ser las de Barbagia (BAR), Galicia (GALI) y País Vasco (VAS).

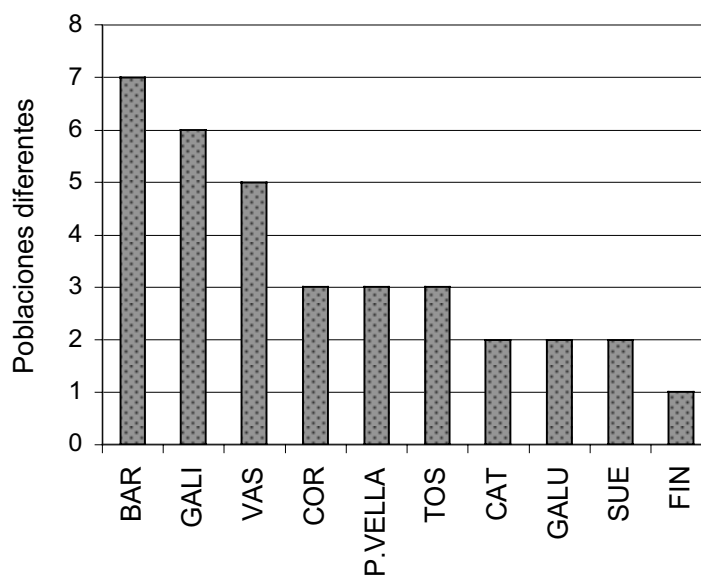


Figura R-5 Histograma del número de poblaciones de las que difiere significativamente ($P < 0.05$) cada una de las poblaciones analizadas, según la prueba exacta de diferenciación poblacional (Tabla R-35).

Distancias genéticas

Para reconstruir las relaciones filogenéticas de las poblaciones analizadas en base a su distribución de haplogrupos, se estimaron las distancias de Reynolds y col. (1983) y la de la medida de la cuerda (Cavalli-Sforza y Edwards, 1967). De acuerdo con la distancia de Reynolds (Tabla R-36) la población más distanciada del resto es la de Barbagia, con un promedio de 0.06404, seguida de la población de Galicia con 0.04833; en tanto que el promedio de distancias más bajo (0.02356) lo presentó la población actual de Cataluña (Fig. R-6a).

	PV	CAT	TOS	COR	GALU	VAS	FIN	SUE	GALI
CAT	0.0246								
TOS	0.0262	0.0072							
COR	0.0231	0.0089	0.0059						
GALU	0.0442	0.0236	0.0252	0.0284					
VAS	0.0536	0.0323	0.0376	0.0375	0.0209				
FIN	0.0243	0.0083	0.0068	0.0122	0.0143	0.0216			
SUE	0.0301	0.0224	0.0307	0.0260	0.0280	0.0414	0.0332		
GALI	0.0798	0.0380	0.0452	0.0439	0.0580	0.0358	0.0410	0.0731	
BAR	0.0897	0.0468	0.0662	0.0609	0.0907	0.0585	0.0659	0.0775	0.0202

Tabla R-36 Matriz de distancias genéticas de Reynolds y col. (1983) obtenida con el programa GENDIST del paquete informático Phylip.

En general, el patrón de diferenciación es algo distinto al que se deduce de la prueba exacta de diferenciación poblacional, pues las poblaciones de Plaça Vella, Suecia y Galura muestran una mayor distancia genética que la que se deduce de la prueba exacta. La población del País Vasco (VAS) en cambio, presenta una menor distancia genética con las demás poblaciones, comparada con la diferenciación que presenta mediante la prueba exacta. En la Tabla R-37 se presentan las distancias de las poblaciones antigua (PV) y actual (CAT) de Cataluña con el resto de poblaciones analizadas, ordenadas de menor a mayor.

Es interesante observar que tanto la población antigua como la actual de Cataluña presentan las menores distancias genéticas con las poblaciones Corza, Toscana, Finlandesa y Sueca, no obstante, la población actual de Cataluña presenta menor distancia con otras poblaciones antes que con la población de la Plaça Vella, lo que no ocurre a la inversa, pues

	P.VELLA		CAT
COR	0.0231	TOS	0.0072
FIN	0.0243	FIN	0.0083
CAT	0.0246	COR	0.0089
TOS	0.0262	SUE	0.0224
SUE	0.0301	GALU	0.0236
GALU	0.0442	P.V.	0.0246
VAS	0.0536	VAS	0.0323
GALI	0.0798	GALI	0.0380
BAR	0.0897	BAR	0.0468
\bar{x}	0.0439		0.0235

Tabla R-37 Distancia genética de Reynolds entre las poblaciones de la Plaça Vella y actual de Cataluña con el resto de poblaciones analizadas; \bar{x} = promedio de las distancias.

la población antigua presenta una distancia menor tan sólo con dos poblaciones antes que con la población actual. Estas pequeñas diferencias entre las poblaciones antigua y actual de Cataluña pueden deberse a las diferencias en el tamaño muestral.

Con relación a la distancia de la medida de la cuerda (Tabla R-38), resulta destacable que la población de Galicia no presente mayores distancias con el resto de poblaciones y que la población de Suecia presente un promedio de diferencias similar al de la población de Barbagia, lo que está en discrepancia tanto con lo observado en la prueba exacta de diferenciación poblacional, como con los resultados de la estimación de distancias genéticas según el método de Reynolds y col. (1983).

	PV	CAT	TOS	COR	GALU	VAS	FIN	SUE	GALI
CAT	0.0315								
TOS	0.0590	0.0292							
COR	0.0596	0.0170	0.0188						
GALU	0.0486	0.0336	0.0332	0.0419					
VAS	0.0621	0.0391	0.0633	0.0591	0.0552				
FIN	0.0379	0.0197	0.0163	0.0326	0.0342	0.0363			
SUE	0.0407	0.0366	0.0686	0.0572	0.0419	0.0404	0.0587		
GALI	0.0421	0.0259	0.0309	0.0389	0.0351	0.0300	0.0137	0.0531	
BAR	0.0488	0.0288	0.0802	0.0649	0.0696	0.0528	0.0539	0.0519	0.0325

Tabla R-38 Matriz de distancias genéticas de la medida de la cuerda (Cavalli-Sforza y Edwards, 1967) obtenida con el programa GENDIST del paquete informático Phylip.

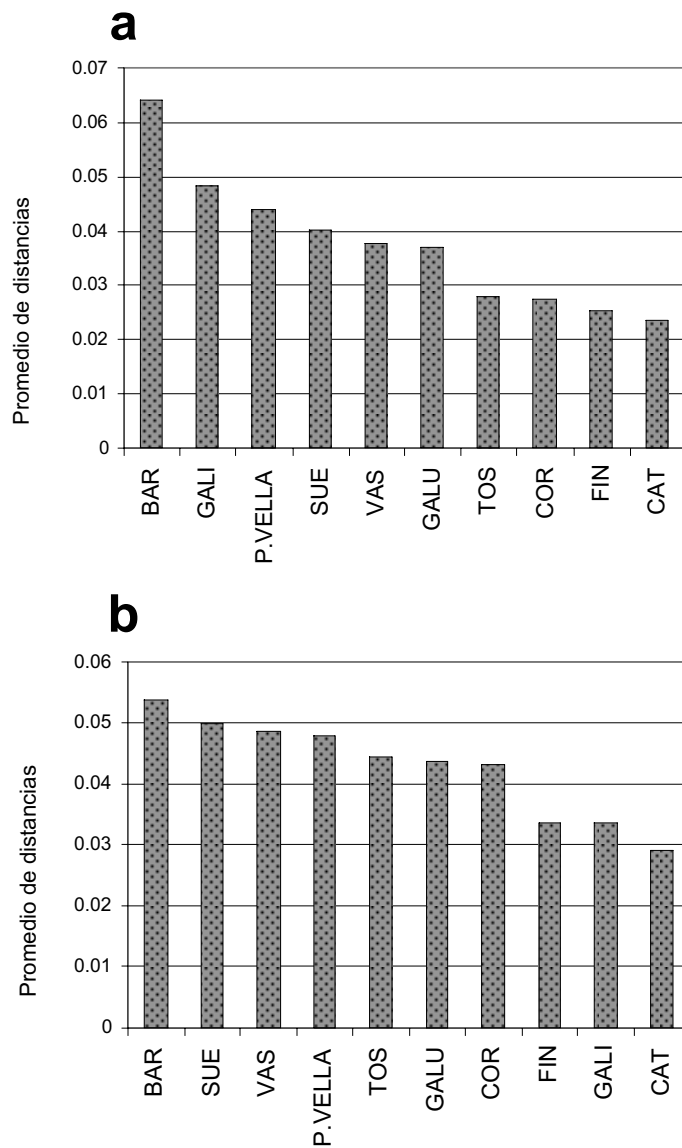


Figura R-6 Representación gráfica de los promedios de las distancias genéticas que presenta cada población en relación a las restantes. **(a)** distancia de Reynolds y col. (1983), **(b)** distancia de la medida de la cuerda (Cavalli-Sforza y Edwards, 1967).

Asimismo, también es notable que cuando se utiliza la distancia genética de la medida de la cuerda, las poblaciones resultan más homogéneas (Figura R-6b). Por otra parte, el patrón de distancias que presentan las poblaciones de la Plaça Vella y Catalana actual, también es ligeramente diferente, según se muestra en la Tabla R-39.

	P.VELLA		CAT
CAT	0.0315	COR	0.0170
FIN	0.0379	FIN	0.0197
SUE	0.0407	GALI	0.0259
GALI	0.0421	BAR	0.0288
GALU	0.0486	TOS	0.0292
BAR	0.0488	P.V.	0.0315
TOS	0.0590	GALU	0.0336
COR	0.0596	SUE	0.0366
VAS	0.0621	VAS	0.0391
\bar{x}	0.0478		0.0290

Tabla R-39 Distancia genética de la medida de la cuerda entre las poblaciones de la Plaça Vella y actual de Cataluña con el resto de poblaciones analizadas; \bar{x} = promedio de las distancias.

La población de la Plaça Vella presenta la menor distancia genética con la población Catalana actual y la máxima distancia con la población Vasca. Por su parte, la población actual, presenta la menor distancia con la población de Córcega y la mayor también con la Vasca después de la Sueca. Es de esperar que la población de la Plaça Vella presente la menor distancia con la población actual de Cataluña, pues en principio se trata de la misma población. Sin embargo es notable que al igual que con la distancia de Reynolds, la población actual presente menores distancias con otras poblaciones actuales antes que con la población antigua. Este efecto podría deberse al bajo número de muestras de la población antigua, como ya hemos mencionado.

Para valorar en conjunto las distancias de las distintas poblaciones es necesario construir un árbol filogenético ya que de esta forma se consideran las distancias de todas las poblaciones y su interacción. Este análisis se expone en la siguiente sección.

Árboles filogenéticos

El árbol filogenético sin raíz construido con el método de Fitch-Margoliash a partir de la matriz de distancias de Reynolds (Fig. R-7a) muestra características compatibles con los resultados de la prueba exacta de diferenciación poblacional, ya que

las poblaciones de Galicia y Barbagia aparecen considerablemente más alejadas del resto de poblaciones.

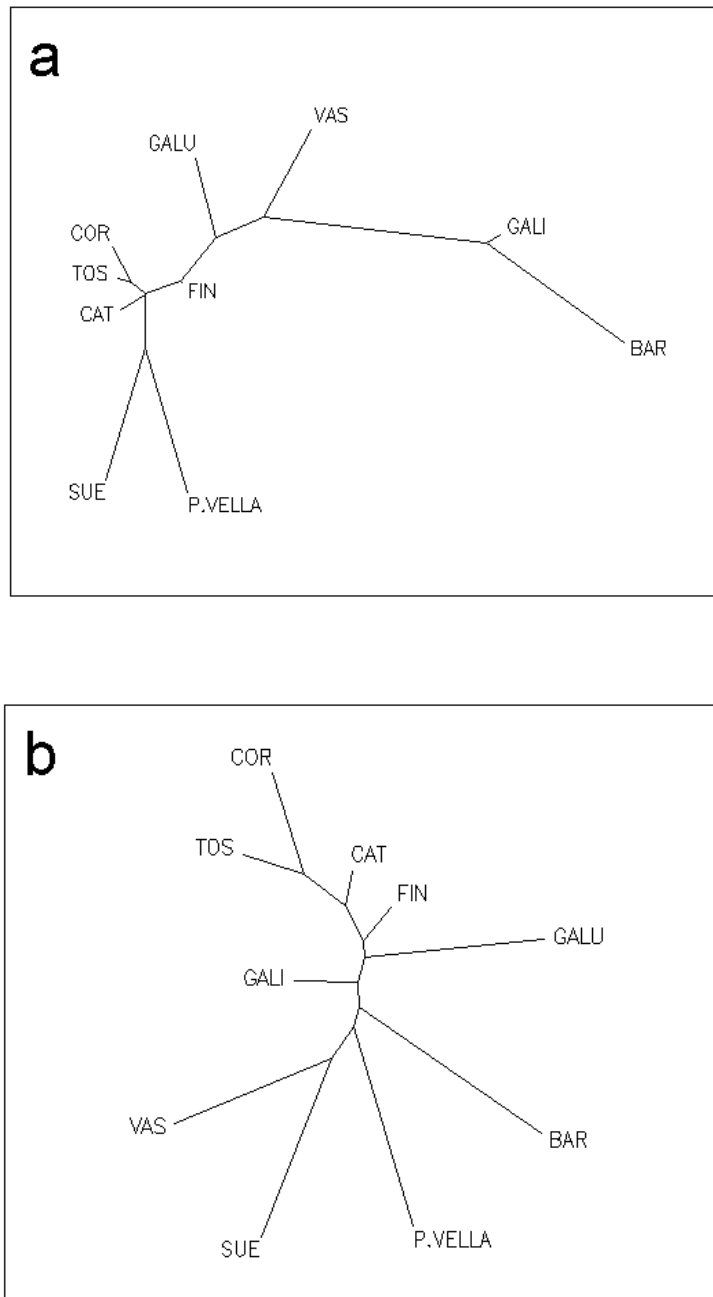


Figura R-7 Árboles filogenéticos sin raíz contruidos con el método de Fitch-Margoliash mediante el programa FITCH del paquete informático Phylip, utilizando la opción G y la opción J (225, 10), a partir de las matrices de distancias **(a)** de Reynolds y col. (1983) y **(b)** de la medida de la cuerda (Cavalli-Sforza y Edwards, 1967).

Llama la atención la situación de la población de la Plaça Vella, ya que aparece relacionada con la población de Suecia. Esta característica puede ser un tanto engañosa, ya que estos árboles no están enraizados, por lo que dependiendo de dónde se sitúe la

raíz, la relación de la población de la Plaça Vella con la población sueca y con el resto de las poblaciones podría ser distinta, sobretodo considerando que la distancia entre la población de la Plaça Vella y la población Sueca es considerable.

El árbol de Fitch-Margoliash construido utilizando las distancias de la medida de la cuerda (Fig. R-7b), presenta una mayor homogeneidad entre las poblaciones, tal como se deduce de la matriz de distancias. La mayoría de las poblaciones aparecen equidistantes y destacan por su cercanía las poblaciones de Córcega y Toscana, que también aparecían cercanamente relacionadas en el árbol construido a partir de la distancia de Reynolds. Este hecho podría ser explicado por la relación histórica de la población Corza con la Toscana, ya que por un período la isla de Córcega perteneció a Italia y de hecho uno de los dialectos que se hablan en la isla está relacionado con el Toscano.

Las poblaciones actual de Cataluña y la de Finlandia ocupan una posición intermedia entre las poblaciones de Córcega y Toscana y el resto de poblaciones analizadas. Esta característica también se observa en el árbol de las distancias de Reynolds, en el que estas cuatro poblaciones forman un grupo cercanamente relacionado. La población Catalana ha tenido una influencia importante, en tiempos históricos, en la isla de la Cerdeña, lo que podría explicar esta similitud. Por otra parte, podría parecer extraño que la población de Finlandia esté incluida en este grupo, sin embargo otros estudios han encontrado que esta población prácticamente no se diferencia del resto de poblaciones europeas en cuanto a su acervo mitocondrial (Sajantila et al. 1995).

En contraste con el árbol inferido a partir de las distancias de Reynolds, la población de la Plaça Vella se separa de la población Sueca, que aparece más relacionada con la población Vasca, sin embargo, la distancia también es considerable y dependiendo dónde se sitúe la raíz, la topología del árbol puede definirse de una u otra manera. El árbol obtenido con el método de máxima verosimilitud (Fig. R-8) presenta una topología con características de los árboles anteriores pero sin ser totalmente similar a ninguno de ellos. En un grupo central aparecen las poblaciones de Galicia, Finlandia, Toscana,

Galura, Córcega y actual de Cataluña, quedando en extremos opuestos la población del País Vasco por un lado y las poblaciones de Suecia y de la Plaça Vella por el otro.

Como se había mencionado anteriormente, la posición de la raíz determinará si la población de la Plaça Vella está más relacionada con la población de Suecia o con el grupo formado por las poblaciones CAT, COR, TOS y FIN.

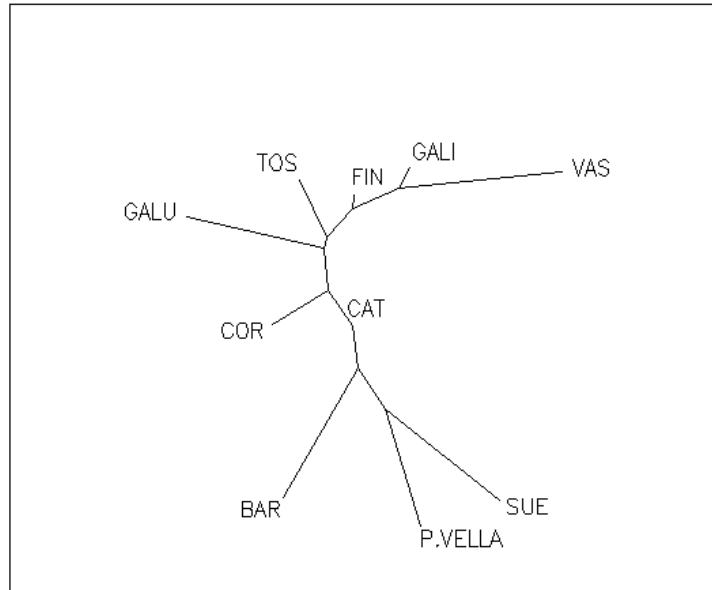
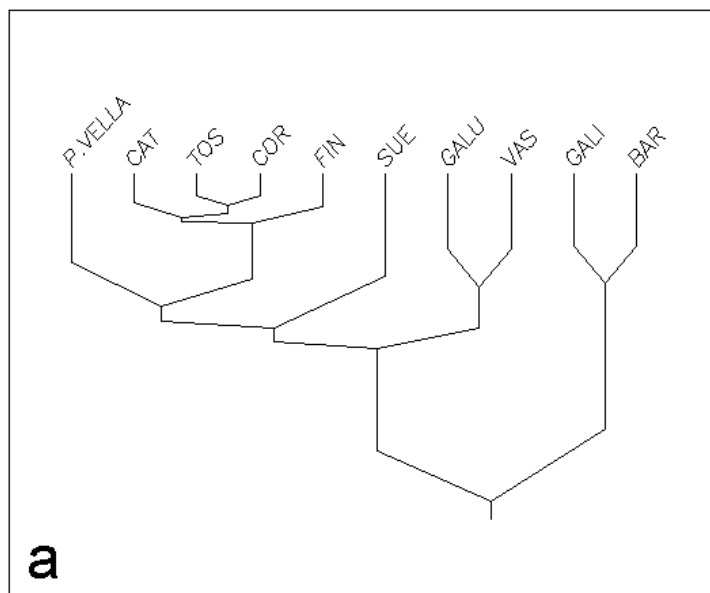


Figura R-8 Árbol filogenético sin raíz construido con el método de máxima verosimilitud mediante el programa CONTML del paquete informático Phylip, utilizando la opción G y la opción J (225,1000).

Por este motivo, se construyeron árboles con raíz, utilizando el método UPGMA asumiendo la misma tasa evolutiva para todas las poblaciones. En el árbol obtenido a partir de las distancias de Reynolds (Fig. R-9a) se observa que existe una primera separación entre las poblaciones de Galicia y Barbagia por un lado y el resto de poblaciones por el otro. También destaca la agrupación de las poblaciones del País Vasco y de Galura. A continuación, la población sueca se separa del resto y la población de la Plaça Vella quedaría en un grupo de dos ramas principales, una de ellas conduce a las poblaciones FIN, COR, TOS, y CAT y la otra exclusivamente a la población de la Plaça Vella. También se observa que las poblaciones TOS, COR y CAT, forman un grupo monofilético.

El árbol desarrollado a partir de las distancias de la medida de la cuerda (Fig. R-9b), es similar al descrito anteriormente y se diferencia principalmente por la exclusión de las poblaciones de Suecia y País Vasco después de la separación de la población BAR y el resto. Destaca también la posición de Galicia que en este árbol queda incluida en el grupo de las poblaciones CAT, COR, TOS y FIN. La población de la Plaça Vella continúa relacionada con este último grupo, aunque separada por una distancia un poco mayor. En ambos árboles, la población de Suecia está más alejada del grupo principal que la población de la Plaça Vella, lo que concuerda con la situación geográfica y la historia de las poblaciones.



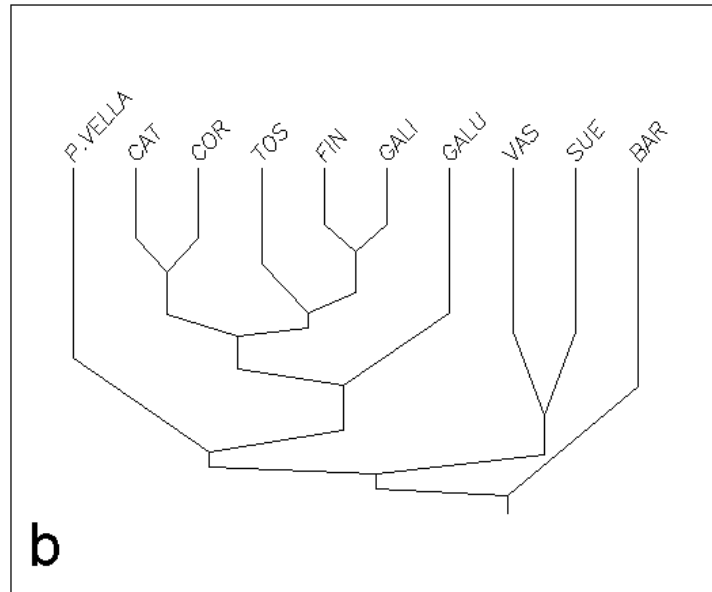


Figura R-9 Árboles filogenéticos con raíz construidos con el método UPGMA mediante el programa NEIGHBOR del paquete informático Phylip, a partir de las matrices de distancias **(a)** de Reynolds y col. (1983) y **(b)** de la medida de la cuerda (Cavali-Sforza y Edwards, 1967).

Un punto importante a tener en cuenta es que estos métodos asumen que las poblaciones analizadas son contemporáneas y la población de la Plaça Vella no cumple esta presunción. Sería interesante investigar hasta que punto esta característica puede explicar la relativa separación de la población de la Plaça Vella, aunque el hecho de que presente una inusual baja frecuencia del haplogrupo H es una alternativa bastante plausible. Más aún, la estructura de estos árboles, debe interpretarse con cierta precaución, ya que las poblaciones pueden parecer filogenéticamente cercanas cuando en realidad no lo están debido al hecho de que procesos similares o equivalentes pueden conducir a distribuciones de haplogrupos similares, como en el caso de las poblaciones de Barbagia y Galicia. Si interpretáramos el árbol de la Figura R-9a de una forma directa, tendríamos que concluir que estas poblaciones están más relacionadas filogenéticamente entre sí que con ninguna otra de las poblaciones analizadas, lo que sería difícil de sostener si se considera la historia de ambas poblaciones o si se revisan otro tipo de evidencias.

Estas consideraciones nos hacen pensar que la reconstrucción filogenética a partir de la frecuencia de los haplogrupos en las poblaciones es un método más bien pobre, ya que además del problema consabido de la base común del acervo mitocondrial de todas las poblaciones europeas, se presenta el problema de la sensibilidad de las frecuencias a la deriva potenciada por el aislamiento y a la pérdida de información que produce el agrupamiento de distintos tipos mitocondriales dentro de un mismo haplogrupo.

Secuencias de la región de control

Índices de diversidad

Porcentaje de haplotipos diferentes

En la Tabla R-40 se muestran los porcentajes de haplotipos distintos ($\%k$) que presentan las poblaciones para cada uno de los segmentos analizados. Las poblaciones se muestran ordenadas de menor a mayor según el $\%k$ observado en el segmento de 360pb, que abarca las posiciones 16024 a 16383.

Población	N	Tamaño del segmento (pb)							
		391		360		302		153	
		<i>k</i>	$\%k$	<i>k</i>	$\%k$	<i>k</i>	$\%k$	<i>k</i>	$\%k$
HAD	17	-	-	4	23.52	4	23.52	4	23.52
ARG	85	-	-	-	-	30	35.29	23	27.05
GALI	92	-	-	53	57.60	51	55.43	35	38.04
VAS	45	-	-	27	60	27	60	20	44.44
SAR	69	-	-	45	65.21	44	63.76	30	43.47
P.VELLA	24	-	-	-	-	-	-	12	50
BRI	100	68	68	68	68	68	68	52	52
FIN	47	-	-	32	68.08	32	68.08	22	46.80
GALU	27	-	-	21	77.77	20	74.07	16	59.25
TOS	48	-	-	39	81.25	39	81.25	34	70.83
CAT	25	22	88	22	88	21	84	17	68
TRK	45	-	-	40	88.88	40	88.88	33	73.33
ASI	22	-	-	22	100	22	100	21	95.45

Tabla R-40 Número de haplotipos diferentes (*k*) que presenta cada población en segmentos de distintos tamaños. El porcentaje de haplotipos ($\%k$) fue calculado en relación al número de individuos analizados (N). Ubicación de los segmentos: 391, 16001-16391; 360, 16024-16383; 302, 16069-16370 y 153, 16210-16362, siguiendo la numeración de la secuencia de referencia de Cambridge.

La población con menor número de haplotipos diferentes es la población subsahariana de Hadza (23.5%), aunque también es la de menor tamaño de muestra con 17 individuos. La población con mayor diversidad de haplotipos es la Asiática, con un 100%; hecho explicado por la procedencia heterogénea de la muestra. La población de Galicia, muestra un índice relativamente bajo y el resto de poblaciones se sitúa entre el 60% de haplotipos diferentes de la población Vasca y el 88.9% de la población Turca. La población Catalana actual (CAT) presenta un alto porcentaje (88%) a pesar del bajo tamaño muestral (n=25).

En general, el patrón de diversidad no cambia para el $%k$ observado en el segmento de 302pb (16069-16370), sin embargo, si es diferente el que se observa para el segmento que incluye los sitios 16210 a 16362 de 153pb. En este segmento las poblaciones de Finlandia y actual de Cataluña presentan una menor diversidad relativa, pues la primera se sitúa antes que la población Británica y la segunda antes de la población Toscana, contrariamente a lo observado en los segmentos más grandes. La población de la Plaça Vella muestra una diversidad intermedia (50%), sin embargo, no existe diferencia significativa entre su proporción de haplotipos diferentes y la proporción que presenta la población actual de Cataluña (68%) en este segmento (prueba exacta de Fisher bilateral, $P= 0.2516$).

Finalmente cabe destacar que el porcentaje de haplotipos diferentes disminuye en todas las poblaciones a medida que se reduce el tamaño del segmento analizado, lo cual es de esperar debido a que existen secuencias que siendo diferentes para segmentos grandes no lo son para segmentos más cortos.

Diversidad genética de Nei

El índice de diversidad genética de Nei (H) se basa en la frecuencia de los haplotipos y es un estimador más robusto que el simple porcentaje de haplotipos. En general, el rango de variabilidad es menor y en las poblaciones europeas va de 0.9295 de

Galicia a 0.9939 de Turquía (Tabla R-41). Asimismo, aunque este índice también muestra la disminución de la diversidad a medida que disminuye el tamaño del segmento, las diferencias son menores. Algunas poblaciones presentan una disminución más pronunciada en el segmento de 153pb, como ocurre con la población de Galicia, sin embargo, para la mayoría de poblaciones esta disminución no es tan notable. En el caso de la población Catalana actual este índice disminuye de 0.9867 para el segmento de 391pb a 0.95 para el de 153pb, en tanto que el $\%k$ disminuía del 88 al 68% en los respectivos segmentos.

Población	N	Tamaño del segmento (pb)			
		391	360	302	153
		$_ \pm$ d.e.	$_ \pm$ d.e.	$_ \pm$ d.e.	$_ \pm$ d.e.
HAD	17	-	0.5956 \pm 0.0991	0.5956 \pm 0.0991	0.5956 \pm 0.0991
GALI	92	-	0.9295 \pm 0.0226	0.9233 \pm 0.0239	0.7387 \pm 0.0516
SAR	69	-	0.9437 \pm 0.0220	0.9369 \pm 0.0239	0.8610 \pm 0.0389
ARG	85	-	-	0.9431 \pm 0.0103	0.9311 \pm 0.0101
VAS	45	-	0.9485 \pm 0.0214	0.9485 \pm 0.0214	0.8747 \pm 0.0415
P.VELLA	24	-	-	-	0.8949 \pm 0.0453
GALU	27	-	0.9658 \pm 0.0253	0.9516 \pm 0.0320	0.9060 \pm 0.0459
TOS	48	-	0.9672 \pm 0.0197	0.9672 \pm 0.0197	0.9184 \pm 0.0361
FIN	47	-	0.9732 \pm 0.0123	0.9732 \pm 0.0123	0.9214 \pm 0.0285
BRI	100	0.9752 \pm 0.0092	0.9752 \pm 0.0092	0.9752 \pm 0.0092	0.9016 \pm 0.0274
CAT	25	0.9867 \pm 0.0167	0.9867 \pm 0.0167	0.9833 \pm 0.0171	0.9500 \pm 0.0293
TRK	45	-	0.9939 \pm 0.0065	0.9939 \pm 0.0065	0.9768 \pm 0.0127
ASI	22	-	1.0000 \pm 0.0137	1.0000 \pm 0.0137	0.9957 \pm 0.0153

Tabla R-41 Diversidad genética ($_$) y su desviación estándar (\pm d.e.) según la fórmula de Nei (1987), calculada para los distintos segmentos disponibles de cada población en base a sus secuencias mitocondriales. Se han ordenado de menor a mayor diversidad según el segmento de 360pb.

Mediante este índice, se observa que la población Catalana actual presenta una elevada diversidad, superada únicamente por la población de Turquía (y la asiática), en contraste con las otras dos poblaciones ibéricas analizadas (VAS y GALI) y con la población antigua de la Plaça Vella, que continúa situada en una situación más o menos intermedia, con mayor diversidad que las poblaciones de Galicia, Cerdeña y País Vasco. Un cambio notable, es la posición de la población Toscana que se sitúa en un sitio intermedio con menos diversidad que las poblaciones FIN y BRI.

Por lo que respecta a las diferencias en relación a los distintos segmentos analizados, el patrón de diversidad poblacional definido al analizar el segmento de 360pb es prácticamente el mismo para el segmento de 302pb y presenta pocos cambios para el segmento de 153pb. Destaca la pérdida de variabilidad de la población BRI que se sitúa antes que las poblaciones TOS y FIN para este segmento. Esta diferencia puede deberse a que en la población BRI exista un mayor número de linajes definidos por substituciones ocurridas en la zona no comprendida en el segmento de 153pb, en comparación con el que presentan la mayoría de poblaciones. En el análisis de la distribución de haplogrupos expuesto en la sección anterior, se observó que esta población mostraba un alto porcentaje de tipos mitocondriales del haplogrupo H. Este es un aspecto interesante que valdría la pena investigar en el futuro.

Porcentaje de sitios variables

Los porcentajes de sitios variables (%v) en las poblaciones analizadas se presentan en la Tabla R-42, en la que las poblaciones han sido ordenadas de menor a mayor según el porcentaje que presentan en el segmento de 360pb.

Población	N	Tamaño del segmento (pb)							
		391		360		302		153	
		v	%v	v	%v	v	%v	v	%v
HAD	17	-	-	11	3.05	11	3.64	10	6.54
P.VELLA	24	-	-	-	-	-	-	12	7.84
GALU	27	-	-	31	8.61	28	9.27	16	10.46
VAS	45	-	-	32	8.88	32	10.59	18	11.76
FIN	47	-	-	37	10.27	36	11.92	21	13.73
CAT	25	41	10.48	38	10.55	35	11.58	19	12.42
ARG	85	-	-	-	-	37	12.25	22	14.38
SAR	69	-	-	52	14.44	49	16.22	26	16.99
TOS	48	-	-	55	15.27	54	17.88	36	23.53
TRK	45	-	-	56	15.55	54	17.88	36	23.53
GALI	92	-	-	56	15.55	55	18.21	34	22.22
ASI	22	-	-	58	16.11	56	18.54	32	20.92
BRI	100	69	17.64	68	18.88	65	21.52	44	28.76

Tabla R-42 Número de sitios variables (v) que presenta cada población en segmentos de distintos tamaños. El porcentaje de sitios variables (%v) fue calculado en relación al número total de bases que comprende el segmento analizado. Las poblaciones se han ordenado de menor a mayor según los %v que presentan en el segmento de 360pb.

El rango en las poblaciones europeas va del 8.61% de Galura al 18.9% de la población Británica, la cual presenta un porcentaje más alto incluso que la población Asiática. Sería interesante investigar hasta que punto este resultado puede deberse al tamaño de muestra, pues la muestra Británica comprende 100 individuos, sobretodo teniendo en cuenta que el tamaño de la muestra no parece afectar al porcentaje de haplotipos distintos.

En situación muy similar se encuentra la población de Galicia que presenta un alto porcentaje de sitios variables y un bajo porcentaje de haplotipos distintos y que también consiste de un número importante de individuos (N=92). Más aún, la población P.VELLA, presenta el segundo porcentaje más bajo (7.84%) después de la población Hadza, y la población CAT también presenta una diversidad menor y ambas poblaciones presentan un número de muestra reducido, lo que podría justificar el bajo porcentaje de sitios variables que presentan.

Así pues, este índice de diversidad parece estar más influido por el tamaño de muestra que los índices basados en el número de haplotipos. Por otra parte, cabe mencionar que no existe diferencia significativa entre P.VELLA y CAT en cuanto a la proporción de sitios variables que presentan en el segmento de 153pb (prueba exacta de Fisher bilateral, $P= 0.2554$).

El orden de las poblaciones no cambia si se analiza el segmento de 302pb y varía muy poco si se analiza el de 153pb. Sin embargo, si es observable que en todas las poblaciones a medida que disminuye el segmento analizado, el porcentaje de sitios variables aumenta, contrariamente a la tendencia que presenta el porcentaje de haplotipos. Esto indica que la zona comprendida por el segmento de 153pb es la que presenta menor estabilidad dentro del segmento hipervariable I, lo que está en relación con una mayor tasa de mutación.

Diversidad nucleotídica

En la Tabla R-43 se han ordenado las poblaciones de menor a mayor diversidad nucleotídica (π) que presentan para el segmento de 360pb. De acuerdo con este índice, la población de Galicia es la que presenta una menor diversidad entre los nucleótidos de sus secuencias mitocondriales (0.009211), a pesar de contener un alto porcentaje de sitios variables. Estas características son compatibles con una expansión reciente de la población, tal como han descrito los autores del artículo en el que originalmente se analizaron estas secuencias.

La población VAS presenta el segundo índice más bajo (0.009515) y en general el orden de las poblaciones presenta características similares a las descritas para los índices de diversidad haplotípica, con la excepción de la población P.VELLA que aparece en una posición de baja diversidad, más similar a la que ocupa según el porcentaje de sitios variables. Este dato es interesante en tanto que este índice no parece estar influido por el tamaño de la muestra como lo está el $\%v$, pues poblaciones como BRI y GALI con un número importante de individuos no ocupan sitios de mayor diversidad.

Un punto a destacar, es que la población Hadza, que ocupa el sitio de menor diversidad en los anteriores índices, se sitúa en una posición más intermedia de acuerdo a su diversidad nucleotídica. La población de Argelia también muestra una relativa mayor diversidad con este índice. Estas características son compatibles con un estado de equilibrio o con una expansión poblacional más antigua.

Población	N	Tamaño del segmento (pb)			
		391	360	302	153
		$\pi \pm \text{d.e.}$	$\pi \pm \text{d.e.}$	$\pi \pm \text{d.e.}$	$\pi \pm \text{d.e.}$
GALI	92	-	0.009211 \pm 0.005291	0.010998 \pm 0.006316	0.012627 \pm 0.008021
VAS	45	-	0.009515 \pm 0.005502	0.011495 \pm 0.006634	0.014241 \pm 0.008923
P.VELLA	24	-	-	-	0.014598 \pm 0.009290
FIN	47	-	0.011890 \pm 0.006657	0.014238 \pm 0.007966	0.018405 \pm 0.010977
HAD	17	-	0.012276 \pm 0.007229	0.014247 \pm 0.008374	0.024003 \pm 0.014293
SAR	69	-	0.012608 \pm 0.006957	0.014969 \pm 0.008264	0.015839 \pm 0.009641
BRI	100	0.012441 \pm 0.006777	0.013238 \pm 0.007228	0.015719 \pm 0.008587	0.019898 \pm 0.011577
GALU	27	-	0.013295 \pm 0.007465	0.015322 \pm 0.008638	0.014673 \pm 0.009282
CAT	25	0.014316 \pm 0.007924	0.014946 \pm 0.008308	0.016956 \pm 0.009477	0.021353 \pm 0.012675
ARG	85	-	-	0.017458 \pm 0.009439	0.021789 \pm 0.012516
TOS	48	-	0.015148 \pm 0.008237	0.018253 \pm 0.009914	0.022948 \pm 0.013201
TRK	45	-	0.016252 \pm 0.008785	0.019254 \pm 0.010415	0.024782 \pm 0.014119
ASI	22	-	0.027381 \pm 0.014550	0.032337 \pm 0.017179	0.040602 \pm 0.022371

Tabla R-43 Diversidad nucleotídica (π) y su desviación estándar (\pm d.e.) estimada para los distintos segmentos disponibles de cada población. Las poblaciones se presentan ordenadas de menor a mayor diversidad según el segmento de 360pb.

En cuanto a las diferencias según el tamaño del segmento, es observable que el orden de las poblaciones no se ve alterado para el segmento de 302pb, pero si para el de 153pb, destacando la población Hadza, que para este segmento presenta una mayor diversidad que la mayoría de las poblaciones para situarse justo detrás de Asia y Turquía. Por su parte, las poblaciones de Galura y Cerdeña pierden variabilidad para este segmento en relación a otras poblaciones. Estos cambios de posiciones en la jerarquía de diversidad deben tenerse en cuenta cuando se reconstruyen árboles filogenéticos utilizando los distintos tamaños de segmento, ya que podrían conducir a diferencias en la topología según el segmento analizado, como se verá en la sección correspondiente.

Por otra parte, al igual que con el porcentaje de sitios variables, se observa un incremento importante de diversidad nucleotídica en todas las poblaciones a medida que el segmento analizado es más pequeño.

Así, en los índices basados en haplotipos, la diversidad disminuye con el tamaño del segmento, en tanto que los índices que reflejan la variabilidad a nivel de nucleótido, la diversidad aumenta en los segmentos de menor tamaño (Tabla R-44). Se debe ser

cauto en la interpretación del aumento referido, ya que en realidad el aumento en la diversidad puede estar más relacionado con la tasa de mutación (por sitio) del segmento analizado que con el tamaño en sí. La zona centro – derecha del segmento hipervariable I presenta una divergencia mayor en las secuencias de las poblaciones analizadas.

Índice	Segmento (pb)		
	360	302	153
$%k$	70.74237	69.32651	52.54081
$\bar{—}$	0.96123	0.95740	0.88519
$%v$	12.81250	14.65232	16.41249
π	0.01248	0.01474	0.01718

Tabla R-44 Promedio de cuatro índices de diversidad en poblaciones europeas: BRI, CAT, FIN, GALI, GALU, SAR, TOS y VAS. Para el segmento de 153pb se incluye además P.VELLA.

Estos resultados pueden ser de interés a la hora de escoger el segmento a amplificar, sobre todo en los estudios de poblaciones antiguas en los que la dificultad de recuperar segmentos grandes aumenta con su tamaño. El hecho de que los segmentos de menor longitud aquí analizados presenten una mayor diversidad nucleotídica, los hace interesantes para métodos de análisis que utilizan esta información, como pueden ser los métodos de reconstrucción filogenética basados en distancias (o el método de las redes medias), sin embargo, al presentar una menor diversidad haplotípica, los segmentos pequeños pueden resultar poco adecuados para estudios basados directa o indirectamente en la frecuencia de los haplotipos.

Historia demográfica: distribución de las diferencias por parejas

La distribución de las diferencias entre pares de individuos de la población Catalana actual (CAT) presenta una forma acampanada similar a la que muestran el resto de poblaciones europeas. La forma general de la distribución no cambia para los segmentos de 391, 360 y 302pb, sin embargo, para el segmento de 153pb la curva se desplaza notablemente hacia la izquierda (Fig. R-10). Este fenómeno se observa también en otras poblaciones (Fig. R-11) y seguramente se debe a la pérdida de diversidad haplotípica que ocurre con la disminución del fragmento analizado.

La distribución observada en cada uno de los segmentos fue ajustada al modelo de expansión súbita (Figuras R-12 y R-13). La bondad de ajuste al modelo fue estimada mediante una prueba de χ^2 y las diferencias entre los valores observados y los esperados son significativas para los fragmentos de 391 y 360 pb, sin embargo, el ajuste mejora para los segmentos de menor tamaño, en los que no existen diferencias significativas entre lo observado y lo esperado, especialmente para el fragmento de 153pb (Tabla R-45). No deja de ser notable que el peor ajuste se obtuvo para el fragmento de 360pb.

Un aspecto interesante de la distribución de la población Catalana es la pendiente tan pronunciada de la cara derecha de la curva que se observa en los fragmentos grandes. De acuerdo con Rogers y Harpending (1992) una de las diferencias entre la forma de la distribución producida por una expansión súbita a partir de un estado de equilibrio y la forma que produciría una expansión después de un cuello de botella de corta duración, radica precisamente en las características de la cara derecha de la curva. Una pendiente más pronunciada correspondería al modelo del cuello de botella. No obstante, este modelo no ha sido estudiado en profundidad, por lo que no se pueden obtener conclusiones definitivas. Por otra parte, asumiendo el modelo de expansión súbita, una pendiente más pronunciada indica que la población al inicio de la expansión presentaba un tamaño mayor en relación al tamaño inicial de las poblaciones que muestran una pendiente más suave.

Otro escenario de historia demográfica es el modelo de crecimiento exponencial continuo. De acuerdo con Slatkin y Hudson (1991), las poblaciones que han atravesado por un proceso demográfico de este tipo, presentan una distribución de diferencias por parejas muy similar a la distribución de Poisson.

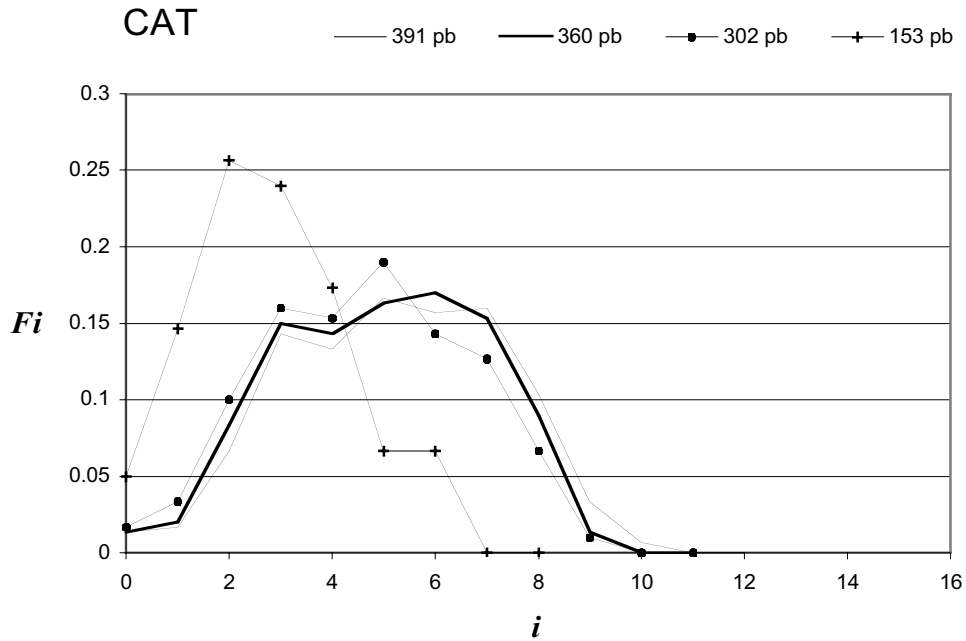


Figura R-10 Distribución de las diferencias por parejas observada en 4 segmentos de distintos tamaños de la población Catalana actual (CAT); i , número de diferencias; F_i , frecuencia relativa de i .

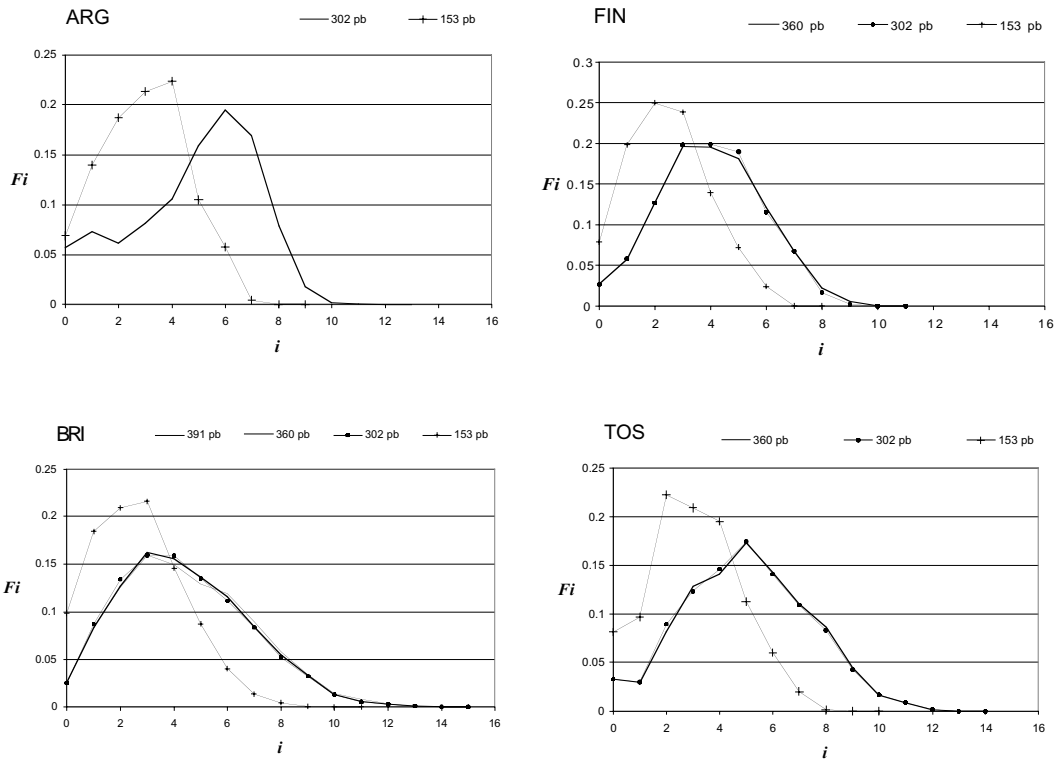


Figura R-11 Distribución de las diferencias por parejas observada en segmentos de distintos tamaños de 4 poblaciones; i , número de diferencias; F_i , frecuencia relativa de i .

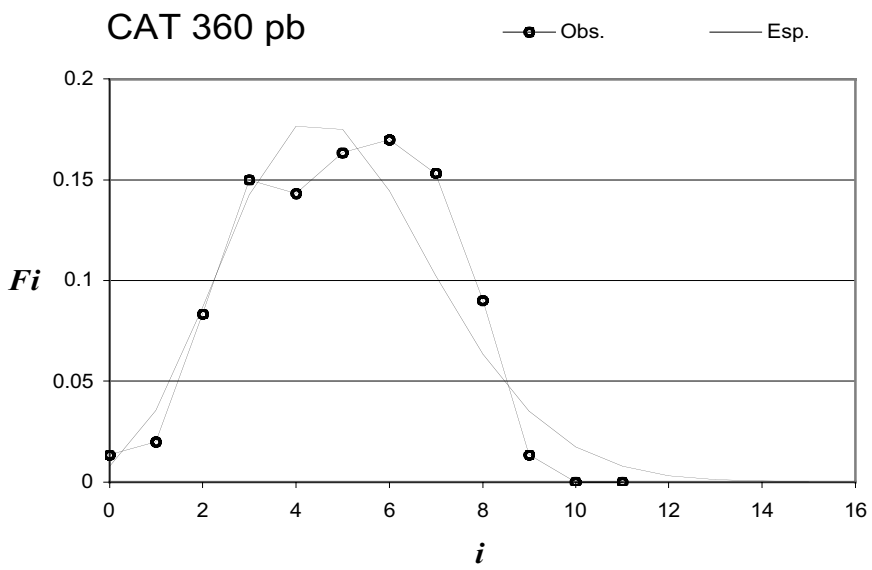
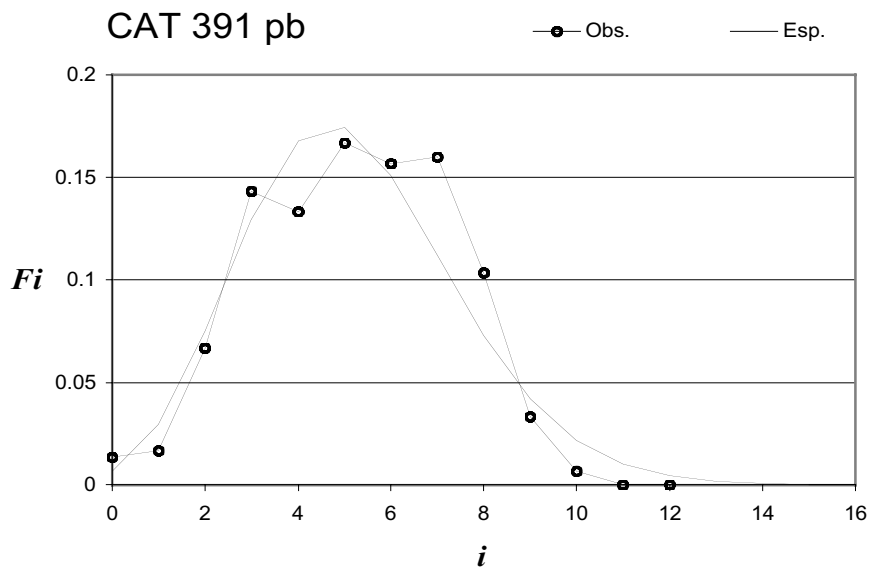


Figura R-12 Distribución observada (Obs.) de las diferencias por parejas y su distribución esperada (Esp.) bajo el modelo de expansión súbita (según el método de Rogers, 1995), para los fragmentos de 391 (arriba) y 360 pb (abajo) de la población actual de Cataluña; i , número de diferencias por parejas; F_i , frecuencia relativa de i .

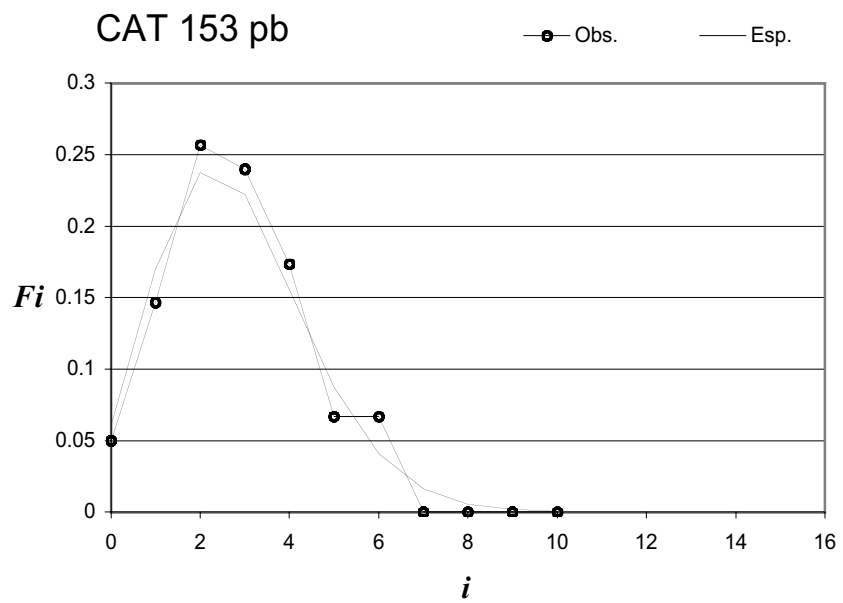
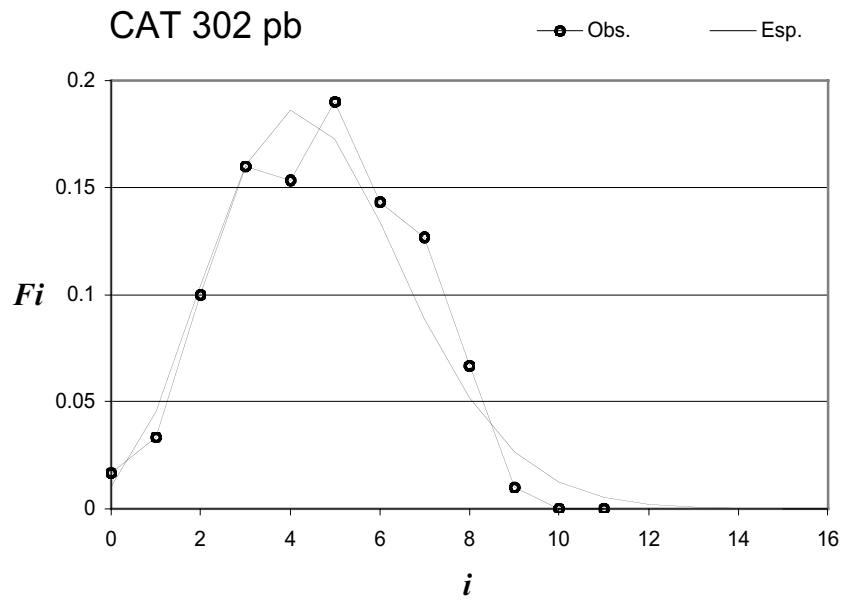


Figura R-13 Distribución observada (Obs.) de las diferencias por parejas y su distribución esperada (Esp.) bajo el modelo de expansión súbita (según el método de Rogers, 1995), para los fragmentos de 302 (arriba) y 153 pb (abajo) de la población actual de Cataluña; i , número de diferencias por parejas; F_i , frecuencia relativa de i .

	391			360			302			153		
	χ^2	g.l.	P	χ^2	g.l.	P	χ^2	g.l.	P	χ^2	g.l.	P
CAT												
Rogers	20.5573	9	0.014768	21.6837	8	0.005537	13.8249	8	0.086444	9.3661	5	0.095323
Poisson	21.3750	9	0.011086	21.9565	8	0.004997	14.0846	8	0.079587	9.1951	5	0.101532
P.VELLA												
Rogers	-	-	-	-	-	-	-	-	-	4.7538	6	0.575752
Poisson	-	-	-	-	-	-	-	-	-	4.6917	6	0.583916

Tabla R-45 Prueba χ^2 de bondad de ajuste al modelo de expansión súbita (Rogers, 1995) y a la distribución de Poisson con parámetro μ = media de las diferencias por parejas. El valor de P indica la probabilidad de que los valores observados y esperados sean iguales; g.l., grados de libertad.

El ajuste de la distribución observada en los distintos segmentos a esta distribución (Tabla R-45) es muy parecido al ajuste que se obtiene con el modelo de expansión súbita, sin embargo es un poco menos adecuado en todos los segmentos excepto para el de 153pb, hecho que confirma que con este segmento se pierde información importante para el análisis de la distribución. Sin embargo, las diferencias en el ajuste a las dos distribuciones teóricas son muy reducidas (en todos los segmentos) y esto contrasta con las observaciones hechas en otras poblaciones, como la vasca, en las que el ajuste a la distribución de Poisson es mucho más pobre que el ajuste al modelo de expansión súbita (Bertranpetit et al. 1995).

Así pues, las diferencias en la forma de la distribución de la población Catalana respecto a la Vasca y a otras poblaciones europeas podrían indicar una historia demográfica ligeramente diferente, aunque no se deben olvidar las limitaciones inherentes a la prueba χ^2 de bondad de ajuste cuando es aplicada en este contexto (ver Slatkin y Hudson, 1991).

Una explicación alternativa para la forma particular de la distribución en la población CAT, podría estar relacionada con el bajo número de muestras analizadas ($n=25$), no obstante, Bertranpetit y col. (1995) demostraron que incluso con 20 individuos la forma de la distribución resulta idéntica a la obtenida con 45 individuos. Además, los errores estándar (obtenidos mediante *bootstrap*) que presenta esta distribución son muy reducidos (Fig. R-14), por lo que la forma particular de la curva no se debe a una falta de robustez. Cabe apuntar, que para observar el efecto de la reducción del tamaño del segmento en la robustez de la distribución, se derivaron los errores

estándar para el fragmento de 153pb y el resultado demuestra que la reducción del tamaño no incrementa el error en modo alguno, a pesar de que la forma de la

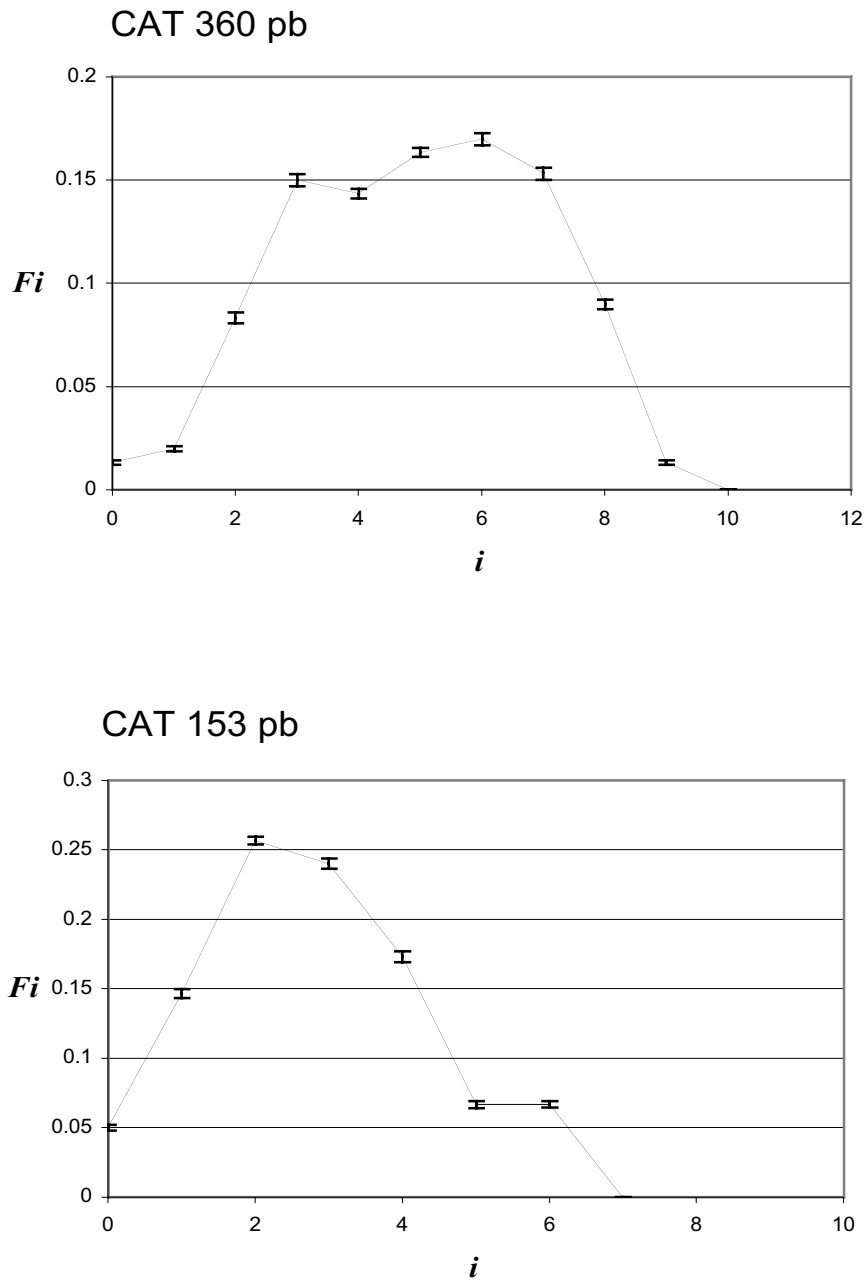


Figura R-14 Errores estándar para cada valor de i derivados mediante 1000 iteraciones de “bootstrap” en los fragmentos de 360 (arriba) y 153 (abajo) pares de bases de la población actual de Cataluña; i , número de diferencias por parejas; Fi , frecuencia relativa de i .

curva si se vea alterada (Fig. R-14). Todo esto nos indica que la pérdida de información relevante para el análisis de las diferencias por parejas que ocurre con este segmento no puede evitarse con un número de muestras mayor.

Otro aspecto interesante es el posible carácter bimodal de la distribución de la población CAT, que presenta picos en 3 y 6 diferencias. Algunos autores (Bertranpetit et al. 1995; ver también Richards et al. 1998) han apuntado la posibilidad de que las distribuciones bimodales se deban a superposiciones de dos eventos de expansión súbita. Sin embargo, la teoría predice que una vez que la población ha incrementado substancialmente su tamaño, el crecimiento posterior no tiene un efecto notorio (Rogers y Harpending, 1992), por lo que una expansión inicial obscurecería los efectos de expansiones posteriores por un período de tiempo muy largo (Rogers, 1995). Otras posibilidades destacadas por los mismos autores (Bertranpetit et al. 1995), relacionadas con la coalescencia o con la dinámica del proceso de crecimiento, parecen pues, más plausibles.

Finalmente, cabe destacar que a pesar de las diferencias que presenta la distribución de la población CAT, no se puede considerar que su historia demográfica sea totalmente diferente al resto de poblaciones no africanas, pues dista mucho de ser una población en equilibrio. De hecho, el índice de irregularidad de Harpending de la distribución de esta población (0.0193), se sitúa dentro del rango observado para el conjunto de poblaciones europeas aquí analizadas, que va de 0.01 de la población BRI a 0.0410 de la población de Galura, para el segmento de 360pb (Tabla R-46).

Población	N	Segmento (pb)			
		391	360	302	153
		<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
ASI	22	-	0.0089	0.0103	0.0221
BRI	100	0.0100	0.0100	0.0100	0.0190
TOS	48	-	0.0110	0.0110	0.0279
TRK	45	-	0.0126	0.0132	0.0332
SAR	69	-	0.0160	0.0120	0.0240
GALI	92	-	0.0190	0.0180	0.0160
ARG	85	-	-	0.0180	0.0273
CAT	25	0.0186	0.0193	0.0191	0.0420
FIN	47	-	0.0196	0.0214	0.0344
VAS	45	-	0.0290	0.0290	0.0436
GALU	27	-	0.0410	0.0290	0.0626
P.VELLA	24	-	-	-	0.0665
HAD	17	-	0.4020	0.4020	0.4490

Tabla R-46 Índice de irregularidad de Harpending (*r*) de las distribuciones de los distintos segmentos de las poblaciones analizadas.

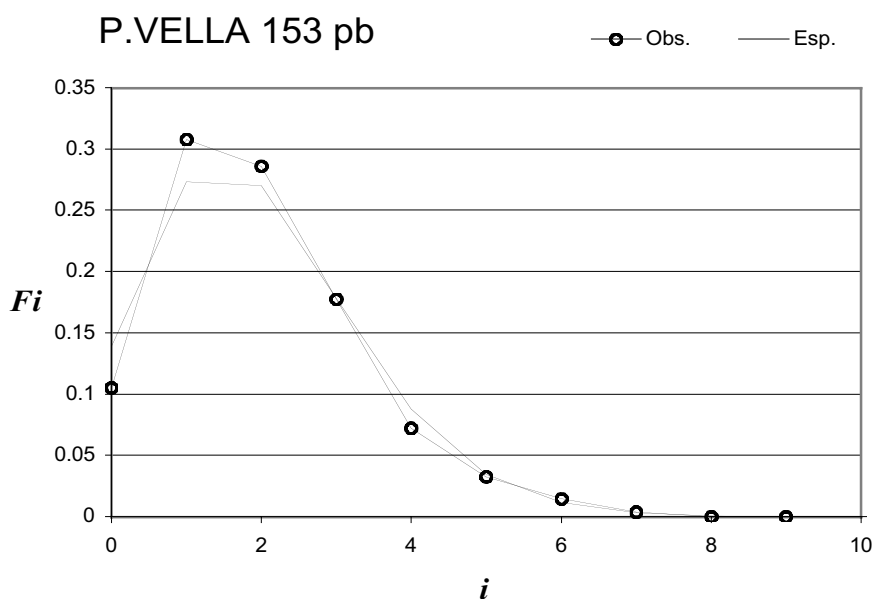
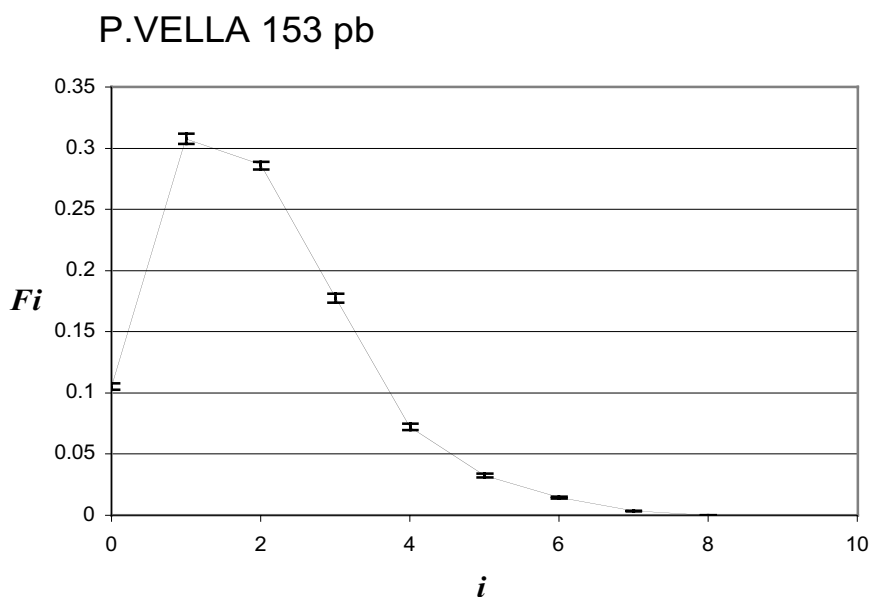


Figura R-15 Distribución de las diferencias por parejas en la población de la Plaça Vella. El gráfico superior muestra los errores estándar para cada valor de i derivados mediante 1000 iteraciones de “bootstrap”; en el gráfico inferior se comparan los valores observados (Obs.) con los valores esperados (Esp.) bajo el modelo de expansión súbita (según el método de Rogers, 1995); i , número de diferencias; F_i , frecuencia relativa de i .

La distribución de las diferencias por parejas de la población de la Plaça Vella presenta una forma general y una robustez similar a la observada para el fragmento de 153pb de la población actual. Asimismo, presenta un buen ajuste al modelo de expansión súbita (Fig. R-15) y un ajuste ligeramente mejor a la distribución de Poisson, al igual que para el segmento de la misma longitud de la población CAT (ver Tabla R-45, Pag. 335).

Es evidente que la información obtenida con segmentos pequeños utilizando este método es más bien limitada y por este motivo, la curva obtenida de P.VELLA no puede usarse para contrastar las peculiaridades mostradas por los fragmentos grandes de la población Catalana actual. Sin embargo, no deja de ser estimulante el hecho de haber encontrado similitudes en los fragmentos pequeños de ambas muestras poblacionales.

El índice de irregularidad de Harpending para la población de la Plaça Vella (0.0665), resultó ser el más alto de todas las poblaciones analizadas para el fragmento de 153pb (ver Tabla R-46, Pag. 337); sin embargo, es bastante similar al estimado para la población Galuresa (0.0626) y muy distinto al obtenido para la población Hadza (0.449), una población que muestra una distribución típica de las poblaciones en equilibrio.

Tiempo transcurrido desde la expansión

El modelo de expansión súbita, permite determinar el tiempo desde que ocurrió la hipotética expansión. Para este cálculo es necesaria la estimación del promedio de la diferencias por parejas y su varianza (ver Material y Métodos). En la Tabla R-47 se muestran los promedios calculados para las poblaciones analizadas.

Población	N	Tamaño del segmento (pb)							
		391		360		302		153	
		<i>M</i>	<i>V</i>	<i>M</i>	<i>V</i>	<i>M</i>	<i>V</i>	<i>M</i>	<i>V</i>
GALI	92	-	-	3.134	4.622	3.090	4.599	1.747	2.211
VAS	45	-	-	3.236	3.995	3.236	3.995	1.978	1.603
P.VELLA	24	-	-	-	-	-	-	1.978	1.818
HAD	17	-	-	3.662	9.781	3.662	9.781	3.147	7.460
FIN	47	-	-	4.008	3.419	3.966	3.233	2.474	2.118
SAR	69	-	-	4.223	6.555	4.136	6.592	2.153	2.206
BRI	100	4.449	5.989	4.390	5.745	4.334	5.701	2.648	2.911
GALU	27	-	-	4.422	4.593	4.199	4.909	2.011	1.520
CAT	25	5.200	4.308	4.960	3.965	4.647	3.942	2.807	2.304
ARG	85	-	-	-	-	4.822	5.469	2.950	2.594
TOS	48	-	-	5.025	5.600	4.983	5.587	3.031	2.903
TRK	45	-	-	5.378	5.109	5.246	5.019	3.257	2.747
ASI	22	-	-	8.610	6.726	8.429	6.829	5.061	3.901

Tabla R-47 Diferencia media por parejas (*M*) y su varianza (*V*) en segmentos de distintos tamaños. El valor de _ se iguala a la media si la varianza es menor que la media (Rogers, 1995).