

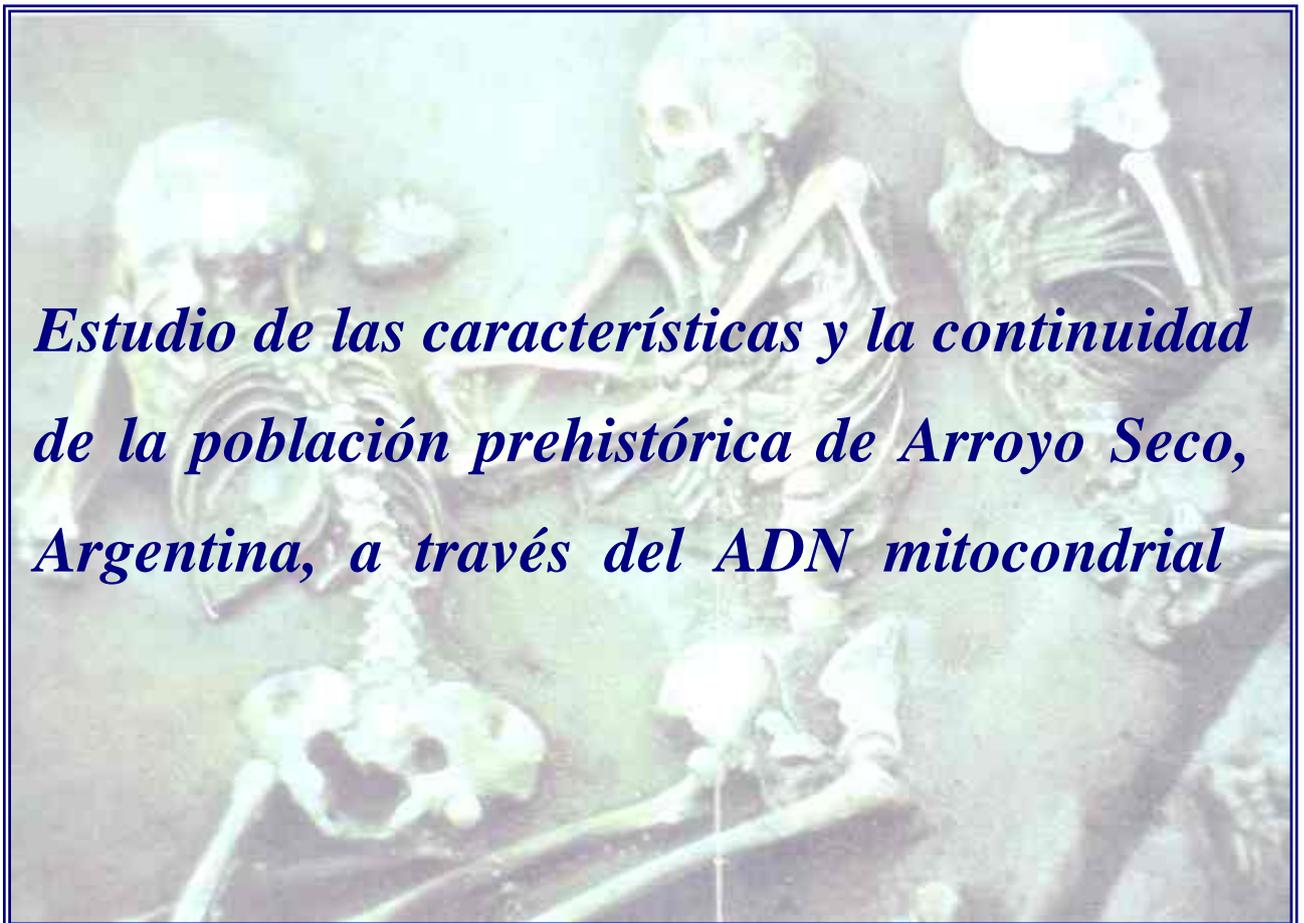
Programa de Desarrollo de las Ciencias Básicas (PEDECIBA)

Área Biología

Subárea Genética

Tesis para la obtención del grado de Magister

Lic. Gonzalo Figueiro



***Estudio de las características y la continuidad
de la población prehistórica de Arroyo Seco,
Argentina, a través del ADN mitocondrial***

Orientadora: *Dra. Mónica Sans*

Tribunal: *Dra. Graciela García*

Dra. Susana González

Dr. Peter Sprechmann

Montevideo, 2006

Índice

Resumen	7
Agradecimientos	9
1. Introducción	11
1.1. Sobre el origen de las poblaciones indígenas americanas	11
1.2. Los análisis del cromosoma Y y el poblamiento de América.....	15
1.3. El ADN mitocondrial en el estudio de las poblaciones americanas	16
1.4. ADN antiguo: Aspectos técnicos.....	23
1.4.1. Sobre la estabilidad y preservación del ADN	23
1.4.2. Extracción de ADN de muestras antiguas.....	26
1.4.2.1. Breve historia	26
1.4.2.2. Protocolos de extracción y criterios de autenticidad	29
1.5. El análisis de ADN mitocondrial antiguo en América.....	32
1.6. La región pampeana: Poblamiento y prehistoria	36
1.7. El sitio Arroyo Seco 2.....	39
2. Objetivos e hipótesis	45
2.1. Objetivos.....	45
2.2. Hipótesis	45
3. Materiales y métodos	46
3.1. Muestra	46
3.2. Métodos	49
3.2.1. Extracción de ADN	49
3.2.1.1. Prevención de contaminación.....	49
3.2.1.2. Protocolo de extracción.....	50
3.2.2. Análisis del ADN	50
3.2.2.1. Procedimiento general de PCR.....	50
3.2.2.2. Control de éxito de las extracciones	52
3.2.2.3. RFLP-PCR y análisis de secuencias.....	52

3.2.3. Análisis de variabilidad y afinidad con otras poblaciones	54
3.2.3.1. AMOVA.....	58
2.2.3.2. Análisis de distancias	61
3.2.3.2.1. Elaboración y <i>bootstrap</i> de los dendrogramas.....	62
3.2.3.2.2. Análisis de coordenadas principales	64
4. Resultados	66
4.1. Extracción de ADN.....	66
4.2. Análisis del ADN.....	66
4.2.1. Determinación de haplogrupo de los individuos.....	66
4.2.2. Amplificación de los controles de extracción	70
4.2.3. Contaminación de los extractos.....	71
4.3. Análisis de variabilidad y afinidad con otras poblaciones.....	74
4.3.1. AMOVA.....	74
4.3.2. Análisis de distancias	76
4.3.2.1. Poblaciones modernas	76
4.3.2.1.1. D_A de Nei	76
4.3.2.1.2. D_R de Rogers.....	79
4.3.2.2. Poblaciones antiguas	83
4.3.2.2.1. D_A de Nei	83
4.3.2.2.2. D_R de Rogers.....	85
5. Discusión	89
5.1. Extracción y análisis de ADN.....	89
5.1.1. Tasa de éxito de las extracciones	89
5.2. Análisis de variabilidad y afinidad con otras poblaciones.....	90
5.2.1. Poblaciones modernas.....	90
5.2.2. Poblaciones antiguas	96
6. Conclusiones y perspectivas	101
7. Referencias citadas.....	104
Anexos	116
I – Protocolo de extracción de ADN.....	116
II – Determinación de haplogrupo de los individuos de la muestra de Arroyo Seco 2	118
III – Procedimientos de remuestreo implementados con el paquete R 2.3.1	122

Índice de tablas

1 – Frecuencias compartidas entre Siberia y América para diversos marcadores clásicos	13
2 – Patrones de restricción de los cinco haplogrupos principales de las poblaciones americanas	19
3 – Polimorfismos de la región hipervariable I correspondiente a los cinco haplogrupos principales.....	19
4 – Cantidad de migraciones y tiempos de divergencia propuestos para los cuatro haplogrupos mayoritarios	20
5 – Localidades, tamaños muestrales y haplogrupos registrados en muestras precolombinas o de poblaciones americanas extintas	22
6 – Vida media y tiempo de degradación de una molécula de ADN	25
7 – Factores favorables para la preservación de ADN antiguo en materiales arqueológicos.....	26
8 – Aplicaciones del ADN antiguo en estudios antropológicos	29
9 – Criterios para garantizar la autenticidad del ADN antiguo	31
10 – Dataciones radiocarbónicas sobre restos humanos de Arroyo Seco 2	46
11 – Entierros e individuos recuperados en el sitio Arroyo Seco 2.....	47-48
12 – Cebadores utilizados para la amplificación de segmentos informativos del ADNmt	53
13 – Muestras modernas utilizadas en este trabajo	55
14 – Muestras antiguas utilizadas en este trabajo.....	56
15a – Agrupaciones de poblaciones modernas para el AMOVA con la muestra de Arroyo Seco 2.....	60
15b – Agrupaciones de poblaciones antiguas para el AMOVA con la muestra de Arroyo Seco 2.....	60
16 – Resumen de los resultados de extracción y diagnóstico de la muestra de Arroyo Seco 2.....	68
17 – Cantidad de experimentos realizados para lograr la primera amplificación de las muestras con los marcadores de RFLP empleados	69
18 – Control de contaminación de los controles de extracción.....	70
19 – Componentes, porcentajes de varianza e índices Φ resultantes del AMOVA con muestras modernas.....	75
20 – Componentes, porcentajes de varianza e índices Φ resultantes del AMOVA con muestras antiguas.....	75
21 – Tasa de éxito en diversos trabajos sobre muestras antiguas americanas.....	90
22 – Grupos indígenas censados en el Chaco y la Pampa-Patagonia en 1977	94

Índice de figuras

1 – Mapa del genoma mitocondrial humano	16
2 – Haplogrupos registrados en muestras precolombinas en América.....	21
3 – Sitios de daño al ADN.....	23
4 – Puntas de proyectil “cola de pescado” recuperadas en Cueva Fell, Chile.....	38
5 – Ubicación del sitio Arroyo Seco 2 y otros sitios tempranos de la región pampeana	40
6 – Arroyo Seco 2: Ubicación esquemática de los hallazgos.....	42
7 – Arroyo Seco 2: Entierro de un individuo infantil con ajuar de colmillos de cánido	42
8 – Entierro n° 2 del sitio Arroyo Seco 2	43
9 – Indumentaria utilizada en el interior del laboratorio de ADN antiguo.....	50
10a – Ubicación de las poblaciones modernas utilizadas para el análisis comparativo	57
10b – Ubicación de los sitios de origen de las muestras antiguas utilizadas para el análisis comparativo	57
11 – Gel de poliacrilamida de productos de digestión	67
12 – Secuencias de AS05 y AS20	73
13 – Árbol <i>neighbor-joining</i> sobre distancias D_A de Arroyo Seco 2 y poblaciones modernas (<i>bootstrap</i> de individuos).....	77
14 – Árbol <i>neighbor-joining</i> sobre distancias D_A de Arroyo Seco 2 y poblaciones modernas (<i>bootstrap</i> de “loci”).....	78
15 – Análisis de coordenadas principales de la matriz de distancias D_A entre las poblaciones de Arroyo Seco 2 y modernas	79
16 – Árbol <i>neighbor-joining</i> sobre distancias D_R de Arroyo Seco 2 y poblaciones modernas (<i>bootstrap</i> de individuos).....	80
17 – Árbol <i>neighbor-joining</i> sobre distancias D_R de Arroyo Seco 2 y poblaciones modernas (<i>bootstrap</i> de “loci”).....	81
18 – Análisis de coordenadas principales de la matriz de distancias D_R entre las poblaciones de Arroyo Seco 2 y modernas	81
19 – Árbol <i>neighbor-joining</i> sobre distancias D_A de Arroyo Seco 2 y poblaciones antiguas (<i>bootstrap</i> de individuos).....	83
20 – Árbol <i>neighbor-joining</i> sobre distancias D_A de Arroyo Seco 2 y poblaciones antiguas (<i>bootstrap</i> de “loci”).....	84
21 – Análisis de coordenadas principales de la matriz de distancias D_A entre las poblaciones de Arroyo Seco 2 y antiguas	85
22 – Árbol <i>neighbor-joining</i> sobre distancias D_R de Arroyo Seco 2 y poblaciones antiguas (<i>bootstrap</i> de individuos).....	86

23 – Árbol *neighbor-joining* sobre distancias D_R de Arroyo Seco 2 y poblaciones antiguas (*bootstrap* de “*loci*”) 86

24 – Análisis de coordenadas principales de la matriz de distancias D_R entre las poblaciones de Arroyo Seco 2 y antiguas 87

Resumen

El análisis de ADN proveniente de restos antiguos ha agregado una dimensión diacrónica a los análisis genéticos, permitiendo en el caso de restos humanos ahondar en aspectos tales como relaciones de parentesco dentro de sitios arqueológicos, movimientos poblacionales y relaciones de ascendencia entre grupos antiguos y grupos actuales, mientras que el análisis de ADN de restos antiguos no humanos ha contribuido a cuestiones de reconstrucción paleoambiental y relaciones filogenéticas, así como a esclarecer la domesticación de varias especies de plantas y animales.

El presente trabajo da cuenta de la extracción y análisis de ADN mitocondrial de los restos de 23 individuos recuperados en el sitio Arroyo Seco 2, localizado en la provincia de Buenos Aires, Argentina. Los restos han sido asignados a cuatro eventos de inhumación, los primeros tres acotados por fechados radiocarbónicos entre 7800 y 6300 años antes del presente, y el cuarto asignado por asociación de elementos arqueológicos al Holoceno tardío. Las extracciones se realizaron en instalaciones exclusivas y con instrumental dedicado, tomando una serie de precauciones para evitar la contaminación de las muestras con ADN moderno. Se extrajo ADN a partir de hueso y diente, empleando un protocolo de fenol/cloroformo – proteinasa K con suspensión de sílice; a partir del extracto fueron amplificados cuatro segmentos polimórficos portando los sitios diagnósticos de los haplogrupos indígenas A, B, C y D. Un segmento de la región hipervariable I fue secuenciado en los extractos negativos para los cuatro marcadores, a fin de confirmar o descartar su contaminación con ADN exógeno.

De las extracciones, 15 rindieron ADN; sin embargo se pudo establecer el haplogrupo de 8 muestras, en tanto que 4 extracciones fueron demostrada y 2 presuntamente contaminadas por ADN del investigador. Por último, en una muestra no se ha podido determinar el haplogrupo y no se ha podido confirmar contaminación, por lo que permanece a la fecha en una situación indeterminada. Las frecuencias de haplogrupos fundadores indígenas A (n = 0), B (n = 3), C (n = 4) y D (n = 1) de la muestra final fueron comparadas con las frecuencias de poblaciones antiguas y modernas de Sudamérica. Los análisis de distancias indican una relación de semejanza y posible ascendencia con las poblaciones modernas del sur de Chile, la Patagonia y el Chaco, planteándose la posibilidad de que Arroyo Seco 2 represente un tronco poblacional común a estas poblaciones. Las relaciones de Arroyo Seco 2 con poblaciones antiguas es más ambigua, agrupándose alternativamente con el extremo norte

—

y sur del subcontinente sudamericano, si bien debe tenerse en cuenta que la cronología de Arroyo Seco 2 supera en antigüedad a cualquiera de las demás muestras americanas.

La posibilidad de extraer y analizar ADN perteneciente a una población del Holoceno temprano, con escasos antecedentes tanto a nivel de América como del resto del mundo, es una fuente potencial de respuestas a interrogantes relacionadas con el poblamiento de todo el Cono Sur sudamericano y, por ende, también de nuestro territorio.

Agradecimientos

Este trabajo no habría sido realizado sin el acicate, la paciencia y el apoyo de muchas personas, pertenecientes a varias instituciones, a las que mencionaré procurando no olvidar a nadie:

- Bernardo “*ya está, ponete a escribir*” Bertoni, Mónica Cappetta y demás integrantes del Departamento de Genética de la Facultad de Medicina (UdelaR), quienes además de permitirme el uso de sus facilidades en el Laboratorio de Oncología Básica y Biología Molecular (LOBBM) me enseñaron desde los cimientos diversas técnicas y me aclararon infinitas dudas.
- Gustavo Politis, Clara Scabuzzo y el resto del equipo del INCUAPA de la Facultad de Ciencias Sociales (Universidad Nacional del Centro de la Provincia de Buenos Aires, sede Olavarría) quienes me facilitaron las muestras cuyo análisis se presenta y los datos de sexo, edad, fechados y contexto de las mismas. Espero que este trabajo haga honor al tesoro invaluable que representa el sitio Arroyo Seco 2 tanto por los restos que conserva como por el exquisito cuidado en su recuperación.
- Los integrantes de la Sección Genética Evolutiva de la Facultad de Ciencias de la UdelaR, con especial énfasis en mis camaradas de armas Lucía Calleros, Mariana Cosse y María José Ferreiro, que toleraron ciertas licencias de mi parte en épocas de fuerte carga docente y en especial durante la redacción de esta Tesis.
- Toda la gente del ámbito académico y extraacadémico que creyó en mí más que yo mismo y me dio un apoyo más allá de lo cognitivo o financiero, sobre todo en épocas en que estuve por renunciar a mucho más que este trabajo: Mi madre, Cristina Lastreto; Patricia Amorín, quien supo acompañarme un largo trecho hasta que nuestros caminos se abrieron; Fernanda Bértola, Victoria Bonanata, Carlos Bordolli, Alessandro Maradei y la gente que pasó – y aún pasa – por la Casa Azul; Pablo Alí y Gastón Bustillo, quienes literalmente me vieron caer; Isabel Barreto, Silvia Borges, Ana Egaña, Soledad dos Santos y el resto de la gente del Instituto de Ciencias Antropológicas de la Facultad de Humanidades y Ciencias de la Educación (UdelaR). Y mil más que saben quiénes son y que perdonarán mi omisión.
- Mi orientadora, maestra y mentora, la Dra. Mónica Sans, quien desde hace casi ocho años viene depositando en mí su confianza incondicional. A ella le corresponde el mérito de

haber leído y releído varios trabajos de mi autoría – de los que éste seguramente no será el último –, puliéndolos, corrigiéndolos y bajándolos a tierra hasta el hartazgo. Asimismo, corresponde mencionar la retroalimentación teórica brindada por la Lic. Isabel Barreto y el Dr. Pedro C. Hidalgo Si bien agradezco a mis colegas por su ayuda, soy el único responsable de las omisiones o inexactitudes que pueda contener este trabajo.

Este trabajo está dedicado a la memoria de mi padre, Luis Ramón Figueiró.

1. Introducción

1.1. Sobre el origen de las poblaciones indígenas americanas

La afirmación de Antonio de Ulloa en el siglo XVIII, de que “*visto un indio de cualquier región se puede decir que se han visto todos en cuanto al color y contextura*” (en Comas 1974: 35) recibió apoyo de antropólogos como Samuel G. Morton, Arthur Keith y Ales Hrdlicka entre otros, siendo Hrdlicka el representante más influyente de esta corriente. Hrdlicka, quien planteaba la existencia hipotética de un “indio americano medio” sobre la base de una serie de características somáticas, sostenía la idea de homogeneidad de los ancestros de los indígenas americanos. Éstos tendrían en su totalidad origen mongoloide y habrían llegado a América desde Siberia, por el estrecho de Bering, hacía como máximo 5.000 años. Esta edad se mantuvo en pie hasta el descubrimiento de artefactos en asociación con megafauna en Folsom, Nuevo México en 1927. Gradualmente se extendió la cronología de la presencia del ser humano en América hasta 12.000 años antes del presente (A.P.).

En el campo de quienes, al contrario que Hrdlicka, plantearon la heterogeneidad de las poblaciones americanas, se puede distinguir dos vertientes: Por un lado estaban quienes estimaban que la variabilidad somática de las poblaciones americanas se debía a influencias medioambientales posteriores al ingreso de los antepasados de éstas en el continente. Por el otro lado estaban aquellos que, dentro del marco teórico del difusionismo histórico-cultural, afirmaban la inmigración de tipos humanos diversos, cada uno de los cuales representaría una de las “razas” o “variedades” existentes al momento de la conquista (lo cual implicaba que éstas no habían sufrido modificaciones en América). Desde esta perspectiva, la variabilidad encontrada por el europeo en América es inseparable de las circunstancias que condujeron al poblamiento del continente: Cada tipo humano habría ingresado al continente proveniente de una región con características físicas y culturales semejantes a las encontradas en América.

Diversas teorías pueden incluirse en esta perspectiva. Paul Rivet (1943) sostenía que la población indígena americana es el resultado de migraciones provenientes de Asia, por el Estrecho de Bering, y de Oceanía, a través del Océano Pacífico, aportando cuatro tipos raciales. La tesis de Mendes Corrêa, por su parte, afirmaba que el elemento que Rivet luego llamaría Australo-Tasmanoide habría llegado a América utilizando la vía antártica. Sin rechazar la tesis de Mendes Corrêa, Montandon plantea la Isla de Pascua como etapa

intermedia de la vía transpacífica (Comas 1974). Imbelloni (1938), en una revisión de planteos anteriores, plantea diez variantes, agrupadas en siete oleadas poblacionales.

Las teorías presentadas se basaban fundamentalmente en datos morfológicos, utilizando un criterio tipológico según el cual todas las poblaciones se pueden clasificar según su ajuste a “tipos” ideales asimilables en muchos casos a “razas”. Este criterio fue cayendo en desuso junto con los estudios morfológicos, a medida en que se demostraba la plasticidad de la morfología de acuerdo al ambiente. En América del Norte, por ejemplo, Franz Boas sería uno de los primeros y más influyentes detractores de los estudios raciológicos, demostrando la baja heredabilidad de diversas variables cefalométricas en un estudio de inmigrantes europeos y sus descendientes; estos estudios han sido objeto de debate en trabajos recientes (*cf.* Sparks y Jantz 2002; Gravlee et al. 2003a, 2003b).

Sin desconocer la variabilidad biológica de los indios americanos, datos morfológicos y genéticos apuntan al origen de las poblaciones americanas en uno o unos pocos grupos reducidos de individuos provenientes de Asia, que habrían ingresado por el Estrecho de Bering. A nivel morfológico, asiáticos y amerindios comparten una serie de rasgos visibles comunes, como el cabello oscuro, la relativa falta de vello facial y corporal, arcos cigomáticos anchos (pómulos salientes) y pliegue epicántico (ojos rasgados); a nivel craneal, los trabajos de análisis multivariado de, entre otros, W.W. Howells (craneometría) y H. Ishida (rasgos discontinuos) agrupan a siberianos e indígenas americanos. La alta frecuencia de rasgos dentarios como incisivos en forma de pala y primeros molares inferiores con tres raíces es también compartida entre poblaciones asiáticas y amerindias (Crawford 1998).

A nivel genético, los amerindios se asemejan a las poblaciones asiáticas, y en especial a las siberianas, en frecuencias de varios grupos sanguíneos, proteínas séricas, enzimas eritrocitarias y polimorfismos de ADN (véase tabla 1 para un resumen de las frecuencias de marcadores genéticos clásicos compartidas por las poblaciones americanas y siberianas).

Sistema	Alelo(s)	Frecuencia en Siberia	Frecuencia en América
<i>Grupos sanguíneos</i>			
RH	RH*R1; RH*R2	>50%	>50%
Diego	DI*A	2-6%	2-6% *
Duffy	FY*A	6-95%	5-92%
<i>Proteínas séricas</i>			
Transferrinas	TF*B	Ausente	Presente
	TF*B0-1	Presente	Presente
	TF*D	Ausente	Presente
	TF*DCHI	Presente	Presente**
Componente grupo-específico	GC*1S	Alta	Alta
	GC*2	Baja	Baja
Factor properdina B	BF*S	92%***	80-100%
Seudocolinesterasa (CHE2)	C5+	0-13%	2-14%
<i>Proteínas eritrocitarias</i>			
Fosfoglucomutasa	PGM*1	54-92%	59-98%
Uridin monofosfato kinasa	UMPK*3	13%	9-20%
Esterasa D	ESD*2	13-34%	3-28%

Tabla 1: Frecuencias compartidas entre Siberia y América para diversos marcadores genéticos clásicos (Crawford 1998).

* Excepcionalmente se encuentran poblaciones con frecuencias de alrededor de 20%.

** TF*DCHI está ausente en prácticamente todas las poblaciones del mundo estudiadas por este *locus*.

*** Una sola población estudiada.

En cuanto a la interrogante de la antigüedad del hombre en América, desde filas de la arqueología norteamericana se afirma que el ingreso del hombre moderno en el continente americano a través del Estrecho de Bering no pudo ser anterior a un época en la que el puente de tierra de Bering estuviera al descubierto al tiempo que el casquete glacial se hubiera retirado lo suficiente como para permitir el paso a los inmigrantes, o sea hace como mucho 12.000 años. Sin embargo varios sitios arqueológicos aportan evidencias de una posible entrada anterior del hombre en América. Entre los sitios más mencionados aparecen Meadowcroft, EE.UU. (19.000-13.000 A.P.), Wilson Butte Cave, EE.UU. (14.500 A.P.), Fort Rock Cave, EE.UU. (13.200 A.P.), Valsequillo, México (22.000 A.P.), Tlapacoya, México (22.000 A.P.), Pikimachay, Perú (21.000-13.000 A.P.), Toca do Boqueirão da Pedra Furada, Brasil (34.000-25.500 A.P.) y Monte Verde, Chile (13.000 A.P.; *cf.* Fiedel 1996 para una revisión completa del debate sobre el poblamiento de América desde el punto de vista arqueológico – y una visión defensora de una fecha de entrada reciente). El fechado de 13.000 años A.P. de Monte Verde, ubicado al sur de Chile (Dillehay 1997; el fechado de 13.000 A.P. ha sido aceptado por la comunidad arqueológica – Meltzer 1997; no obstante otro fechado de 33.000 años A.P. es aún objeto de debate), significa un desafío importante para los defensores

de un poblamiento tardío, en la medida en que los habitantes del sitio tendrían que haber entrado al continente con anterioridad a esa fecha. Considerando que el casquete glacial aún bloqueaba la entrada continental a Norteamérica, hay que postular una entrada por vía costera (Dixon 2002). Sin embargo, sobre la base de simulaciones, Surovell (2003) afirma que para que Monte Verde anteceda en 1500 años a los fechados a los sitios continentales norteamericanos aceptados, la población se debe haber movido a lo largo de la costa absteniéndose durante un milenio y medio de migrar tierra adentro. Eso habría sido posible únicamente si las tasas de retorno del ambiente costero superaran ampliamente a las tasas de retorno continentales, cosa considerada poco plausible sobre bases paleoecológicas (Surovell 2003).

En cuanto al carácter de la inmigración, específicamente, la cantidad de contingentes que habrían ingresado, así como el tamaño de los mismos, el trabajo seminal de Greenberg et al. (1986) plantea tres oleadas de poblamiento del continente americano, basándose en datos lingüísticos, dentarios y genéticos. El trabajo plantea fechas estimadas de ingreso al continente sobre bases lingüísticas; las entradas de las oleadas, nombradas en términos lingüísticos como la amerindia, na-dene y esquimal, fueron fechadas entre 9000 y 11000 años, 4500 o 9000 años y 2500 años, respectivamente. La teoría propone también que una única oleada llegó a Sudamérica, lo que reduce la variabilidad de los indígenas sudamericanos a aquélla que pudiera obtenerse *a posteriori* del evento de poblamiento.

Este trabajo fue fuertemente criticado, especialmente en lo referente al fechado glotocronológico (basado en tasas de divergencia entre lenguas y grupos de lenguas) y ciertos criterios en el orden lingüístico que habrían sobresimplificado el panorama interno del grupo llamado “amerindio” (Crawford 1998; Nótese que el trabajo de Greenberg et al. [1986] otorga el nombre de amerindio a un grupo lingüístico de los indígenas americanos, en tanto que en los trabajos anteriormente citados se hace referencia a “amerindio” como indígena americano en sentido amplio). No obstante, en el campo genético el modelo de oleadas se siguió empleando como hipótesis, mientras los estudios se empezaban a aproximar al tema desde marcadores uniparentales, específicamente, polimorfismos de ADN del cromosoma Y, de transmisión paterna, y polimorfismos en el ADN mitocondrial, de transmisión materna.

1.2. Los análisis del cromosoma Y y el poblamiento de América

El cromosoma Y humano se compone de tres regiones distintas: Dos regiones pseudoautosómicas, subteloméricas, que recombinan con el cromosoma X, y comprenden 3 megabases (MB) de un total de 63 MB. El restante 95% del cromosoma Y corresponde a la región no recombinante, que se transmite únicamente a través de linajes masculinos (Lahn et al. 2001).

Los primeros estudios exitosos de polimorfismos a nivel de cromosoma Y se centraron en análisis de secuencias cortas repetidas en tándem (STR), llamadas también microsatélites. Los microsatélites son series en el orden de las decenas o centenares de motivos de di- tri- o tetranucleótidos dispersos en millares de copias por todo el genoma. Los STR son altamente variables, con alelos consistentes en diferencias en la cantidad de repetidos y tipificables con relativa facilidad mediante la reacción en cadena de la polimerasa (PCR). Los análisis evolutivos y poblacionales basados en microsatélites son un poco más complejos que para polimorfismos de secuencia, en virtud de los variados modelos de mutación existentes en la actualidad para microsatélites, que determinan a su vez diferentes índices de distancia evolutiva (Bennett et al. 1998).

El primer microsatélite polimórfico para el cromosoma Y, llamado actualmente DYS19, mostró ser extremadamente informativo al estudiarse en conjunto con polimorfismos del sistema alfoide, cuyas distintas combinaciones (haplotipos) mostraban una fuerte asociación con poblaciones de distintas regiones geográficas. En el caso de América, el haplotipo IIA, combinación del tipo alfoide II con el alelo A de DYS19, mostró ser mayoritario en poblaciones americanas, sugiriéndose que sería el haplotipo llevado por la mayor parte de los primeros pobladores americanos (Pena et al. 1995). Poco tiempo después, el descubrimiento de un polimorfismo de nucleótido simple (SNP) en el locus DYS199, conocido actualmente como linaje Q*3 (caracterizado por la mutación Q-M3*), cuya transición C→T es mayoritaria en amerindios, contribuyó a fortalecer la idea de un haplogrupo fundador mayor para los indígenas americanos (Underhill et al. 1996).

Ulteriores trabajos empleando microsatélites han mostrado considerable variabilidad en poblaciones americanas. Asimismo, el aumento en los tamaños muestrales y los marcadores empleados (microsatélites y SNPs) llevó al descubrimiento en América de dos haplogrupos adicionales: P*M45, con el nucleótido ancestral C en el marcador M3 (Ruiz-Linares et al. 1999) y otro caracterizado por la mutación S4Y (Karafet et al. 1999). Con ello y

con el análisis de las frecuencias de haplogrupos en poblaciones asiáticas, se desató un debate, análogo al suscitado a partir de datos de ADNmt, acerca de la cantidad, el origen y la cronología de las oleadas poblacionales que habrían alcanzado el continente americano, debate aún no resuelto (*cf.* Lell et al. 2002; Tarazona-Santos y Santos, 2002; Zegura et al. 2004).

1.3. El ADN mitocondrial en el estudio de las poblaciones americanas

Las mitocondrias y los cloroplastos constituyen sistemas genéticos secundarios de las células eucariotas, y presentan el fenómeno de herencia citoplasmática, caracterizada por proporciones de segregación no mendelianas. La herencia de los genes mitocondriales en mamíferos es materna, en virtud que las mitocondrias del embrión son proporcionadas por el óvulo. Los genomas mitocondriales conocidos codifican proteínas mitocondriales esenciales, y ARN ribosomales y de transferencia suficientes para sintetizar dichas proteínas. No

obstante, las mitocondrias no son autónomas en la medida en que muchas proteínas esenciales para su funcionamiento son codificadas en el núcleo y sintetizadas en el citoplasma (Lodish et al. 2000).

El genoma mitocondrial humano es una molécula circular de 16.568 pares de bases (el número de bases corresponde a la revisión hecha por Andrews et al. 1999 de la secuencia original publicada por Anderson et al. 1981; la numeración de las posiciones del genoma se siguen sin

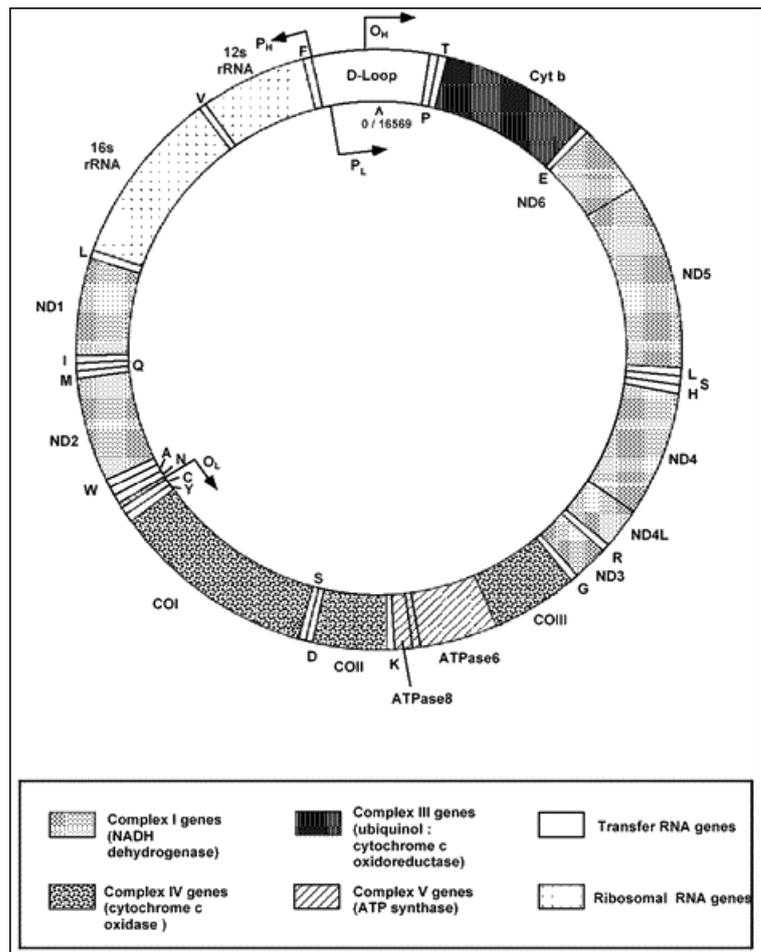


Figura 1: Mapa del genoma mitocondrial humano. Tomado de MITOMAP (Wallace v Lott 2003)

embargo numerando sobre las 16.569 bases originales). Codifica 22 ARN de transferencia, sus propios ARN ribosomales, y 13 polipéptidos de más de 50 aminoácidos (Figura 1). A semejanza de la mayor parte de los genomas mitocondriales de vertebrados, es extremadamente compacto, con regiones intergénicas reducidas y ausencia de intrones (Lodish et al. 2000).

A pesar de observarse un alto grado de conservación de los genes mitocondriales de vertebrados, se ha constatado un rápido ritmo evolutivo del ADN mitocondrial (ADNmt), especialmente en lo que respecta a la región control, una zona de aproximadamente 1000 pb relacionada con el origen de replicación. Dentro de ésta, las llamadas regiones hipervariables (HVR) I y II (ubicadas en las posiciones 16024-16383 y 57-372, respectivamente; las posiciones siguen la secuencia de referencia – Anderson et al. 1981) presentan una alta tasa de mutación. En el ser humano, las tasas de mutación promedio estimadas para las regiones hipervariables I y II son de 11,7 mutaciones por sitio y generación, con un intervalo de confianza del 95% de 2,1–31,6 (Heyer et al. 2001). Esta alta variabilidad intersitio se debería fundamentalmente a la presencia de sitios “calientes” (*mutational hotspots*) y no a eventos recombinatorios (Macaulay et al. 1999; Stoneking 2000). Sin embargo, los argumentos a favor de la existencia de recombinación (Eyre-Walker et al. 1999; Hagelberg et al. 1999) no pueden aún ser descartados por razones de muestreo (Wiuf 2001). Sobre la base del conocimiento disponible a la fecha, no obstante, se puede asumir que toda variación en el genoma mitocondrial humano proviene de eventos de mutación. Por lo tanto, es posible calcular diversidad intrapoblacional y distancias interpoblacionales aún en casos de poblaciones que han divergido en tiempos relativamente recientes.

Ha de tenerse en cuenta sin embargo que las características que determinan la utilidad del ADNmt son también sus principales desventajas: La ausencia de recombinación hace que el genoma mitocondrial se herede como un solo *locus*, y en virtud de su herencia uniparental es especialmente sensible a fenómenos de deriva (Bennett et al. 1998).

Los primeros estudios de variabilidad del ADNmt fueron basados en polimorfismos de sitios de restricción (Brown 1980; Johnson et al. 1983), que mostraron una serie de patrones con aparente especificidad poblacional. Posteriormente, la información generada en cuanto a patrones de restricción se sumó hacia fines de la década de 1980 al debate - aún no resuelto - acerca de la época y el lugar de origen del *Homo sapiens* (Cann et al. 1987). Más recientemente, la comparación de secuencias de las regiones hipervariables agregó argumentos a la discusión en torno al destino evolutivo de las poblaciones “arcaicas” europeas

de *Homo neanderthalensis* (Krings et al. 1997, 2000; Ovchinnikov et al. 2000; Schmitz et al. 2002; Lalueza-Fox et al. 2005).

Aportando al debate en torno a las fechas y el modo de poblamiento del Nuevo Mundo, los análisis tempranos de polimorfismos de fragmentos de restricción (RFLP) de muestras de ADNmt de distintos continentes mostraron que los patrones de restricción para una endonucleasa en particular (llamados “morfos”) se asociaban con el origen étnico del individuo portador (Denaro et al. 1981). A su vez, Johnson et al. (1983) observaron que los patrones de restricción polimórficos generados por seis enzimas se podían representar como ramificaciones basadas en mutaciones únicas a partir de dos tipos centrales, especialmente frecuentes en Europa y Asia. Posteriormente, el trabajo de Wallace et al. (1985), sobre una muestra de 81 nativos norteamericanos, mostró que los ADNmt amerindios presentaban tipos asiáticos, pero en frecuencias radical y significativamente diferentes, lo cual era explicado por un efecto cuello de botella que apoyaría el escenario de que mujeres asiáticas cruzaron el puente de tierra de Bering y se dispersaron en el continente en pequeñas bandas de cazadores-recolectores.

El trabajo de Schurr et al. (1990) sobre ADNmt de amerindios (que constituirían la primera oleada postulada en el trabajo de Greenberg et al. 1986) arrojó la presencia de diez haplotipos exclusivos, de los cuales cuatro fueron designados por su alta frecuencia como “haplogrupos fundadores”, todos ellos de origen asiático: una ganancia de sitio de *HaeIII*, una delección entre la subunidad II de la citocromo oxidasa c (COII) y el gen del tRNA(Lys), una pérdida de sitio de *HincII* y una pérdida de sitio de *AluI* (Tabla 2). Estos datos sugerían que el ADN mitocondrial amerindio derivaba de por lo menos cuatro linajes maternos, y que se habrían acumulado variantes en su dispersión por el continente. A los cuatro haplogrupos mencionados, llamados actualmente A, B, C y D (Torroni et al. 1992), se agrega el haplogrupo X, derivado de una variante actualmente presente en poblaciones europeas (Easton et al. 1996; Brown et al. 1998; Smith et al. 1999), lo cual en un principio dio lugar a un debate en cuanto al carácter “fundador” del haplogrupo (este debate habría sido resuelto al descubrirse la presencia del haplogrupo X en muestras prehistóricas – Ribeiro-dos-Santos et al. 1996; Malhi y Smith 2002) y en cuanto a la posibilidad de una ruta trasatlántica de migración prehistórica desde Europa.

Haplogrupo	Sitio de restricción						
	HaeIII np 663	Delección 9 pb*	HincII np 13259	AluI np 13262	AluI np 5176	AccI np 14465	DdeI np 1715
A	+	-	+	+	+	-	+
B	-	+	+	-	+	-	+
C	-	-	-	+	+	-	+
D	-	-	+	-	-	-	+
X	-	-	+	-	+	+	-

Tabla 2: Patrones de restricción de los cinco haplogrupos principales de las poblaciones americanas. Los grupos A, B, C y D corresponden a los “haplogrupos fundadores” definidos por Schurr et al. (1990) Basado en Kaestle y Smith (2001) y Eshleman et al. (2003).

(*) Delección de 9 pb en la región intergénica COII/tRNA^{Lys}.

Posteriormente, diversos trabajos realizados sobre la región hipervariable del ADN mitocondrial humano demostraron una alta variabilidad intrapoblacional en lo referente a mutaciones de esta región; esta variabilidad se agrupó igualmente en *clusters* (Ward et al. 1991; Horai et al. 1993; Bailliet et al. 1994) que pudieron identificarse posteriormente con los patrones de restricción A, B, C y D definidos anteriormente (Forster et al. 1996; Tabla 3).

Np	Haplogrupo				
	A	B	C	D	X
16111	<u>T</u>	C	C	C	C
16189	T	<u>C</u>	T	T	T
16217	T	<u>C</u>	T	T	T
16223	<u>T</u>	C	<u>T</u>	<u>T</u>	T
16278	C	C	C	C	<u>T</u>
16290	<u>T</u>	C	C	C	C
16298	T	T	<u>C</u>	T	T
16319	<u>A</u>	G	G	G	G
16325	T	T	<u>C</u>	T	T
16327	C	C	<u>T</u>	C	C
16362	<u>C</u>	T	T	<u>C</u>	T
16519	<u>C</u>	T	T	<u>T</u>	T

Tabla 3: Polimorfismos de la región hipervariable I correspondiente a los cinco haplogrupos principales. Los nucleótidos subrayados corresponden a las mutaciones características del haplogrupo. Tomado de Forster et al. (1996).

Si bien existe acuerdo en cuanto a las características del ADNmt de las poblaciones americanas, no lo hay en cuanto al significado de éstas para la forma y la época del poblamiento de América (Tabla 4): Las primeras interpretaciones (Torroni et al. 1992) indicaban que las poblaciones amerindias (representadas por los haplogrupos A, B, C y D) y las poblaciones na-dene (representadas fundamentalmente por el haplogrupo A y en menor medida por el B) representaban distintas migraciones, ocurridas la primera entre 21.000 y 42.000 años A.P. y la segunda entre 5.250 y 10.500 años A.P. (el trabajo de Torroni et al. [1992] no incluyó esquimales, por lo que los autores se abstienen de hacer inferencias sobre

los mismos). Por su parte, Horai et al. (1993) consideraron a cada haplogrupo una migración independiente, situación altamente improbable considerando que incluso una muestra pequeña de cualquier población siberiana oriental contendrá representantes de más de uno de los haplogrupos fundadores (Merriwether et al. 1995; Eshleman et al. 2003). Por añadidura, Merriwether et al. (1995) demostraron que los tres grupos lingüísticos considerados por Greenberg et al. (1986) presentan los cuatro haplogrupos en cuanto se emplean tamaños muestrales grandes, circunstancia que raramente podría repetirse en tres eventos migratorios, por lo que arguyen que las Américas habrían sido pobladas por un solo evento migratorio desde Asia. Esta afirmación es apoyada por Bonatto y Salzano (1997a, 1997b), quienes plantean una fecha de entrada única mayor a 22.000 años, y que incluiría a los antepasados de los indígenas de los tres grupos lingüísticos.

Autores	Nº de Eventos	Fechas de divergencia
Torrioni et al. 1992	2 (Amerindio y Na-dene)	A (Na-dene): 5.250–10.500 A (Amerindio): 17.750–35.500 B (Na-dene): 8.000–16.000 B (Amerindio): 9.250–18.500 C (Amerindio): 25.500–51.000 D (Amerindio): 18.250–36.500
Horai et al. 1993	4 (uno por haplogrupo)	A, B, C, D: 14.000–21.000
Torrioni et al. 1994	2 (un evento “temprano” con haplogrupos A, C y D y uno “tardío” con haplogrupo B)	A (Amerindio): 25.862–34.091 B (Amerindio): 11.724–15.456 C (Amerindio): 33.105–43.636 D (Amerindio): 18.276–24.091
Bonatto y Salzano 1997a	1	A: 34.887–47.142 o 23.949–32.377 B: 36.240–44.569 o 22.972–30.611 C: 34.150–47.210 o 23.443–32.423 D: 38.979–54.576 o 26.759–37.482
Stone y Stoneking 1998	1	A: 25.000–57.000 o 12.000–30.000 B: 16.000–41.000 o 8.000–21.000 C: 13.000–40.000 o 6.000–21.000 D: 19.000–51.000 o 9.000–27.000
Schurr et al. 1999	1	A: 26.969–35.550 B: 13.483–17.773 C: 40.972–54.009 D: 19.483–25.682
Silva et al. 2002	1	A: 16.400–24.600 B: 14.600–21.700 C: 17.300–25.900 D: 19.300–28.300

Tabla 4: Cantidad de migraciones y tiempos de divergencia propuestos para los cuatro haplogrupos mayoritarios (intervalo de confianza del 95%).

Varios estudios, al plantear la fecha de entrada de los primeros pobladores, procuran utilizar aquellos de sus resultados que impliquen el menor conflicto posible con los datos

paleoclimáticos disponibles, si bien prácticamente siempre los estimados sobre ADNmt superan por varios milenios a los estimados arqueológicos generalmente aceptados asumiendo una entrada continental de los pobladores. La única manera de conciliar los datos paleoclimáticos con los genéticos es postular una entrada en América previa a la retirada del casquete glacial, probablemente por vía costera (Schurr y Sherry 2004). Fix (2005) demostró a partir de simulaciones que la variabilidad regional en cuanto a frecuencias de los cuatro haplogrupos fundadores es perfectamente plausible planteando un modelo de colonización costera del continente.

Los estimados de fechas de entrada se basan en la mayor parte de los casos en cálculos de diversidad nucleotídica de las HVR I y II para cada haplogrupo, y asumen que la diversificación de los haplogrupos sucedió tras el evento de inmigración que, al incluir un cuello de botella – tema también objeto de discusión –, habría “puesto a cero” el cronómetro evolutivo. Esta asunción debe hacerse con prudencia ya que, si los primeros pobladores ingresaron con algún grado de diversidad haplotípica por mínima que ésta fuese dentro de los haplogrupos, las edades de ingreso estarán sobrestimadas (Ward et al. 1991; Merriwether et al. 1995). Malhi et al. (2002) sostienen que las mediciones de diversidad nucleotídica de la región hipervariable no aportan datos concluyentes, entre otras cosas por la hipermutabilidad de la misma, pero los análisis de otras regiones del genoma mitocondrial no han, hasta ahora, arrojado mayor luz sobre el caso: Los datos aportados por Silva et al. (2002) de secuencias de 8000 pb de la región codificante apuntan a un solo evento migratorio, en tanto que Bandelt et al. (2003), basándose en secuencias del genoma mitocondrial completo afirman que un poblamiento inicial hace más de 30.000 años debe haber sido sucedido por otros tantos eventos menores, incluso hasta el principio del Holoceno.

Las inferencias mencionadas en cuanto a cantidad y antigüedad de los contingentes que ingresaron al Nuevo Mundo se basan

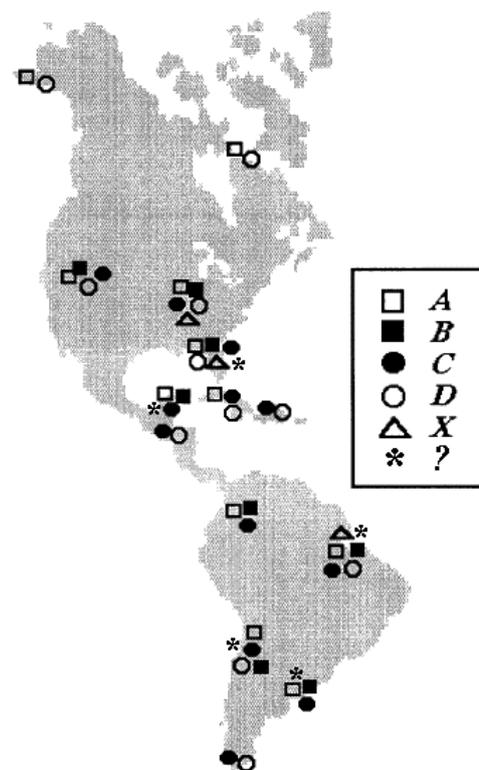


Figura 2: Haplogrupos registrados en muestras precolombinas en América. Modificado de Jones (2003).

enteramente en muestras de ADN mitocondrial de individuos modernos. Sin embargo, en cuanto fue posible analizar ADN de restos antiguos, se analizaron restos de América en busca de testar si las poblaciones pasadas poseían los mismos haplogrupos que los encontrados en las poblaciones actuales o si, por el contrario, hubo haplogrupos en tiempos precolombinos extintos luego de la llegada del europeo. Trabajos tempranos, como el de Pääbo et al. (1988) hallaron mutaciones del ADN mitocondrial que no correspondían ninguno de los cuatro haplogrupos fundadores. Trabajos posteriores (e.g. Hauswirth et al. 1994; Ribeiro-dos-Santos et al. 1996) hallaron individuos que no correspondían con ninguno de esos haplogrupos. En algunos casos, estos individuos resultaron tener el haplogrupo X, que fue definido con detalle posteriormente; en otros, sigue siendo un haplogrupo distinto a los cinco hasta ahora identificados (Tabla 5).

Localidad	Cronología aprox. (años A.P.)	N	A	B	C	D	X	Otros	Referencia
Florida, EE.UU.	8100–7000	14	0%	7%	0%	21%	0%	72%	Hauswirth et al. 1994
Colombia	1500–100	6	50%	17%	33%	0%	0%	0%	Monsalve et al. 1996
Amazonia	4000–500	18	28%	6%	22%	6%	17%	22%	Ribeiro-Dos-Santos et al. 1996
Sudoeste de EE.UU.	9200–350	94	10%	51%	13%	25%	0%	0%	Parr et al. 1996; Carlyle et al. 2000; Kaestle y Smith 2001
Patagonia	5000–100	15	0%	0%	27%	73%	0%	0%	Lalueza et al. 1997
Tierra del Fuego	100	69	0%	0%	49%	51%	0%	0%	Lalueza et al. 1997; García-Bour et al. 2004
Illinois, EE.UU.	650	108	32%	12%	42%	8%	6%	0%	Stone y Stoneking 1998
Yucatán	1350–400	25	84%	4%	8%	0%	0%	4%	González-Oliver et al. 2001
República Dominicana	1300–300	24	0%	0%	75%	25%	0%	0%	Lalueza-Fox et al. 2001
Norte de Chile	3900–500	93	28%	30%	20%	3%	0%	18%	Moraga et al. 2001; Moraga et al. 2005
Cuba	4700–1600	15	7%	0%	60%	33%	0%	0%	Lalueza-Fox et al. 2003
Uruguay	2000–100	9	11%	33%	45%	0%	0%	11%	Bertoni et al. 2004; Sans 2004; Sans y Figueiro 2005; datos sin publicar
Noroeste argentino	1500–1000	16	6%	50%	0%	44%	0%	0%	Dejean et al. 2005

Tabla 5: Localidades, tamaños muestrales y haplogrupos registrados en muestras precolombinas o de poblaciones americanas extintas.

En general, sin embargo, las muestras antiguas analizadas en América presentan los haplogrupos fundadores registrados en poblaciones modernas (Figura 2). La diversidad intra-

haplogrupo suele no obstante ser mayor en las muestras antiguas: Análisis de la región hipervariable I de una muestra del sitio Norris Farms, EE.UU., arrojaron una diversidad a nivel de linajes únicos más alta que la que se halla en poblaciones modernas; esta mayor diversidad puede haber sido mermada por efecto de la reducción poblacional poscontacto (Stone y Stoneking 1998, 1999).

1.4. ADN antiguo: Aspectos técnicos

1.4.1. Sobre la estabilidad y preservación del ADN

De acuerdo a Lindahl (1993a), los mayores daños ocurridos al ADN en la célula viva son daños hidrolíticos y oxidativos (Figura 3). Los daños por hidrólisis atacan el enlace N-glicosídico entre el azúcar y la base, que es relativamente débil en aras de la resistencia del enlace fosfodiéster. Este daño implica fundamentalmente la pérdida de purinas, que son veinte veces más sensibles al daño que las pirimidinas. Tras la pérdida de la base se suele producir un corte en la cadena de azúcar-fosfato (β -eliminación), con lo que se da una ruptura de la cadena de ADN. Otro daño producido por hidrólisis es la deaminación, que afecta a la citosina (convirtiéndola en uracilo) y a su homólogo la 5-metilcitosina (convirtiéndola en timina). este daño no afecta generalmente a las purinas.

El daño oxidativo se basa en el modo en que radicales hidroxilo y peróxido modifican las bases o distorsionan la hélice. La modificación de bases oxidativa ataca en principio a las purinas, consistiendo fundamentalmente en la conversión de la guanina en 8-hidroxiguanina por parte de un radical hidroxilo; esta base aparea mejor con la adenina que con la citosina. El

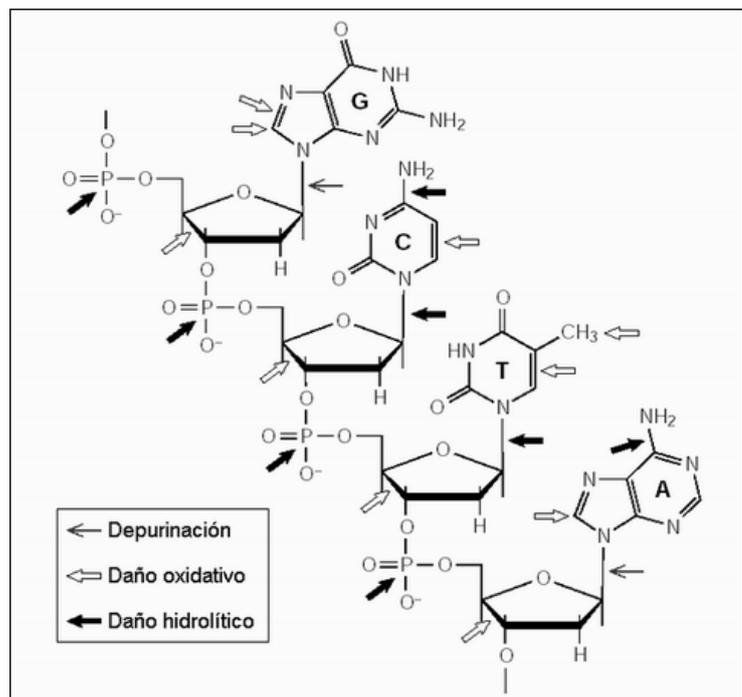


Figura 3: Sitios de daño al ADN. Modificado de Lindahl (1993a).

daño oxidativo a las pirimidinas produce en última instancia la desintegración de las mismas. Por último, las distorsiones de la hélice producidas por el daño oxidativo se debe al establecimiento de enlaces covalentes dobles entre las purinas y la cadena de azúcar-fosfato, y a la formación de dímeros de purinas.

La mayor parte de los daños mencionados se revierten con relativa eficacia en la célula viva por una serie de mecanismos de reparación. No obstante, una vez muerto el organismo, estos daños se acumulan. Se supone que los daños son más intensos en el período inmediatamente siguiente a la muerte, con la liberación de enzimas autolíticas por los lisosomas y la ruptura de la lámina nuclear, que presumiblemente deja susceptibles a la mayor parte del ADN y el ARN al ataque de nucleasas. Por añadidura, la morfología de los ácidos nucleicos extraídos de restos momificados revelan uniones interhebra (*inter-strand cross-linking*; Pääbo 1989). La imposibilidad de unión por parte de agentes intercalantes como el bromuro de etidio al ADN antiguo es consistente con la posibilidad de que la hélice esté fuertemente deformada (Rogan y Salvo 1990).

Se puede esperar que el ADN totalmente hidratado, desprovisto de mecanismos de reparación, se fragmente por completo en fragmentos pequeños (< 10 pb) en unos pocos miles de años. La velocidad de degradación puede reducirse en presencia de sales, y la unión del ADN a la hidroxiapatita reduce la velocidad de desintegración a la mitad. Por ende, en el tejido óseo bien preservado y a bajas temperaturas, se puede esperar una preservación mayor.

Estas consideraciones están directamente relacionadas con la amplificación por reacción en cadena de la polimerasa (PCR). La deaminación oxidativa hace que las citosinas (convertidas en uracilo) sean reconocidas como timinas por la *Taq* polimerasa, y por consiguiente en la extensión se agregará una adenina. Como resultado, se produce en la PCR la sustitución de un par CG por otro AT; un exceso de mutaciones G→A o C→T debe ser observado como producto del daño por deaminación (Pääbo et al. 1990). Estudios recientes atribuyen a este mecanismo la causa principal de modificaciones conducentes a errores de amplificación, en función de la reducción de la frecuencia de errores tras el tratamiento del ADN con Uracilo-N-Glicosilasa (UNG) (Hofreiter et al. 2001). Otro fenómeno resultante en una sustitución CG→AT es el "cambio de molde" o "salto de cebador" (*primer jumping*), basado en la tendencia de la *Taq* polimerasa a agregar residuos A en el extremo 3' de los productos de extensión. Como resultado, la interrupción en la extensión por la fragmentariedad del molde puede dar lugar a que, en el siguiente ciclo, el producto

parcialmente extendido se una a otro molde correspondiente a otro blanco o a un producto de contaminación, dando lugar a una amplificación “recombinante” (Pääbo et al. 1990).

Las modificaciones que llevan a la eliminación de bases o ruptura de la cadena azúcar-fosfato del ADN llevan ulteriormente a un grado de degradación que impide la amplificación de fragmentos mayores a 300-500 pares de bases (O'Rourke et al. 2000a), y a una relación inversa entre tamaño del fragmento y éxito en la amplificación. Obviamente, el grado de degradación depende del tiempo y de las condiciones en que actúan los agentes degradantes, entre los cuales se considera que la temperatura es un factor crítico, pudiéndose establecer la vida media del ADN de un organismo muerto sobre esta base (Garzón 2002). Así, considerando por ejemplo que la vida media del ADN a 1°C es de 6210 años (Tabla 6), la reducción hipotética de una molécula de 10^9 pares de bases a 256 pares de bases se produciría en 100 mil años aproximadamente si la temperatura media en ese período fuera de 1°C, y en un millón de años en caso de ser la temperatura de 10°C bajo cero. Por ende, en un ambiente con temperaturas medias de aproximadamente 17°C, es de esperar que una molécula de ADN del orden de magnitud del genoma humano se reduzca a fragmentos de 256 pb en unos 4100 años.

t (°C)	Vida media (Z) (Años)	Tiempo t_w (Años)
-10	65487	$1.06 \cdot 10^6$
-8	42062	$6.84 \cdot 10^5$
-6	27194	$4.42 \cdot 10^5$
-4	17695	$2.87 \cdot 10^5$
-2	11587	$1.88 \cdot 10^5$
0	76341	$1.24 \cdot 10^5$
1	6210	$1.01 \cdot 10^5$
2	5060	$8.23 \cdot 10^4$
3	4128	$6.71 \cdot 10^4$
4	3373	$5.48 \cdot 10^4$
5	2760	$4.49 \cdot 10^4$
6	2262	$3.68 \cdot 10^4$
7	1856	$3.02 \cdot 10^4$
8	1525	$2.48 \cdot 10^4$
9	1255	$2.04 \cdot 10^4$
10	1034	$1.68 \cdot 10^4$

Tabla 6: Vida media y tiempo de degradación de una molécula del orden de magnitud de un genoma humano a 256 pb (Garzón 2002).

Debe hacerse notar que las cifras mencionadas refieren a un pH de 7.6, y que la presencia de apatita puede producir un efecto *buffer* en las variaciones de acidez, estabilizando con ello el medio (Garzón 2002). Asimismo, hay otros factores, relacionados con el medio en el cual se encuentra la muestra, que pueden influir positivamente en la

preservación del ADN (Tabla 7). A grandes rasgos, se considera que la preservación de la materia orgánica sigue trayectorias semejantes, por lo que una baja tasa de racemización de aminoácidos será buena señal (aunque no garantía) de preservación de ADN (Poinar y Stankiewicz 1999). Otros autores consideran que en el caso del hueso, el grado de cristalinidad de hidroxiapatita es mejor indicador que la preservación del colágeno (proteína mayoritaria del hueso) (Götherström et al. 2002).

Factor favorable	Comentarios
Ausencia de microorganismos	Los microorganismos y su metabolitos son capaces de destruir el ADN por completo.
Aridez	La ausencia de agua reduce el daño oxidativo e hidrolítico.
Adhesión del ADN a superficies minerales	Las superficies minerales como la hidroxiapatita estabilizan el ADN.
Inhumación rápida tras la muerte	La formación de gases y la degradación de los tejidos acelera la invasión por microorganismos.
Sequedad / dureza de los tejidos	Los tejidos duros y secos previenen reacciones físicas y químicas. Huesos y dientes proveen de protección contra reacciones químicas y ataque por microorganismos.
Temperaturas bajas	Las temperaturas reducidas retardan la mayor parte de las reacciones químicas, así como la reproducción de microorganismos.
pH neutro o levemente alcalino	El aumento del pH conlleva la destrucción de los tejidos y con ellos del ADN.
Presencia de ácidos húmicos y fúlvicos	Los geopolímeros fenólicos previenen la actividad aeróbica.

Tabla 7: Factores favorables para la preservación de ADN antiguo en materiales arqueológicos (modificado de MacHugh et al. 2000).

1.4.2. Extracción de ADN de muestras antiguas

1.4.2.1. Breve historia

Los primeros análisis de ADN extraído de restos arqueológicos se remontan a 1984, año en el que se extrajo y clonó un fragmento de 229 pb de ADN mitocondrial de la piel desecada del *quagga*, un caballo extinto. El árbol filogenético resultante del análisis lo coloca en una rama común con la cebra, y separado del caballo y el burro (Higuchi 1984 en Rogan y Salvo 1990). Al año inmediato siguiente, se extrajo y clonó un fragmento de 300 pb de un repetido *AluI* de una momia egipcia, determinándose que el elemento clonado era representativo de toda la familia de insertos *Alu*. (Pääbo 1985 en Rogan y Salvo 1990). En retrospectiva, considerando la relación inversa entre temperatura ambiental y preservación de

ADN, y a la luz del trabajo de Marota et al. (2002), quienes consideran imposible la conservación de ADN en el medio egipcio más allá de unos pocos siglos, el trabajo seminal de Pääbo habría sido producto de contaminación con ADN exógeno, quedando así invalidado (véase no obstante Zink y Nerlich 2003).

La técnica de clonado es especialmente problemática al trabajar con ADN degradado como el que se recupera de restos arqueológicos: El ADN modificado puede inhibir o no ser reconocido por las polimerasas, endonucleasas y ligasas empleadas; los productos de ligación con moléculas alteradas puede afectar la capacidad de replicación del plásmido y la eficiencia de transformación; por último, los tamaños de los fragmentos de ADN recuperados (en el orden de los 10^2 pb) hacen que se requerirían bibliotecas enormes (en el orden de los millones de clones) para recuperar información útil. Es por eso que no fue sino hasta la implementación de la reacción en cadena de la polimerasa (PCR) (Mullis et al. 1986; Mullis y Faloona 1987) que se produjo un aumento en los intentos de recuperación y estudio de ADN conservado en tejidos antiguos durante la década de 1990 (Hagelberg 1994). La elección más frecuente en cuanto a ADN a amplificar fue el mitocondrial, en virtud de que está presente en miles de copias por célula, haciendo asimismo miles de veces más probable la amplificación de un segmento mitocondrial que uno nuclear.

En retrospectiva, los estudios pioneros trabajaron sobre dos conceptos erróneos (Jones 2003): el primero, que el ADN podría recuperarse únicamente de restos antiguos de excepcional conservación; el segundo, que en caso de conservarse, el ADN se conservaría en forma indefinida. Los primeros intentos de extraer ADN antiguo se hicieron sobre tejidos blandos excepcionalmente conservados (*e.g.* Pääbo et al. 1988), pero trabajos posteriores (*e.g.* Cooper et al. 1992) mostraron que el tejido esquelético suele rendir mejores resultados que el tejido epitelial o muscular al momento de la extracción de ADN. Esto se debe a que las bajas proporciones de agua inhiben rápidamente los procesos químicos que alteran el ADN tras la muerte del individuo. A esto se agrega la presencia de minerales como la hidroxiapatita (fracción mineral del hueso), que estabilizan la molécula de ADN (MacHugh et al. 2000; Götherström et al. 2002). El segundo preconcepto dio lugar a un optimismo desmedido en cuanto a las antigüedades de los restos de los que se podía recuperar ADN amplificable, olvidando en muchos casos que la sensibilidad de la PCR es, a la vez que una ventaja, su principal inconveniente. Lindahl (1993b) procuró disminuir este optimismo al contestar a las pretensiones de recuperación de ADN de insectos conservados en ámbar (Cano et al. 1993). Menciona la importancia de la reproducibilidad de los datos, y comenta que es

extremadamente improbable que las extracciones reportadas de ADN de restos de 100 millones de años sean legítimas. Agrega su decepción respecto a que no se haya probado antes con restos en ámbar de edades más recientes ($10^5 - 10^6$ años). De hecho, si se logra amplificar ADN antiguo de la muestra que sea, no hay manera posible de saber si es realmente antiguo (Lindahl 1993b). Las réplicas al trabajo de Lindahl hacen referencia a que éste se basa fundamentalmente en trabajos experimentales con ADN en solución, y que pierde de vista la extraordinaria capacidad de preservación de algunos medios como, por ejemplo, el ámbar (DeSalle y Grimaldi 1994). Dos factores serían los responsables de la extraordinaria preservación de tejidos en este medio: la desecación extrema producida por la resina, y el ambiente antimicrobiano y anóxico que proporciona (Poinar et al. 1994). Asimismo, las tareas de extracción y análisis de ADN en el trabajo de Poinar (1993) cumplen con todos los controles recomendados, incluidos blancos múltiples en variadas etapas. No obstante, el intento de reproducir extracciones de ADN de insectos incluidos en ámbar fracasó, dando lugar a las lapidarias afirmaciones de que, probablemente, aislar y amplificar ADN de especímenes de antigüedad mayor a 100 mil e incluso 50.000 años es una tarea imposible (Austin et al. 1997), y que todas las secuencias de antigüedad mayor a un millón de años son artefactos (Pääbo et al. 2004).

Trabajando con más prudencia, se ha ido sin embargo rompiendo barreras en cuanto a la antigüedad de los especímenes de los cuales se ha recuperado y analizado ADN, siendo probablemente el análisis más notable y debatido el de Krings et al. (1997), que parecía descartar definitivamente la posibilidad de que el hombre de Neandertal hubiera aportado al acervo genético del hombre moderno. Por otra parte, el análisis de ADN antiguo ha generado resultados útiles en cuanto a la relación de especies extintas (como la moa, el lobo marsupial y el ganso hawaiano entre otras 50 especies) con especies existentes, a los orígenes de la domesticación de distintas especies de plantas y animales (como el maíz y el perro), a la reconstrucción dietaria y ambiental a través de coprolitos y muestras sedimentarias, al estudio de patógenos en restos antiguos, etcétera (Pääbo et al. 2004). Estos campos tienen en muchos casos vínculos y aplicaciones en estudios antropológicos (Tabla 8).

Aplicación	Implicancias	Marcadores utilizados
Determinación del sexo	Patrones de matrimonio y entierro Tasas de mortalidad diferenciales entre sexos Patrones diferenciales por sexo de enfermedades, dieta, estatus y bienes materiales	Marcadores de los cromosomas X e Y
Análisis de ADN no humano	Patrones de caza y dieta Domesticación Reconstrucción ambiental Historia y patrones de aparición de enfermedades prehistóricas e históricas	ADN de mitocondrial, de cloroplastos y nuclear
Parentesco materno y paterno	Estructuras sociales Estatus Patrones de matrimonio Costumbres funerarias Migraciones	ADN mitocondrial y de cromosomas sexuales; microsatélites autosomales
Continuidad / reemplazo poblacional	Rastreo de migraciones prehistóricas Relaciones de ascendencia / descendencia entre grupos Relaciones entre grupos antiguos y modernos con semejanzas morfológicas o culturales	ADN mitocondrial y nuclear
Reconstrucción filogenética	Evolución de especies Parentesco entre especies extintas y vivientes Orígenes del humano moderno	ADN mitocondrial y nuclear

Tabla 8: Aplicaciones del ADN antiguo en estudios antropológicos (tomado de Kaestle y Horsburgh 2002).

1.4.4.2. Protocolos de extracción y criterios de autenticidad

Hay dos variantes fundamentales de formas de extracción de ADN antiguo (ADNa) de restos óseos (O'Rourke et al. 2000a; Kaestle y Horsburgh 2002): El método de fenol/cloroformo/proteinasas K y el método de suspensión de sílice. El primero es una modificación del protocolo de fenol/cloroformo (Sambrook et al. 1989) y consiste en la digestión de la muestra (previamente descalcificada con EDTA) usando proteinasas K, y el uso de fenol/cloroformo/isoamyl alcohol 25:24:1 (*phenol/chloroform/isoamyl alcohol*, p/c/i) y cloroformo a fin de separar el ADN (que queda disuelto en la fase acuosa resultante de agregar el p/c/i o el cloroformo al producto de digestión) de los demás compuestos del tejido de base. La precipitación se realiza utilizando etanol o isopropanol, y el precipitado es disuelto en *buffer* TE.

El segundo procedimiento (Höss y Pääbo 1993) es una modificación del método inicialmente planteado por Boom et al. (1990) y se basa en el uso de tiocianato de guanidinio

(GuSCN) para lisar las proteínas y facilitar la unión del ADN a partículas de sílice. Tras un período de incubación a 60°C, se centrifuga la solución de GuSCN para poder eliminar restos celulares, y el sobrenadante es transferido a una suspensión de sílice conteniendo un *buffer* con GuSCN. Tras una incubación breve y centrifugado, el *pellet* de sílice es lavado con un *buffer* de extracción modificado, etanol y acetona en ese orden. El *pellet* lavado se disuelve en *buffer* TE.

Ambos métodos tienen ventajas y desventajas. El protocolo de fenol/cloroformo y proteinasa K tiene por ventaja el extraer cantidades mayores de ADN que el protocolo de sílice, pero coextrae elementos que actúan como inhibidores de la reacción de PCR (entre ellos productos de Maillard, resultantes de la hidrólisis de azúcares). El protocolo de sílice rinde a su vez ADN más puro, pero el propio sílice es un inhibidor de la PCR, y el GuSCN presenta una fuerte afinidad a los ácidos nucleicos y por ende una alta propensión a la contaminación por ADN exógeno. Se han diseñado también protocolos modificados, en los que se combina el procedimiento de fenol/cloroformo con el de sílice, como el protocolo empleado en este trabajo.

Por causa de las exiguas cantidades de ADN amplificables en restos arqueológicos (en teoría hasta una sola copia del fragmento deseado), el riesgo de contaminación por ADN exógeno es constante. En los trabajos pioneros de ADN antiguo, la mayor parte del ADN recuperado provenía de organismos ajenos al de interés (ADN bacteriano en casos de restos conservados en *permafrost*, o vegetal en restos recuperados de turberas). El problema de la contaminación es especialmente crítico en el caso del estudio de ADN humano, en la medida en que la amplificación del ADN de la muestra excluyendo el ADN contaminante no puede realizarse mediante la utilización de cebadores específicos. Las dudas sobre contaminación se pueden disipar mediante la comparación de la secuencia amplificada con las secuencias homólogas de los investigadores, pero siempre se han de extremar las precauciones a fin de evitar contaminación. Conviene aplicar medidas preventivas desde la misma recuperación de los restos, que incluyen evitar la limpieza (en seco o con agua) de los especímenes antes de su introducción al laboratorio, no utilizar sustancias conservantes, no mezclar especímenes antiguos y modernos, y excavar con guantes, colocando las piezas elegidas para análisis individualmente en bolsas plásticas secas (Yang y Watt 2005). Las precauciones en el laboratorio – que debe ubicarse en un ambiente aislado – son el uso de materiales descartables y vestimenta protectora durante el proceso de extracción. El material no descartable debe ser

tratado con radiación ultravioleta y lavado con soluciones de hipoclorito de sodio o ácido clorhídrico.

Criterio	Utilidad(es)
1. Clonado de productos de amplificación y secuenciado de clones múltiples	Detección de heterogeneidad en los productos de amplificación, sea por contaminación, daño al ADN o PCR saltatorio (“ <i>jumping PCR</i> ”)
2. Controles de extracción y de PCR: Cada juego de extracciones debería incluir por lo menos un control de extracción sin material de muestra y tratamiento idéntico. Igualmente, para cada juego de PCR debe incluirse varios controles negativos.	Identificar contaminación y distinguir entre contaminación introducida durante la extracción de ADN y contaminación introducida durante la preparación de la PCR.
3. Repetición de amplificaciones a partir del mismo o diferentes extractos	a) Detección de contaminación esporádica b) Detección de cambios consistentes debidos a lesiones inductoras de error (“ <i>miscoding lesions</i> ”) en extractos con número muy bajo de moléculas de molde.
4. Cuantificación del número de moléculas amplificables de ADN: Debe realizarse para cada par de cebadores a utilizarse puesto que el número de moléculas amplificables varía drásticamente con el tamaño del fragmento a amplificar, su composición de bases y la sensibilidad del cebador.	Considerar la posibilidad de cambios consistentes, los cuales pueden excluirse en extractos con más de 1000 moléculas de molde, bastando con una sola amplificación para obtener resultados fiables.
5. Correlación inversa entre eficiencia de la amplificación y tamaño del fragmento amplificado	Puesto que el ADN antiguo está fragmentado, esta correlación debería darse en cualquier tipo de extracto.
6. Ensayos bioquímicos de preservación macromolecular	La mala preservación bioquímica indica baja probabilidad de contener ADN. La buena preservación bioquímica es un argumento a favor de la autenticidad de una secuencia antigua.
7. Exclusión de insertos nucleares de ADNmt	Es altamente improbable que varios pares de cebadores amplifiquen en forma preferencial un inserto nuclear. Por ende, las sustituciones en la parte solapante de dos productos puede advertir sobre la amplificación de un inserto nuclear. La falta de diversidad en estudios poblacionales también puede tomarse como indicación de resultados debidos a insertos nucleares.
8. Reproducción en un segundo laboratorio	El propósito es similar al de los criterios 2 y 3, es decir, detectar contaminación de reactivos o muestras en el laboratorio. Esto no sería necesario siempre, sino cuando se obtienen resultados inesperados. Nótese que los contaminantes presentes en la muestra antes de su manipulación serán reproducidos en el segundo laboratorio.

Tabla 9: Criterios para garantizar la autenticidad del ADN antiguo. Tomado de Pääbo et al. (2004).

A nivel de análisis los criterios – que se han multiplicado en el correr de los últimos años – incluyen la secuenciación de varios productos de PCR, la realización de ensayos para verificar la preservación de otras macromoléculas biológicas, el uso de múltiples controles de extracción y PCR, y la replicación de resultados en distintos laboratorios (Pääbo et al. 2004) (Tabla 9).

1.5. El análisis de ADN mitocondrial antiguo en América

Los análisis genéticos de poblaciones antiguas tienen como dificultad fundamental el lograr una muestra temporalmente consistente: En la medida en que muchas veces se trabaja con muestras atribuibles a períodos amplios, en el orden de los siglos, cualquier generalización que se haga de la muestra a la población debe manejarse con prudencia. No obstante, los análisis de ADN mitocondrial en muestras americanas antiguas se multiplican, y han arrojado resultados relevantes en varios aspectos. Algunos trabajos representativos de las regiones abarcadas y de los avances hechos en la materia son las siguientes:

Little Salt Springs, EE.UU: El sitio de Little Salt Springs está ubicado en una zona anegadiza de la cual se ha recuperado tejidos blandos conservados. Pääbo et al. (1988) extrajeron exitosamente ADN humano de una muestra de tejido cerebral datado en 6900 años A.P. La presencia de dos repetidos en la región intergénica COII- tRNA^{Lys}, la presencia en esta misma región de una mutación rara en amerindios y la presencia del sitio de restricción de *HincII* en la posición 13259 (cuya ausencia caracteriza al haplogrupo C – ver *supra*) lleva a suponer que el individuo analizado era portador de un linaje actualmente ausente en poblaciones americanas.

Windover, EE.UU: El sitio de Windover, en una turbera ubicada en Florida, rindió los restos de 176 individuos, 91 con tejido blando intacto (Hauswirth et al. 1994), con fechados de entre los 8100 y los 7000 años A.P. Para verificar la hipótesis de una población continua, se analizó la HVR I del ADN mitocondrial de 14 individuos; de éstos, 5 compartían una mutación A→G en la posición 16219. Dado que estos 5 individuos abarcan el rango temporal completo de la muestra, se les consideró relacionados por vía materna y, por lo tanto, integrantes de una población temporalmente continua.

Colombia: En un estudio en el cual se comparan sitios de restricción y secuencias de la región hipervariable de 6 momias colombianas de entre 1500 y 100 años de antigüedad, se

pudo atribuir a todos los individuos a haplogrupos amerindios conocidos. No obstante, uno de los individuos, portador de la delección característica del haplogrupo B, presentaba sólo una de las mutaciones típicas en la región hipervariable I, lo cual llevó a los autores a plantear la posibilidad de que se trate de un linaje extinto, o presente en baja frecuencia en las poblaciones amerindias actuales (Monsalve et al. 1996).

Región amazónica: Ribeiro-Dos-Santos et al. (1996) realizaron un análisis de esqueletos provenientes de 11 sitios de la región amazónica, con fechados de entre 4000 y 500 años A.P. En total, se logró analizar la HVR I de 18 individuos, de los cuales 14 se pudieron asignar a los 5 haplogrupos principales, quedando 4 en la categoría “otros” que no obstante contenían mutaciones asiáticas. Al igual que en los otros casos con mutaciones no categorizables, se atribuye su ausencia en muestras modernas a su desaparición asociada a la reducción poblacional tras la llegada del europeo.

Sudoeste Norteamericano: El análisis de 47 esqueletos provenientes de sitios de la cultura Fremont (1450-600 años A.P.), procedente de la “gran cuenca” del Sudoeste norteamericano, rindió 32 individuos en los que se pudo determinar uno de los cuatro haplogrupos por RFLP. La alta frecuencia de haplogrupo B en esta muestra (73%) sirvió para plantear un origen común de los pueblos Fremont y Anazasi – ubicados al sur de los sitios Fremont, y culturalmente distinguibles de éstos – en poblaciones antiguamente residentes más al sudoeste, y representadas hoy día por poblaciones de México (Parr et al. 1996).

Un análisis posterior de restos atribuidos a la cultura Anazasi (Carlyle et al. 2000) mostró frecuencias que los agrupa cerca de los indios Pueblo modernos, con quienes se relacionan por ascendencia. No obstante, al comparar las poblaciones Fremont con los habitantes actuales de la zona anteriormente ocupada por ellas, hallaron que las diferencias en frecuencias de haplogrupos son lo suficientemente importantes como para fortalecer la sospecha de una migración hacia la gran cuenca en épocas recientes (Carlyle et al. 2000).

La posible migración hacia zonas de la gran cuenca por gente de otra región podría corroborarse si se comprobaba la heterogeneidad temporal de los pueblos de esa zona, lo cual fue llevado a cabo por Kaestle y Smith (2001), en un estudio donde se comparó las frecuencias de los haplogrupos fundadores en habitantes modernos de la Gran Cuenca con las frecuencias correspondientes de esqueletos de sitios prehistóricos de Nevada, la mayor parte de ellos fechados entre 6000 y 350 años A.P. Los resultados muestran frecuencias significativamente diferentes a las de los habitantes modernos, cuya semejanza genética con grupos de California estaría apoyando una hipótesis migratoria.

Tierra del Fuego – Patagonia: Mediante un análisis de 60 esqueletos de aborígenes extintos de Tierra del Fuego y Patagonia, la mayor parte de ellos de épocas recientes (aprox. 200 años A.P.), Lalueza et al. (1997) plantean la ausencia de los linajes A y B de los habitantes de Tierra del Fuego. Los análisis de García-Bour et al. (2004) de restos de 24 individuos atribuibles a cuatro grupos étnicos fueguinos (Aonikenk, Selk'nam, Yamana y Kaweskar) dan los mismos resultados, presentando los haplogrupos C y D en estas muestras una alta diversidad. García-Bour et al. (2004) plantean que la ausencia de linajes A y B junto con la alta diversidad interna de los otros dos haplogrupos no puede deberse a un fenómeno de deriva, sino al origen de las poblaciones del sur del continente americano en una primera oleada migratoria, carente de los grupos A y B. Otros resultados sobre muestras menores (Demarchi et al. 2001a) tienden a apoyar estos resultados.

Mesoamérica: La comparación de los datos mitocondriales de los restos de 25 individuos recuperados del sitio del Clásico tardío y Postclásico Maya de Xcaret (600-1500 d.C.) con los datos del emplazamiento Clásico Maya de Copán y con datos modernos muestran, por un lado, la persistencia de la tendencia general de altas frecuencias del haplogrupo A en la península de Yucatán, y por el otro, el carácter posiblemente cosmopolita del emplazamiento de Copán, que si bien era de tradición Maya, estaba rodeado de grupos no Maya (González-Oliver et al. 2001).

Antillas: Lalueza-Fox et al. (2001) propusieron testar las tres hipótesis que se han planteado sobre bases arqueológicas acerca del origen de las poblaciones caribeñas (Norteamérica, América Central y Sudamérica). Analizando la HVR I de una muestra de restos pertenecientes a Taíno (pueblos agricultores habitantes de las Antillas al momento de la llegada del europeo) del sitio La Caleta, República Dominicana (670 a 1680 d.C.), hallaron frecuencias de haplogrupos radicalmente diferentes de las centroamericanas, y características de secuencia que aproxima a los Taíno a poblaciones amazónicas.

Acto seguido, se abordó el análisis de restos óseos pertenecientes a Ciboney (pueblos cazadores-recolectores de las Antillas, que habrían poblado las islas con anterioridad a los Taínos) (Lalueza-Fox et al. 2003). Si bien la distribución de haplogrupos y las mutaciones de la HVR I en la muestra analizada (4700-1600 A.P.) son significativamente diferentes de cualquiera de las tres regiones propuestas de origen, los datos indican que la probabilidad de origen sudamericano es en el orden de las 1000 veces mayor que las otras dos regiones. Se sugiere una región común de origen para todas las poblaciones caribeñas, en torno a la desembocadura del Orinoco, en tres oleadas sucesivas.

Chile: Moraga et al. (2001) analizaron momias provenientes de la región de Tarapacá, en el extremo norte chileno, recuperadas de sitios que abarcan un rango temporal de 1300 años (de 2900 a 600 años A.P.), logrando diagnosticar el haplogrupo por RFLP en 32 casos de 42. Las frecuencias observadas para los cuatro haplogrupos principales es consistente con un posible origen amazónico de las poblaciones andinas.

Acto seguido, Moraga et al. (2005) pusieron a prueba las hipótesis, formuladas sobre evidencia arqueológica, de la influencia en los grupos de la costa Pacífica del norte de Chile de las culturas del Titicaca y los Andes Centrales. Para ello, se analizó la HVR I de 61 individuos de un intervalo de 3.900 años de los valles de Lluta, Azapa y Camarones, observando una tendencia al incremento del haplogrupo B y la disminución de la frecuencia del haplogrupo A a lo largo del período, apuntando a un posible flujo génico.

Noroeste argentino: El análisis realizado por Dejean et al. (2005) de secuencias de la HVR I de 16 individuos del sitio de Pampa Grande (1000 a 1500 A.P.) en la provincia de Salta, arrojó como resultado la ausencia de haplogrupo C en la muestra, pero frecuencias de haplogrupos B y D consistentes con los grupos aborígenes argentinos actuales en general, y con grupos chaqueños en particular, separándose fuertemente de los grupos fueguinos en los análisis de distancias.

Uruguay: En un estudio preliminar realizado sobre ADN extraído de esqueletos recuperados en sitios arqueológicos del Departamento de Rocha, con fechados de entre 2000 y 400 años A.P., se secuenció la HVR I de 5 individuos; 4 de ellos pudieron ser incluidos en los haplogrupos previamente establecidos, perteneciendo 3 al haplogrupo C y 1 al haplogrupo B. Dos de los individuos de haplogrupo C, recuperados en el mismo sitio, presentaron la misma secuencia en el segmento analizado (160 y 164 pb), con una mutación rara en la posición 16288, con lo cual se infiere que se trata de individuos con probable parentesco estrecho entre sí. Del quinto individuo se secuenciaron sólo 150 pb, lo cual deja pendiente la determinación concluyente del individuo, si bien debe destacarse que no presenta ninguna mutación típica de haplogrupos amerindios en el segmento secuenciado (Bertoni et al. 2004, Figueiro y Sans, datos sin publicar). Posteriormente, por RFLP se ha podido caracterizar a otros tres individuos, uno como haplogrupo A y dos como haplogrupo B (Sans y Figueiro 2005).

Poco después, en el marco del análisis de los restos, repatriados en 2002, del charrúa Vaimaca Perú (llevado a Francia en 1833 tras haber sido capturado por el ejército nacional en la campaña de Salsipuedes), se extrajo ADN de dos piezas dentales, analizándose

posteriormente el ADN mitocondrial tanto al nivel de RFLP como de secuencia de la HVR I. Los resultados son consistentes a ambos niveles, perteneciendo Vaimaca al haplogrupo C. Asimismo, la HVR I presenta una mutación en la posición 16288 (Sans 2004; Sans y Figueiro 2005); esta mutación fue, hasta el momento, encontrada únicamente en nuestro país en asociación con las 4 mutaciones típicas del haplogrupo C, en dos de los individuos analizados por Bertoni et al. (2004), en un individuo actual de la ciudad de Melo (Sans et al. 2006) y en otro de Montevideo (Pagano et al. 2005). Esto plantea una continuidad poblacional en nuestro territorio entre los pobladores prehistóricos de la región este y los cazadores-recolectores conocidos en tiempos históricos (Sans 2004; Sans y Figueiro 2005).

En resumen, los estudios de ADN mitocondrial antiguo llevados a cabo hasta ahora en el continente americano muestran: a) la estabilidad de las frecuencias de los haplogrupos mayores a lo largo del tiempo. Esto se ha observado en América del Norte, pero aún está por probarse en América del Sur; b) la subsiguiente utilidad de estos datos a la hora de la reconstrucción de historias de movimientos y continuidad poblacionales; c) una nota constante de precaución en lo que respecta a que, a mayor rango temporal abarcado, mayor margen de error tienen las historias poblacionales inferidas, ya que rara vez se está trabajando con una población en el sentido estricto (Hauswirth et al. 1994; Carlyle et al. 2000). El estudio llevado adelante en Norris Farms (Stone y Stoneking 1993, 1998) es hasta el momento el trabajo más riguroso en lo referente al tamaño de la muestra y el rango temporal abarcado.

En la presente investigación se pretende abordar el problema de la permanencia/continuidad de poblaciones en un sitio de la región pampeana Argentina, así como las características de una población del Holoceno temprano.

1.6. La región pampeana: Poblamiento y prehistoria

La región de llanuras que se extienden entre el Océano Atlántico y el pie de las Sierras Centrales y la precordillera Mendocina, a grandes rasgos entre los 54° y 63° de longitud oeste y los 31° y 40° de latitud sur, recibe el nombre de Pampa o Pampas. Su clima actual es templado, caracterizado por un gradiente de humedad que declina en dirección este-oeste, con un promedio anual de lluvias que presenta la misma tendencia: Las precipitaciones más altas (mayores a 900 mm anuales) se registran en el noreste y las más bajas (menos de 500 mm

anuales) en el sudoeste. Las temperaturas medias en el noreste oscilan entre los 24°C en verano y los 10°C en invierno, con una variación general 4°C menor en el sudoeste de la región. En síntesis, el área se clasifica como subhúmeda-húmeda en el este y subhúmeda-seca en el oeste, con un borde semiárido más al sur (Politis et al 2004).

Este territorio, relativamente inexplorado por el europeo hasta el siglo XVII, fue prácticamente desconocido desde el punto de vista arqueológico hasta finales de la década de 1970. Desde entonces, se ha logrado un avance sustancial en las investigaciones de su poblamiento y procesos culturales, que se extienden como mínimo hasta el Pleistoceno final, hace 12.000 años (Berón y Politis 1997). A continuación se presenta una síntesis de la prehistoria de la región pampeana, con énfasis en la región pampeana argentina, basada en Berón y Politis (1997), Politis (2000), Politis y Madrid (2001) y Politis et al. (2004).

Pleistoceno final / Holoceno temprano (hasta 6.500 años A.P.): A fines del Pleistoceno (12.000-10.000 años A.P.) primaban condiciones climáticas frías en ambientes semiáridos a áridos. Los estudios realizados indican que hace 11.000 años el nivel del mar se encontraba a unos 60 m por debajo del nivel actual, con lo cual el actual Río de la Plata se presentaba como una gran planicie costera dominada por una red de drenaje con brazos deltaicos, lagunas costeras y sistemas de dunas (Ayup en Politis et al. 2004). Desde los inicios del Holoceno se registra paulatinamente una tendencia hacia el incremento de los niveles de temperatura y humedad que culminaron en el Holoceno tardío (aprox. 1.000 años A.P.) con el establecimiento de las condiciones actuales.

Es durante finales del Pleistoceno que se produce el poblamiento de la Pampa, con dataciones de hasta 11.200 años en la zona de Tandilia y de 12.300 años en la zona interserrana. Durante el Holoceno temprano, aproximadamente entre 10.000 y 7.000-6.500 años antes del presente, la región pampeana, especialmente en su porción oriental, era todavía un poco más fría y árida que en la actualidad. Durante todo este período, el rasgo arqueológico común consiste en la presencia de puntas de proyectil del tipo “cola de pescado” (Figura 4) y la explotación de guanaco y fauna actualmente extinta (géneros *Equus*, *Doedicurus*, *Megatherium* entre otros). Se registran múltiples sitios relacionados con actividades múltiples y reocupación (campamentos base) y actividades específicas (caza, procesamiento, ocupaciones de corta duración).

Holoceno medio (6.500-3.500 años A.P.): A mediados del Holoceno, el progresivo aumento en la temperatura registró un pico en el cual se desarrollaron los eventos transgresivos del avance marino (de 2 a 4,5 m por encima del nivel actual), cuyas fechas son controvertidas (entre 8.000 y 5.000 años A.P.). Finalmente, aproximadamente entre el 4.000 y el 3.000 A.P. se estableció la actual línea costera.

Hacia esta etapa se registran una serie de eventos clave en la comprensión de la evolución

de los cazadores-recolectores pampeanos: Hacia el inicio de este período se extinguen los últimos representantes de la megafauna pleistocénica. Con la desaparición de un grupo de animales se registra asimismo el afianzamiento de un modo de vida, en donde el guanaco, el venado y el ñandú – en ese orden – se vuelven recursos recurrentemente explotados, con estrategias de caza masiva. Asimismo se registran las primeras ocupaciones de la zona costera y las zonas áridas del oeste así como eventos de inhumación y representaciones simbólicas.

Holoceno tardío (3.500-500 años A.P.): Este período muestra evidencias de intensificación de la ocupación humana de la región, siendo el guanaco, el venado y la nutria los principales recursos faunísticos. Se evidencian cambios significativos en cuanto a la tecnología, como ser la producción de una mayor variedad de instrumentos, la introducción del arco y la flecha, y la adquisición de la alfarería. Asimismo, hacia el final del Holoceno tardío se identifica un incremento importante de contactos interétnicos, producidos por el ingreso de grupos guaraníes por el noreste, así como Tehuelches septentrionales por el sur (hacia el 1.500 A.P.) y Araucanos por sur y oeste hacia el siglo XVI. La interacción con éstos últimos, junto con la adopción del caballo y el ganado traídos por el europeo, produjo cambios sustanciales en la dinámica de los grupos pampeanos.

Una larga serie de sitios, investigados a partir de la década de 1980, sirvió de base para la reconstrucción de más de 10.000 años de prehistoria pampeana. Uno de ellos, con fechados que se extienden hasta el Pleistoceno final y evidencias que incluyen inhumaciones humanas, es el sitio Arroyo Seco 2.

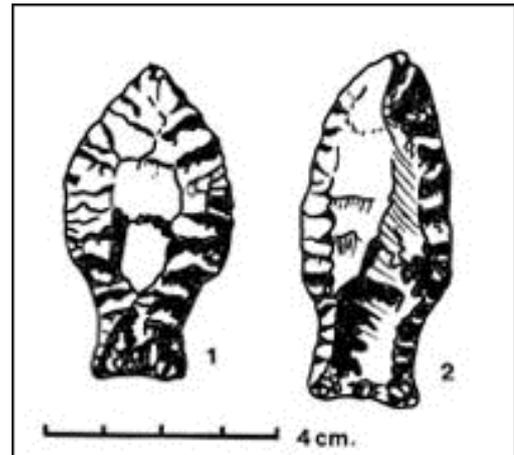


Figura 4: Puntas de proyectil “cola de pescado” recuperadas en Cueva Fell, Chile (Tomado de Massone y Prieto 2004)

1.7. El sitio Arroyo Seco 2

El sitio 2 de la Localidad Arqueológica Arroyo Seco se encuentra localizado en el Área Interserrana Bonaerense, a los 38° 21' 38" de Latitud Sur, y a los 60° 14' 39" de Longitud Oeste (Figura 5), en proximidades de la ciudad de Tres Arroyos, en la provincia de Buenos Aires. Se ubica sobre una pequeña lomada, con una suave pendiente hacia el este y sureste en dirección al cauce del Primer Brazo de los Tres Arroyos. En la actualidad, el sitio funciona como Escuela de Campo de la Facultad de Ciencias Sociales (Olavarría) de la Universidad Nacional del Centro de la Provincia de Buenos Aires, en virtud de un convenio firmado entre esta Universidad y la Municipalidad de Tres Arroyos (Barrientos 1997).

La existencia de material arqueológico en la Localidad Arroyo Seco es conocida desde hace aproximadamente setenta años, cuando el Sr. José A. Mulazzi realizó recolecciones superficiales en el lugar. A comienzos de la década del '70, un grupo de aficionados de la ciudad de Tres Arroyos realizó excavaciones en lo que hoy se conoce como Sitio 1 y Sitio 2, recuperando un total de cinco esqueletos humanos y restos óseos de fauna extinguida. En enero de 1977 comienza una investigación sistemática del sitio 2 bajo la dirección del Dr. Alberto Rex González, jefe de la División Arqueología del Museo de Ciencias Naturales de La Plata. Se llevaron a cabo excavaciones intensivas entre los años 1979 y 1981, abriéndose una superficie total de 183 m² y recuperándose entre otros materiales restos de megafauna extinta y diecisiete esqueletos humanos (Barrientos 1997). Otra serie de excavaciones se realizó entre 1986 y 1988, y posteriormente en 1992, 1995 y 1996, recuperándose por lo menos 20 individuos adicionales.

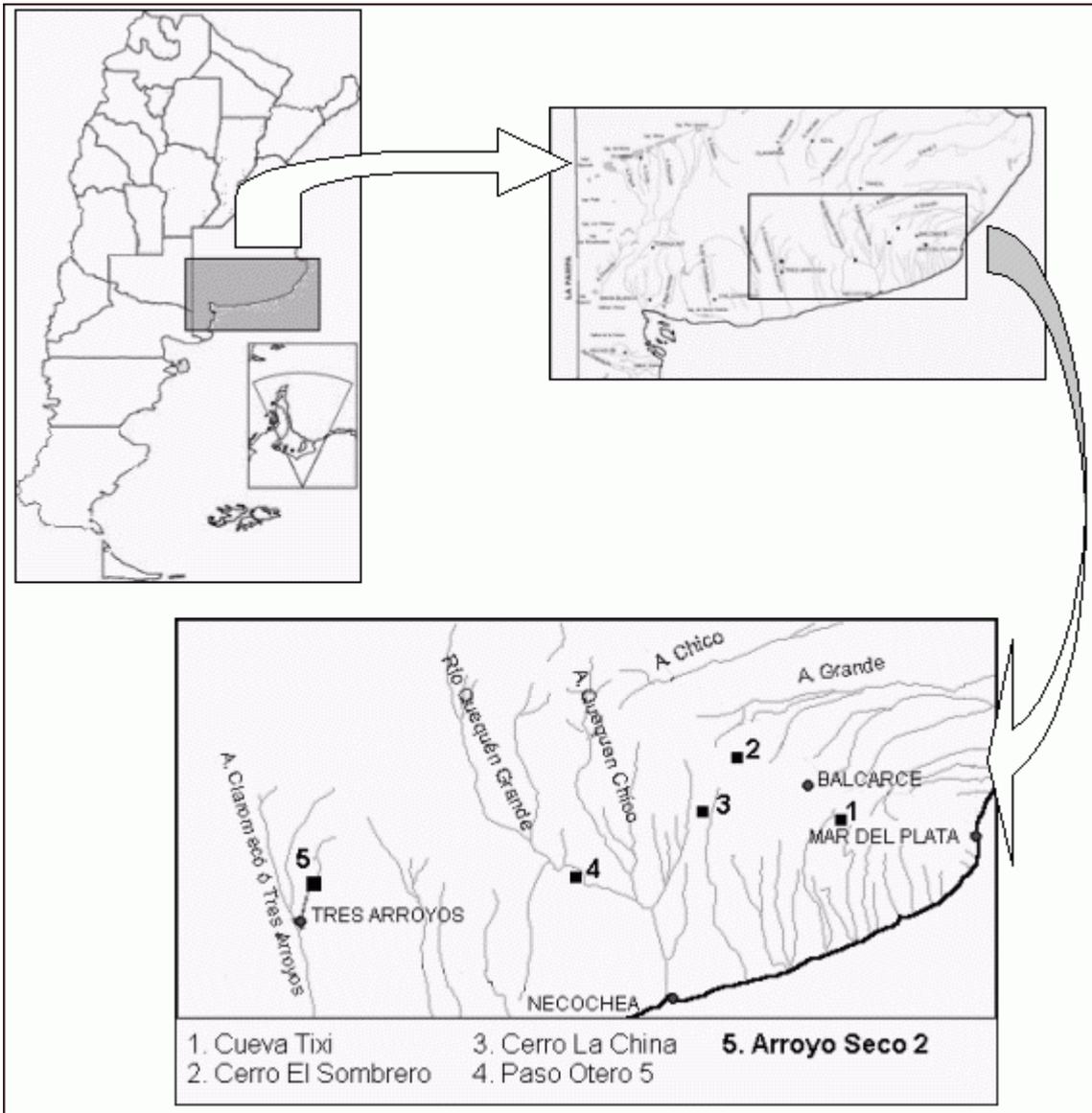


Figura 5: Ubicación del sitio Arroyo Seco 2 y otros sitios tempranos de la región pampeana.

El sitio fue ocupado repetidamente, aunque la identificación de las evidencias de cada ocupación es difícil ya que las asociaciones originales de restos fueron modificadas por diversos procesos, naturales y antrópicos. No obstante, se han identificado asociaciones recurrentes de restos en distintos sectores, las que podrían representar una o varias ocupaciones de bandas cazadoras-recolectoras con características similares. Del análisis de los materiales resultaron tres "componentes" para el sitio 2 de Arroyo Seco, que indican tres respectivos tipos de ocupación humana (Politis 1989; Figura 6).

El *Componente Superior* está formado por los restos recuperados en la unidad superior, conformada por el suelo actual, llamada "X". Se trata de una unidad fuertemente perturbada, por lo cual los restos están removidos, quedando evidencias relacionadas con la

tecnología de los ocupantes (restos líticos y cerámicos) pero poco o nada de su economía o preferencias de caza.

El *Componente Medio*, mejor conservado y caracterizado, se ubica en la parte superior de la unidad "Y" (compuesta por 0,40 a 0,30 m de limo grueso a arena fina), y contiene una variedad importante de artefactos líticos (tanto instrumentos como desechos de talla) asociados con restos de fauna que permiten deducir una economía basada fundamentalmente en la explotación del guanaco como presa principal de caza, y del venado de las pampas y del ñandú como recursos complementarios. Se hallaron también elementos de molienda que sugieren el procesamiento de productos vegetales silvestres (Politis 1989). Un fechado de 7540 ± 80 años A.P. sobre hueso de guanaco estaría asociado al inicio de este componente (Politis y Madrid 2001).

En la parte inferior de la unidad "Y" y en la siguiente (unidad "S") se ubica el *Componente Inferior* del sitio, representado por fechados entre los 12.000 y los 7.000 años A.P, y en el que se hallan las evidencias de asociación entre los mamíferos pleistocénicos y el hombre. Los motivos de la dispersión de los fechados es discutida, pudiendo deberse a problemas de datación o, efectivamente, a una ocupación del sitio en episodios breves pero continuos durante cuatro o cinco mil años. Es claro no obstante que en este período el sitio habría funcionado como campamento base, en el cual se llevarían a cabo etapas de terminación o mantenimiento de artefactos líticos y procesamiento de recursos faunísticos tales como guanaco, venado, ñandú y algunos mamíferos extintos, sobre una base regular y comprendiendo distintas zonas tales como costa atlántica, llanura y sierras (Politis y Madrid 2001).

Los procesos pedogenéticos que tuvieron lugar aproximadamente hacia el 6,000 A.P. provocaron la lixiviación de carbonatos, lo que condujo a una progresiva alcalinización de los sedimentos y a la formación de una capa con alto contenido de carbonato de calcio, conocida como unidad "S" hace por lo menos 5,700 años. Finalmente, por debajo de ella aparece la unidad "Z", similar a la unidad "Y", pero con tonalidades más grisáceas, que se extiende hasta los 4 m de profundidad.

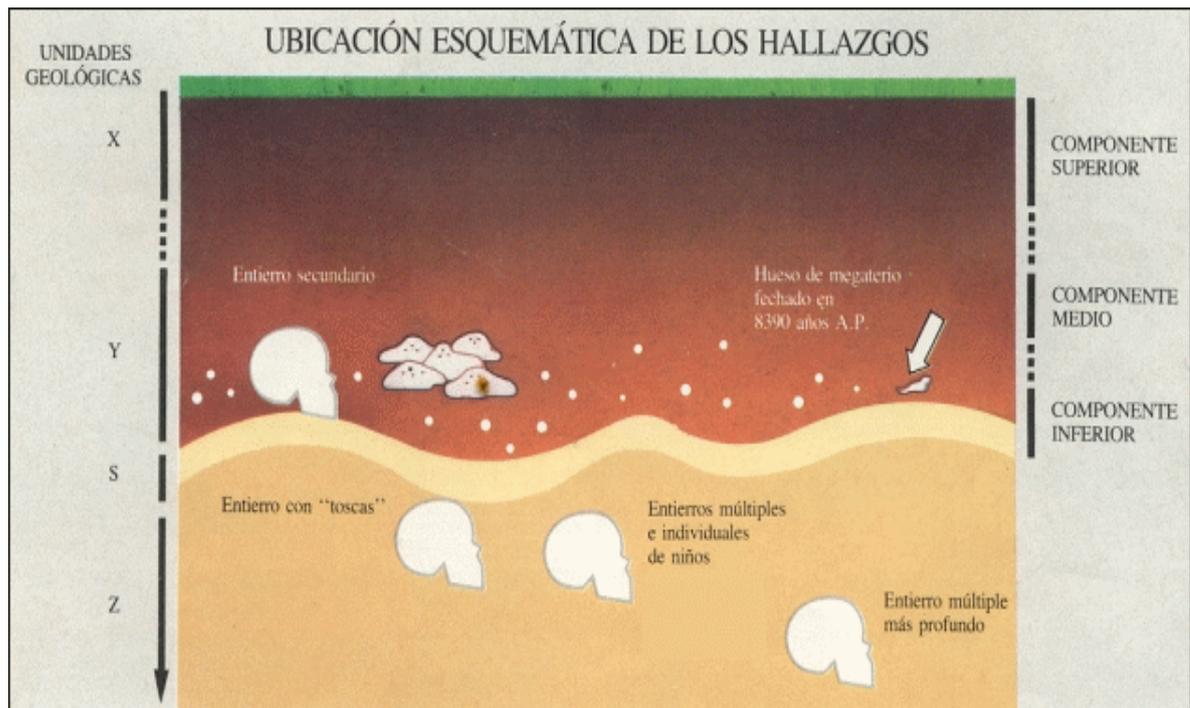


Figura 6: Ubicación esquemática de los hallazgos. Tomado de Politis (1989).

Es en la unidad “Z” que se ubica la mayoría de los entierros humanos recuperados en el sitio Arroyo Seco 2. Éstos se pueden adscribir a cuatro “eventos”, fechados con relativa precisión (Politis y Madrid 2001):

Evento 1: Datado entre 7800 y 7600 años A.P. Representado entre otros por cuatro esqueletos de adultos jóvenes, con un total de once puntas de proyectil clavadas entre los huesos (en un caso 5 de ellas en un mismo esqueleto) que ocasionaron su muerte.

Evento 2: Alrededor de los 7000 años A.P., integrado por entierros simples y múltiples.

Evento 3: 6500 a 6300 años A.P. Representado por lo menos por 8 entierros, siendo varios de ellos niños con un complejo ajuar funerario formado por cuentas de collar de colmillos de cánidos y valvas circulares (Figura 7).

Evento 4: Holoceno tardío. Representado por un solo enterramiento secundario que se encuentra a

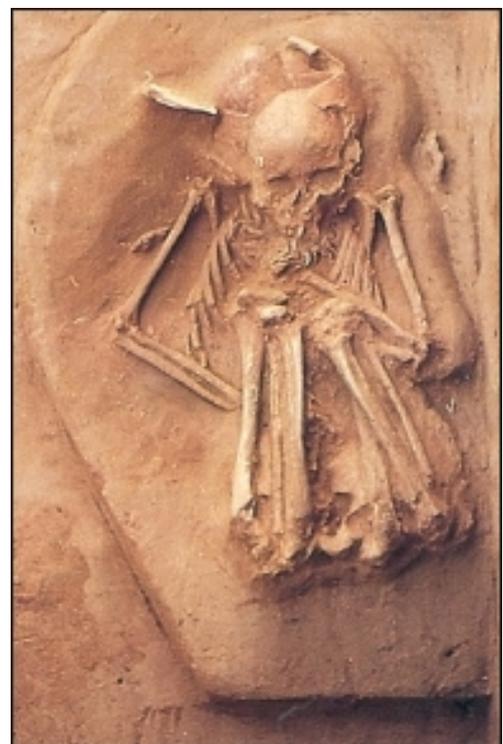


Figura 7: Entierro de un individuo infantil con ajuar de colmillos de cánido. Tomado de Politis (1989).

la base de la Unidad Y. Sin embargo, Barrientos (1997) añade tres inhumaciones de infantes (esqueletos 11, 12 y 13) a este momento.

Como lo expresan Politis y Madrid (2001: 760), “Los tres eventos del Holoceno temprano sugieren la recurrencia en la utilización de un espacio para las prácticas inhumatorias, conducta que hacia el 7800 A.P. no se había registrado. Las implicaciones de esta nueva actitud hacia los muertos indican un cambio en algunos aspectos ideológicos, especialmente con relación a la sacralización del paisaje.”



Figura 8: Entierro n° 2 del sitio Arroyo Seco 2, correspondiente al tercer período de inhumación. Está constituido por dos mujeres adultas (izquierda y centro) y una adolescente (derecha). Tomado de Politis (1989).

Los entierros humanos de Arroyo Seco 2 constituyen la manifestación más temprana en América de la inhumación de los muertos como práctica regular y localizada. Se trata asimismo de una muestra de gran tamaño, y sus entierros tempranos (los correspondientes a las etapas primera, segunda y tercera), constituyen un conjunto bien acotado por fechados radiocarbónicos. La muestra puede considerarse provisionalmente representativa de la primeras poblaciones de la región pampeana argentina, que luego habrían de ser eventualmente reemplazadas por otras poblaciones en el Holoceno medio (Barrientos y Pérez

2002). A su vez, hay razones de carácter arqueológico para considerar a Arroyo Seco como potencial muestra “estándar” para análisis comparativos con muestras posteriores de toda la región pampeana, incluido el Uruguay: El poblamiento de la Pampa argentina se remontaría al límite Pleistoceno-Holoceno, con dataciones de entre 11.200 y 12.300 años (Politis et al. 2004). El poblamiento del actual territorio uruguayo sería contemporáneo y cuenta como rasgo arqueológico común con puntas de tipo “cola de pescado” (Suárez y López 2003). Además, y añadidas a un elemento ergológico común y extendido, están las evidencias de uso en sitios tempranos de la Pampa argentina de materias primas provenientes de Uruguay (Flegenheimer et al. 2001) que indicarían cuando menos una interacción social entre grupos que podría eventualmente implicar flujo génico. Por lo tanto, el análisis de la muestra de Arroyo Seco 2 es una clave potencial para entender el poblamiento de la región pampeana argentina y la continuidad poblacional en la misma, y por extensión, en nuestro propio territorio.

2. Objetivos e hipótesis

2.1. Objetivos

General

Estudiar las características de la población de sitio Arroyo Seco 2 mediante el uso de marcadores del ADN mitocondrial.

Específicos

- a) Estudiar la composición genética de la muestra extraída del sitio Arroyo Seco 2 a través de las frecuencias de los cuatro haplotipos mitocondriales fundadores.
- b) Establecer el grado de similitud/distancia de la muestra con relación a otras muestras antiguas de Sudamérica.
- c) Explorar las relaciones y eventual descendencia de poblaciones indígenas modernas del Cono Sur con respecto a la muestra que, como supuesto exploratorio inicial, se considerará representativa de la composición de las poblaciones prehistóricas de la región pampeana argentina.

2.2. Hipótesis

La *hipótesis nula* H_0 del presente estudio es que la población prehistórica de Arroyo Seco 2 no se relaciona genéticamente con otras poblaciones, antiguas o contemporáneas, de la región sur de Sudamérica. Las hipótesis alternativas, mutuamente excluyentes, son dos:

H_1 : Existe continuidad, desde el Holoceno temprano hasta el pasado reciente, entre la población prehistórica de Arroyo Seco 2 y las poblaciones históricamente registradas en la Pampa y la Patagonia argentina que fueron sometidos por la llamada "conquista del desierto" a fines del siglo XIX.

H_2 : Hay una discontinuidad entre la población prehistórica de Arroyo Seco 2 y las poblaciones históricas del área.

En cualquier caso, las afinidades a testar de la población de Arroyo Seco 2 serían: a) con las poblaciones de Tierra del Fuego, representando un contingente poblacional temprano, separado del que pobló el resto del subcontinente; b) con las poblaciones prehistóricas de la región pampeana en general, incluyendo, sobre la base de los datos arqueológicos disponibles, al actual territorio uruguayo.

3. Materiales y métodos

3.1. Muestra

Hasta el momento, se han recuperado en el sitio Arroyo Seco 2 un total de 40 individuos: 24 adultos (55%), 12 subadultos (27%) y 8 de edad sin determinar (18%). El grupo adulto se compone de 13 individuos masculinos (54%), 10 individuos femeninos (42%) y uno sin determinar (4%). Los tres eventos en los que han sido divididos los entierros del Holoceno temprano y medio se basan en fechados radiocarbónicos hechos directamente sobre esqueletos humanos (Tabla 10), que abarcan un rango cronológico aproximado de 1500 años (ca. 7800 a 6300 años antes del presente).

N° Lab.	Edad C14 A.P.	Muestra
AA-7966	6300 +/- 70*	<i>Homo sapiens</i> – AS03
TO-1505	8980 +/- 100*	<i>Homo sapiens</i> (1) – AS05
LP-55	8560 +/- 320	<i>Homo sapiens</i> (1) – AS05
Beta-80908	7100 +/- 60*	<i>Homo sapiens</i> (1) – AS05
Beta-11251	5250 +/- 110*	<i>Homo sapiens</i> (1) – AS05
LP-186	6560 +/- 60	<i>Homo sapiens</i> – AS06
AA-62517	7043 +/- 82	<i>Homo sapiens</i> – AS07
S1-5482	5910 +/- 55	<i>Homo sapiens</i> – AS07
TO-1503b	7040 +/- 60*	<i>Homo sapiens</i> – AS15
TO-1503a	7000 +/- 80*	<i>Homo sapiens</i> – AS15
CAMS-16170	6970 +/- 60*	<i>Homo sapiens</i> – AS15
NZA-1101	6880 +/- 90	<i>Homo sapiens</i> – AS15
AA-7967	6495 +/- 65	<i>Homo sapiens</i> – AS18
AA-9045	6860 +/- 60*	<i>Homo sapiens</i> – AS19
AA-9046	7800 +/- 115*	<i>Homo sapiens</i> – AS24
Beta-80909	7580 +/- 50*	<i>Homo sapiens</i> – AS26
AA-9048	7615 +/- 90*	<i>Homo sapiens</i> – AS31
AA-19286	7685 +/- 95*	<i>Homo sapiens</i> – AS32
AA-24050	7805 +/- 85*	<i>Homo sapiens</i> – AS36
AA-24051	6940 +/- 75*	<i>Homo sapiens</i> – AS40

Tabla 10: Dataciones radiocarbónicas sobre restos humanos de Arroyo Seco 2. Basado en Politis y Madrid (2001) y Politis (com. pers.). Los individuos de los cuales se tomó muestra para extracción de ADN se indican en negrita. (1) Todos los fechados corresponden al mismo esqueleto humano y son discutibles. (*) Fechados realizados con espectrometría de acelerador de masas (AMS).

Para la asignación temporal de los entierros que no fueron directamente fechados, se procedió a analizar la evidencia contextual disponible (tipo de entierro, ajuar, estructuras, instrumental lítico asociado y profundidad), que arrojó las agrupaciones tentativas resumidas en la tabla 11. Debe agregarse que los individuos AS11, AS12, AS13, AS22 y AS38

presentaban una asociación de elementos que, a partir de la comparación con otros sitios del área, los ubicaría cronológicamente hacia el 2.000 AP, en base a fechados de los sitios La Toma y Laguna Tres reyes 1 (Barrientos 1997). Para aquellos entierros carentes de información contextual suficiente no pudo formularse hipótesis cronológica alguna, quedando los mismos por lo tanto sin asignación temporal.

Entierro	Ind.	Sexo	Edad	Tipo	Material asociado
Unidad cronológica 1: Media ponderada 7.625 ± 40 A.P.					
13	AS20	Masculino	45-49	Primario	Estructura funeraria
17	AS24	Masculino	25-29	Primario	Puntas de proyectil
19	AS26	Femenino	20-25	Primario	Cuentas perforadas de valvas. Estructura funeraria
24	AS31	Masculino	60	Primario	Puntas de proyectil. Estructura funeraria
25	AS32	Femenino	50-54	Primario	Lascas
	AS34	Indet.	1-2	Primario	
27	AS36	Masculino	20-24	Primario	Puntas de proyectil
	AS39	Femenino	15-19	Primario	Puntas de proyectil
Unidad cronológica 2: Media ponderada 6.940 ± 30 A.P.					
10	AS15	Femenino	20-24	Primario	
	AS16	Indet.	2,5-3,5	Primario	
	AS17	Indet.	Indet.	Primario	
12	AS19	Masculino	45-49	Primario	
16	AS23	Masculino	25-29	Primario	
20	AS27	Indet.	4-8	Primario	
31	AS40	Femenino	50-54	Primario	Colmillos de cánidos y cuentas circulares perforadas de valvas
Unidad cronológica 3: Media ponderada 6.495 ± 65 – 6.300 ± 70 A.P.					
1	AS1	Masculino	22-26	Primario	Colmillos de cánidos
	AS2	Femenino	6-10	Primario	Cuentas circulares perforadas de valvas de 3 mm de diámetro
	AS3	Indet.	1-3	Primario	
2	AS4	Femenino	23-26	Primario	
	AS5	Femenino	23-26	Primario	
	AS6	Femenino	9,5-14,5	Primario	
3	AS7	Femenino	30-34	Primario	
4	AS9	Indet.	2-3	Primario	Colmillos de cánido. Cuentas circulares perforadas de valvas
5	AS10	Indet.	Indet.	Primario	Colmillos de cánidos
11	AS18	Indet.	Indet.	Primario	Colmillos de cánidos
21	AS28	Indet.	Indet.	Primario	
29	AS33	Indet.	0,5-1,5	Primario	Colmillos de cánidos

Tabla 11: Entierros e individuos recuperados en el sitio Arroyo Seco 2. La mayoría de las asignaciones a unidades cronológicas son según Barrientos (1997), en tanto que algunas asignaciones (AS22, AS23, AS27 y AS28) corresponden a la aplicación propia de criterios de Barrientos (1997). El individuo AS05 no fue incluido en alguna unidad cronológica en virtud de la disparidad de los fechados (Barrientos 1997), aunque se tomó muestra del mismo para extracción de ADN, al igual que los individuos que aparecen sombreados. (Continúa en página siguiente)

Unidad cronológica 4: Aprox. 2.000 A.P.*					
6	AS11	Indet.	2-3	Primario	Cuentas rectangulares perforadas de valvas
7	AS12	Indet.	Indet.	Primario	Cuentas rectangulares perforadas de valvas
8	AS13	Indet.	Indet.	Primario	Cuentas perforadas de valvas
15	AS22	Masculino	50-54	Primario	Cuentas rectangulares perforadas de valvas
30	AS38	Masculino	Adulto ind.	Secundario	
Entierros sin asignación cronológica					
9	AS14	Masculino	Indet.	Primario	
14	AS21	Indet.	6-10	Primario	
18	AS25	Masculino	35-39	Primario	
22	AS46	Indet.	Indet.	Primario	
23	AS30	Indet.	20-24	Primario	
26	AS35	Indet.	Indet.	Primario	
28	AS8	Masculino	12-16	Primario	
32	AS41	Femenino	50	Primario	
33	AS42	Masculino	Adulto ind.	Secundario	
	AS43	Femenino	Adulto ind.	Secundario	
	AS44	Masculino	Adulto ind.	Secundario	
	AS45	Femenino	Adulto ind.	Secundario	

Tabla 11 (cont.) (*) Cronología basada en fechados de otros sitios.

Las muestras para estudios moleculares fueron tomadas los días 5 y 6 de febrero de 2004 en las instalaciones del grupo de Investigaciones arqueológicas y paleontológicas del Cuaternario Pampeano (INCUAPA) de la Facultad de Ciencias Sociales de la Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA), en Olavarría, con el permiso y bajo la supervisión del Dr. Gustavo Politis, director de la Unidad y responsable del sitio Arroyo Seco 2. Se prefirió en un principio la extracción de piezas dentales, pero las características del sedimento (todas las inhumaciones excepto una corresponden a la unidad “Z” – ver introducción), de color grisáceo y rico en carbonatos, tuvieron por resultado la práctica cementación de las piezas dentales a sus alvéolos, con lo cual la extracción requería la ruptura del alvéolo o del diente. Por lo tanto, se eligieron huesos íntegros de las extremidades (metacarpos, metatarsos o falanges) en ausencia de dientes sueltos (ver resultados, tabla 16).

3.2. Métodos

3.2.1. Extracción de ADN

3.2.1.1. Prevención de contaminación

La extracción de ADN fue realizada entre los meses de marzo y octubre de 2004 en las instalaciones del Laboratorio de ADN Antiguo del Departamento de Antropología Biológica de la Facultad de Humanidades y Ciencias de la Educación (FHCE), Universidad de la República, en Montevideo. Estas instalaciones se encuentran en un área aislada, de aproximadamente 16 m², y cuentan con los implementos necesarios para minimizar la posibilidad de contaminación con ADN moderno: Inyector de aire con filtro, luz UV-C general y localizada e instrumental utilizado exclusivamente para ADN antiguo. Durante el trabajo en el laboratorio se tomaron todas las precauciones prácticas para minimizar la contaminación, incluyendo el uso de mameluco con capucha, cubrezapatos, tapabocas y guantes de látex descartables (Figura 9), encendido de la luz UV 10 minutos antes del ingreso y después de la salida del laboratorio, y limpieza sistemática de las superficies antes y después de cada sesión de trabajo (así como durante su transcurso) con hipoclorito de sodio al 20%. Cada pieza de instrumental fue irradiada con UV localizado, al igual que los reactivos que no contenían material orgánico.

En cuanto a la preparación de la muestra debe señalarse que, a la fecha, se han propuesto varios procedimientos para eliminar cualquier potencial ADN moderno contaminante. Estos procedimientos incluyen la eliminación de la capa exterior de la muestra, la perforación para extraer sólo material interno, el lavado con ácido clorhídrico, etanol, agua oxigenada o hipoclorito de sodio y la inmersión de la pieza en EDTA. En el presente trabajo, las muestras fueron preparadas con una inmersión breve en hipoclorito de sodio al 20% en el caso de dientes y la remoción con hoja de bisturí de la capa exterior de la superficie a trabajar en caso de piezas óseas, seguido de la irradiación de las piezas con luz ultravioleta. Si bien no hay consenso en cuanto al procedimiento más efectivo, los empleados en este trabajo están entre los más mencionados en la literatura (Kemp y Smith 2005).

3.2.1.2. Protocolo de extracción

Posteriormente a la limpieza se extrajo hasta 0,3 g de polvo (dentina en el caso de dientes, hueso compacto proveniente del canal medular en caso de huesos) con un torno de manualidades, cambiando la fresa entre cada pieza a trabajar. Este polvo fue utilizado para la extracción de ADN, utilizando un protocolo adaptado de fenol-cloroformo y proteinasa K con la utilización de una suspensión de sílice. Dicho protocolo, elaborado por D. A. Merriwether (modificado de Baron et al. 1996 y descrito sintéticamente en Merriwether



Figura 9: Indumentaria utilizada en el interior del laboratorio de ADN antiguo.

et al. 2000), se utilizó para la extracción de ADNmt de restos prehistóricos uruguayos (Bertoni et al. 2004; Sans 2004; Sans y Figueiro 2005) (ver Anexo I) y fue puesto a punto antes de comenzar las extracciones.

3.2.2. Análisis del ADN

3.2.2.1. Procedimiento general de PCR

La preparación de las reacciones de PCR se hizo en el laboratorio de ADN antiguo, con instrumental de uso exclusivo para minimizar el riesgo de contaminación. La amplificación por PCR del ADN mitocondrial fue realizada en instalaciones del Laboratorio de Oncología Básica y Biología Molecular (LOBBM) de la Facultad de Medicina. Debe señalarse que no se manipuló ni guardó ADN moderno ni productos de PCR en las instalaciones de ADN antiguo.

Se trabajó con reacciones de 20 µl para los análisis por RFLP y con reacciones de 30 µl para las muestras a secuenciar. Las mezcla de reacción consistió en el *buffer* provisto por el

fabricante (20 mM Tris-HCl, 20 mM KCl y 5 mM (NH₄)₂SO₄), entre 1 y 2 mM de MgCl₂, 0,21 μM de cada cebador, 35 μM de dNTP y 10% (v/v) de dilución de ADN. En virtud de los potenciales factores inhibidores de la reacción de PCR en el extracto (taninos, ácidos húmicos y fúlvicos, productos de Maillard), así como el carácter inhibitorio de la propia suspensión de sílice utilizada en el protocolo de extracción (O'Rourke et al. 2000a), se utilizaron diluciones de 1:20 y 1:50 del extracto original en la reacción de PCR, con lo cual la concentración de extracto en la reacción era 1:200 o 1:500, según la dilución utilizada. La finalidad del uso de diluciones era hallar un equilibrio entre la dilución de ADN y la dilución de los inhibidores; debe observarse que en general, las reacciones utilizando diluciones de 1:20 rindieron producto amplificado con más frecuencia que las diluciones 1:50. En cada tanda de reacciones de empleó tres blancos conteniendo agua, y una muestra moderna (generalmente el ADN del investigador) como control positivo de amplificación y negativo a nivel de análisis por RFLP. El agregado de este ADN moderno se hacía invariablemente fuera del laboratorio de ADN antiguo.

Las reacciones de PCR fueron efectuadas con *Taq* polimerasa diseñada para activarse a temperaturas mayores de 90°C (*Platinum Taq*, Invitrogen corp., o *Hotstart Taq*, Fermentas corp.). Si bien su fidelidad es equivalente a la *Taq* polimerasa común, el hecho de que inicie su actividad en forma brusca y con el ADN a replicar desnaturalizado minimiza la formación inespecífica de productos de partida, como ser los dímeros de cebadores. El uso de esta variedad de polimerasas se emplea en general en la actualidad en sustitución de las reacciones de “inicio en caliente” basadas en barreras de cera entre la mezcla de reacción y la solución de ADN. En un principio se trabajó a base de 0,4 unidades de *Taq* por reacción de 20 μl, pero en los casos de mayor dificultad se usó 0,5 y 0,7 unidades de *Taq* por reacción.

Las reacciones de PCR consistieron en un paso de desnaturalización inicial y activación de la polimerasa a 94°C por 10 minutos, seguido de 45 ciclos de desnaturalización (94°C, 40 segundos), hibridización (40 segundos) y extensión (72°C, 40 segundos). Para aumentar la eficiencia de la reacción, se empleó “*touchdown* PCR” consistente en comenzar la reacción con temperaturas de hibridización 3°C mayores a las temperaturas de desnaturalización del cebador más rico en GC, a fin de obtener extensión únicamente sobre blancos específicos. En ciclos subsiguientes se disminuye la temperatura a razón de 0,5 °C por ciclo (Kaestle y Horsburgh 2002). En el presente trabajo se comenzó con temperaturas de 1°C por encima de la T_m menor del par de cebadores, bajando la temperatura por ciclo hasta un total final de 3°C. Finalmente, se llegó a utilizar temperaturas 4°C menores a la del menor

cebador, disminuyendo un total de 6,75°C la temperatura final. Si bien esto podría producir amplificaciones inespecíficas, éstas se observaron sólo en forma excepcional en poliacrilamida.

3.2.2.2. Control de éxito de las extracciones

El control del éxito de las extracciones se realizó mediante la amplificación por PCR del segmento 13235-13334 del genoma mitocondrial, en cuyo tramo se encuentra el sitio de restricción correspondiente al haplogrupo C. Debe observarse que el segmento de 100 pb mencionado para el control del éxito de la extracción fue elegido por ser el más corto en el juego de cebadores utilizado; por consiguiente, la no amplificación de dicho segmento en dos ocasiones sucesivas utilizando ambas diluciones fue considerado fracaso de la extracción – o para ser más específicos, la inexistencia en el extracto de segmentos de ADN mayores de 100 pb.

3.2.2.3. RFLP-PCR y análisis de secuencias

Para el análisis de RFLP-PCR se empleó inicialmente los pares de cebadores A, B y D tomados de Moraga et al. (2000) y el par C, tomado de Parr et al. (1996) (Tabla 11), que amplifican las regiones que poseen los sitios de restricción típicos de los haplogrupos homónimos. Posteriormente, para reducir los tamaños de los amplicones y aumentar la eficiencia de amplificación, se agregaron los cebadores A2 (Handt et al. 1996), B2 (Parr et al. 1996) y el par D2 (Handt et al. 1996). En casos en los que el RFLP-PCR no fue concluyente para los cuatro marcadores utilizados se amplificó y secuenció el segmento 16288-16355 de la HVR I utilizando el juego de cebadores M4 (dado que la amplificación de la HVR I en su totalidad a partir de ADN antiguo es un evento prácticamente imposible, se diseñaron los pares de cebadores M0-M5 para amplificar la HVR I en un máximo de seis segmentos solapantes; ver tabla 12).

Cebador	Posición (5'-3')	Secuencia	Largo amplicón	Largo fragmentos
RFLP				
A	L: 583-604 H: 727-708	GTA GCT TAC CTC CTC AAA GCA A AGG GTG AAC TCA CTG GAA CG	145 pb	81 pb 64 pb
B	L: 8195-8214 H: 8316-8294	CAC AGT TTC ATG CCC ATC GT ATG CTA AGT TAG CTT TAC AGT GG	122/113 pb	No aplica
C	L: 13235-13256 H: 13334-13313	AAT CGT AGC CTT CTC CAC TTC A GGC GTG GGT ACA GAT GTG CAG G	100 pb	25 pb 75 pb
D	L: 5099-5120 H: 5277-5255	CTA ACT ACT ACC GCA TTC CTA AAT TCT TCG ATA ATG GCC CAT TT	179 pb	78 pb 101 pb
A2	L: 610-635	TGA AAA TGT TTA GAC GGC CTC ACA TC	118pb (con A _H)	54 pb 64 pb
B2	L: 8249-8267	GGG CCC GTA TTT ACC CTA T	68/59 pb (con B _H)	No aplica
D2	L: 5146-5165 H:5210-5189	CGA CCC TAC TAC TAT CTC GC GGG TGG ATG GAA TTA AGG GTG T	65 pb	31 pb 34 pb
HVR I				
M0	F: 16015-16035 R: 16114-16096	TAT TCT CTG TTC TTT CAT GGG GTT GTT GGC GAT ACA TAA A	100 pb	No aplica
M1	F: 16036-16055 R: 16158-16139	GAA GCA GAT TTG GGT ACC AC TAC TAC AGG TGG TCA AGT AT	123 pb	No aplica
M2	F: 16112-16131 R: 16237-16218	CAC CAT GAA TAT TGT ACG GT TGT GTG ATA GTT GAG GGT TG	126 pb	No aplica
M3	F: 16192-16209 R: 16322-16303	CCA TGC TTA CAA GCA AGT TGG CTT TAT GTA CTA TGT AC	131pb	No aplica
M4	F: 16268-16287 R: 16375-16356	CAC TAG GAT ACC AAC AAA CC GTC ATC CAT GGG GAC GAG AA	108 pb	No aplica
M5	F: 16327-16347 R: 16429-16410	CGT ACA TAG CAC ATT ACA GT GCG GGA TAT TGA TTT CAC GG	103 pb	No aplica

Tabla 12: Cebadores utilizados para la amplificación de segmentos informativos del ADNmt.

La amplificación de los fragmentos fue controlada por electroforesis en gel de agarosa al 2% con bromuro de etidio. La visualización de los fragmentos tras su digestión con las enzimas de restricción correspondientes y el control de presencia de la delección de 9 pb diagnóstica del haplogrupo B se realizaron mediante electroforesis en gel de poliacrilamida al 6% con tinción por nitrato de plata. La secuenciación de las muestras seleccionadas para tal fin fue llevada a cabo en el Centro Técnico de Análisis Genéticos (CTAG) de Facultad de Ciencias. Las secuencias fueron alineadas utilizando el programa GeneDoc (Nicholas et al. 1997) y comparadas con la secuencia de referencia de Cambridge (revisión hecha por Andrews et al. 1999 de la secuencia original de Anderson et al. 1981). Eventuales ambigüedades fueron resueltas mediante lectura directa del ferograma.

3.2.3. Análisis de variabilidad y afinidad con otras poblaciones

Los datos de pertenencia a haplogrupos alcanzados mediante la combinación de ambas aproximaciones (RFLP y mutaciones diagnósticas en HVR I) fueron analizados con las siguientes estrategias:

1. Análisis Molecular de la Varianza (AMOVA; Excoffier et al 1992) utilizando las frecuencias de los haplogrupos, a fin de comparar la muestra analizada con poblaciones antiguas y modernas de América del Sur (Tablas 13 y 14). Esto se efectuó mediante el uso del programa ARLEQUIN (Schneider et al. 2000).

2. Cálculo de distancias a partir de las frecuencias de haplogrupos de la muestra, en comparación con las muestras antiguas y modernas consideradas. Estas distancias fueron interpretadas mediante los siguientes procedimientos:

- Elaboración de dendrogramas por el procedimiento de ligamiento entre vecinos o *neighbor-joining* (Saitou y Nei 1987). Los dendrogramas fueron evaluados mediante dos clases de análisis por *bootstrap*, uno por remuestreo de individuos dentro de cada población, y otro por remuestreo de *loci* dentro de cada población, asumiendo cada sitio de restricción como un *locus*. El primer procedimiento se efectuó utilizando los paquetes *ADE4*, *ape* y *cluster* del entorno R versión 2.3.1 (R Development Core Team 2006). El segundo procedimiento de *bootstrap* se efectuó utilizando el paquete DISPAN (Ota 1993) y los paquetes *ADE4*, *ape* y *cluster* del entorno R versión 2.3.1 (R Development Core Team 2006). La elaboración de los árboles consenso de los *bootstrap* elaborados con el entorno R se realizó empleando PHYLIP versión 3.6 (Felsenstein 2004).
- Análisis de coordenadas principales, utilizando el paquete *ADE4* del entorno R versión 2.3.1 (R Development Core Team 2006).
- Test de Mantel para analizar la correlación entre las matrices de distancias genéticas y las distancias geográficas entre las muestras, utilizando el paquete *ADE4* del entorno R versión 2.3.1 (R Development Core Team 2006).

La elección de las poblaciones modernas para comparación (Tabla 13; Figura 10a) se basó especialmente en su proximidad geográfica al sitio, en la medida en que la resolución utilizada en el análisis y la antigüedad de la muestra analizada trae las siguientes limitantes:

- Es prudente analizar los datos disponibles, en forma de frecuencias génicas de un solo *locus*, utilizando análisis comparativos lo más libres posible de asunciones evolutivas. En virtud de ello, la selección de muestras alejadas geográficamente puede resultar en semejanzas basadas en fenómenos evolutivos causalmente coincidentes. La utilización de muestras geográficamente próximas permite dar mayor asidero factual a las eventuales semejanzas que surjan en el análisis.
- La muestra analizada dista 7000 años en el tiempo de las muestras modernas comparativas. En vista de ello, hacer inferencias de fenómenos microevolutivos basados en eventuales aislamientos o contactos es, cuando menos, arriesgado. En lo que respecta a la continuidad entre las poblaciones prehistóricas y las actuales de la región considerada, varios trabajos realizados en el sudoeste norteamericano (Carlyle et al. 2000; O'Rourke et al. 2000b; Kaestle y Smith 2001) indicarían una relativa continuidad en las frecuencias de haplogrupos mayores entre las poblaciones antiguas y las actuales de una región dada. Sin embargo, la historia poscontacto de la región pampeana está caracterizada por la adopción de una alta movilidad en los grupos indígenas locales, y en tiempos recientes, por el exterminio y desplazamiento forzado de las mismas. Por ello, la continuidad observada en grupos norteamericanos está aún por probarse en la región pampeana.

País / Región	Grupo	N	A	B	C	D	Fuente
Bolivia	Aymara	172	6%	67%	12%	14%	Merriwether et al. 1995
Chile	Huilliche	80	4%	29%	19%	49%	
	Atacameño	50	12%	72%	10%	6%	Goicoechea et al. 2001
Chaco	Chorote	20	15%	40%	30%	15%	
	Mataco	70	6%	64%	3%	27%	
	Pilaga	40	5%	38%	28%	30%	
	Toba	54	20%	43%	6%	31%	Demarchi et al 2001b
Argentina	Mapuche	129	14%	37%	22%	26%	Goicoechea et al. 2001 Ginther et al. 1993
	Tehuelche	29	0%	21%	24%	55%	Goicoechea et al. 2002
Sur Chile	Mapuche	111	0%	7%	44%	49%	Moraga et al. 2000
	Pehuenche	105	3%	11%	41%	46%	
Argentina	Guaraní	106	45%	23%	5%	27%	Diez 2004; Altuna 2003
Brasil	Guaraní	217	88%	0%	5%	7%	Marrero 2003

Tabla 13: Muestras modernas utilizadas en este trabajo.

En el caso de las muestras antiguas, se entendió “antiguo” en forma amplia como toda muestra generada a partir de restos arqueológicos. Dichas muestras abarcan en todos los casos un rango temporal en el orden del millar o más de años, y en ningún caso son contemporáneas a la muestra de Arroyo Seco 2. Las antigüedades de las muestras son de entre 9000 y 100 años A.P. Puesto que la información disponible respecto a poblaciones antiguas es más escasa, se incluyó todas las muestras disponibles en América del Sur, más las muestras de las Antillas analizadas por Lalueza-Fox et al. (2001, 2003). Éstas últimas fueron incluidas ya que los grupos que poblaron las Antillas en el entorno del séptimo milenio A.P. provendrían del norte de América del Sur (Lalueza-Fox et al. 2001, 2003) y podrían ofrecer una imagen aproximada de la composición de las poblaciones de dicha región.

Región	Sitio/ Grupo	Cronol. Aprox. (Años A.P.)	N	A	B	C	D	Fuente
Norte Chile		3900 – 500	76	34%	37%	25%	4%	Moraga et al. 2001, 2005
Amazonia		4000 – 500	11	45%	9%	36%	9%	Ribeiro-dos Santos et al. 1996
Tierra del Fuego		100	68	0%	0%	49%	51%	Lalueza et al. 1997; García-Bour et al. 2004
Patagonia		5000 – 100	15	0%	0%	27%	73%	Lalueza et al. 1997
NW argentino	Pampa Grande	1000 – 1500	16	6%	50%	0%	44%	Dejean et al. 2005
Este uruguayo		2000 – 400	7	14%	43%	43%	0%	Bertoni et al. 2004; Sans y Figueiro 2005
Colombia		1500 – 100	6	50%	17%	33%	0%	Monsalve et al. 1996
República Dominicana	Taíno	1300 – 300	24	0%	0%	75%	25%	Lalueza-Fox et al. 2001
Cuba	Ciboney	4700 – 1600	15	7%	0%	60%	33%	Lalueza-Fox et al. 2003

Tabla 14: Muestras antiguas utilizadas en este trabajo.

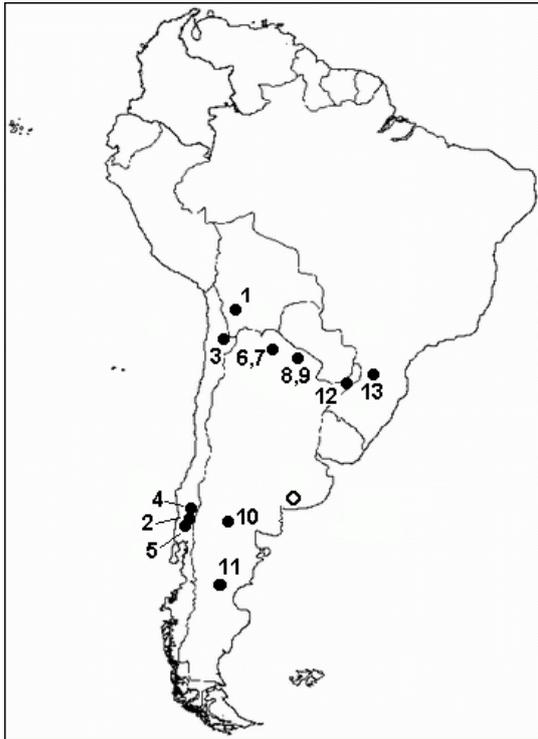


Figura 10a: Ubicación de las poblaciones modernas utilizadas para el análisis comparativo. 1 – Aymara; 2 – Huilliche; 3 – Atacameño; 4 – Pehuenche; 5 – Mapuche (CH); 6 – Mataco; 7 – Chorote; 8 – Pilaga; 9 – Toba; 10 – Mapuche (ARG); 11 – Tehuelche; 12 – Guaraní (ARG); 13 - Guaraní (BR). La ubicación del sitio Arroyo Seco 2 se señala con un círculo vacío.



Figura 10b: Ubicación de los sitios de origen de las muestras antiguas utilizadas para el análisis comparativo. 1 – Norte Chile; 2 – Amazonas; 3 – Tierra del Fuego; 4 – Patagonia; 5 – Noroeste argentino; 6 – Este uruguayo; 7 – Colombia; 8 – República Dominicana; 9 – Cuba. La ubicación del sitio Arroyo Seco 2 se señala con un círculo vacío.

3.2.3.1. AMOVA

El análisis de la varianza molecular (AMOVA) puede definirse como una herramienta relativamente versátil para el análisis de estructura genética, y funciona en un sentido análogo al análisis de la varianza. Incorpora varios niveles de análisis, pudiéndose en función de ello analizar la correspondencia de los datos a mano con diversos panoramas hipotéticos. Adaptado a los datos disponibles en este trabajo, el procedimiento, tal como fue definido por Excoffier et al. (1992) se puede sintetizar del siguiente modo:

Dada una matriz de distancias entre poblaciones D^2 , podemos subdividir la matriz en una serie de submatrices correspondiente a subdivisiones particulares, establecidas por criterios no genéticos como ser geográficos, lingüísticos o ambientales:

$$D^2 = \begin{bmatrix} [D_{11}^2] & [D_{12}^2] & \dots & [D_{1I}^2] \\ [D_{21}^2] & [D_{22}^2] & \dots & [D_{2I}^2] \\ \dots & \dots & \dots & \dots \\ [D_{I1}^2] & \dots & \dots & [D_{II}^2] \end{bmatrix}$$

donde los elementos de las submatrices diagonales contienen distancias al cuadrado entre poblaciones de la misma agrupación, en tanto que los elementos fuera de la diagonal contienen distancias entre poblaciones de la subdivisión i y la subdivisión i' . Considerando que una suma de cuadrados convencional puede escribirse igualmente como la suma de las diferencias al cuadrado entre todos los pares de datos, se puede construir un análisis jerárquico de la varianza molecular directamente de la matriz de distancias. Se emplea un modelo lineal siguiendo el patrón descrito inicialmente por Cockerham (1973):

$$p_{jig} = \mu + a_g + b_{jg} + c_{jig}$$

donde p_{jig} indica el individuo j de la población i en el grupo g , y μ es la p_{jig} esperada promedio de todo el análisis. Los efectos son a para los grupos, b para poblaciones y c para individuos dentro de poblaciones, efectos que se asume son aditivos, aleatorios, no correlacionados, con los componentes de varianza asociados σ_a^2 , σ_b^2 y σ_c^2 , respectivamente. Sobre la base de la descomposición convencional, se puede establecer la siguiente partición de la suma de desviaciones al cuadrado (SSD):

$$SSD_{total} = SSD_{interestratos} + SSD_{intraestratos}$$

lo cual nos coloca dentro del marco del análisis de la varianza tradicional. En este caso, Excoffier et al. (1992) particionan el SSD_{total} en tres componentes: $SSD_{intrapoblacional}$, $SSD_{interpoblacional/intragrupo}$ y $SSD_{intergrupo}$. Sobre la base de los componentes de varianza calculados, y siguiendo a Cockerham (1973), se pueden calcular índices Φ de correlación haplotípica, tales que

$$\sigma_c^2 = (1 - \Phi_{ST})\sigma^2$$

$$\sigma_b^2 = (\Phi_{ST} - \Phi_{CT})\sigma^2$$

$$\sigma_a^2 = \Phi_{CT}\sigma^2$$

donde $\sigma^2 = \sigma_a^2 + \sigma_b^2 + \sigma_c^2$. Φ_{ST} es la correlación de los haplotipos al azar dentro de las poblaciones respecto a pares aleatorios de haplotipos extraídos del total de las poblaciones. Φ_{CT} es la correlación de haplotipos extraídos al azar del grupo de poblaciones considerado respecto al total, y Φ_{SC} es la correlación de la diversidad intrapoblacional respecto a la del grupo de poblaciones. Siguiendo la analogía con los índices F, se puede expresar Φ_{ST} , Φ_{SC} y Φ_{CT} como:

$$\Phi_{ST} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma^2}, \Phi_{CT} = \frac{\sigma_a^2}{\sigma^2}, \Phi_{SC} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_c^2}$$

Desde luego, los requisitos relacionados con los índices F (muestreo aleatorio, acción únicamente de la deriva y ausencia de migración) se cumplen raramente o nunca en poblaciones naturales. Del mismo modo, la interpretación de los índices Φ , análogos a aquellos, debe hacerse con las mismas precauciones. La significación de cada uno de los índices se establece mediante una distribución empírica obtenida reasignando al azar individuos a diferentes poblaciones o a diferentes poblaciones dentro de una agrupación dada para testar Φ_{ST} y Φ_{SC} , y reasignando al azar poblaciones a diferentes agrupaciones para testar Φ_{CT} .

Para analizar las posibles afinidades de la población de Arroyo Seco 2 utilizando el AMOVA, el procedimiento consistió en asignar la muestra a cada una de las agrupaciones

definidas previamente para las poblaciones modernas y antiguas analizadas (Tablas 15a y 15b, respectivamente). Estas agrupaciones fueron analizadas mediante AMOVA a fin de testar la significación de las subdivisiones.

Los criterios para agrupar las poblaciones antiguas y modernas en subdivisiones difirieron ligeramente. En el caso de las poblaciones modernas, el criterio preponderante fue el geográfico, agrupando las poblaciones de acuerdo a su localización geográfica actual. A grandes rasgos, el criterio geográfico coincide con un criterio macrolingüístico, con un par de excepciones:

- a) Los grupos Mapuche (chilenos y argentinos), Tehuelche, Huilliche y Pehuenche, pertenecientes al tronco lingüístico andino (todos los grupos pertenecen a la familia lingüística Araucana excepto los Tehuelche, cuya lengua original es de la familia Patagona) fueron separadas por país, en la medida en que se consideró a la cordillera de los Andes como una barrera geográfica que en gran medida podría haber dado lugar a algún grado de diferenciación entre los grupos de uno y otro lado.
- b) Se agrupó por su proximidad a los grupos Atacameño y Aymara, aún cuando desde el punto de vista lingüístico la lengua del primero (Kunza) no está clasificado (algunos autores lo ubican en la familia macro-Chibcha – *cf.* Gordon 2005) y el segundo pertenece a la familia Quechua-Aymara. Un intento de subdivisión separando ambos

Subdivisión	Poblaciones
1	Aymara Atacameño
2	Guaraní (ARG) Guaraní (BR)
3	Chorote Mataco Pilaga Toba
4	Mapuche (ARG) Tehuelche
5	Mapuche (CH) Pehuenche Huilliche

Tabla 15a: Agrupaciones de poblaciones modernas para el AMOVA con la muestra de Arroyo Seco 2. Para esta subdivisión, $\Phi_{SC} = 0,0881$; $\Phi_{ST} = 0,3177$; $\Phi_{CT} = 0,2518$; $p(\Phi_{CT}) < ,000$

Subdivisión	Poblaciones
1	Amazonia Norte Chile Colombia
2	Este uruguayo
3	Noroeste argentino
4	Tierra del Fuego Patagonia
5	Taíno Ciboney

Tabla 15b: Agrupaciones de poblaciones antiguas para el AMOVA con la muestra de Arroyo Seco 2. Para esta subdivisión, $\Phi_{SC} = 0,0117$; $\Phi_{ST} = 0,2559$; $\Phi_{CT} = 0,2472$; $p(\Phi_{CT}) < ,004$

grupos mostró una variación entre agrupaciones menor, aunque significativa.

La formación de agrupaciones poblacionales para el AMOVA con muestras antiguas resultó, sin embargo, problemática, contribuyendo a ello sin duda cuestiones de tamaño muestral y rango temporal. Se comenzó con una subdivisión de 5 *clusters*, de carácter predominantemente geográfico, sobre la cual se trabajó luego a fin de obtener una agrupación que

maximizara la variación entre subdivisión. Quizá en forma en principio sorprendente, la subdivisión con un mayor grado de reducción de la variabilidad intragrupal y maximización de la variabilidad intergrupala es la que agrupa las muestras amazónicas (Ribeiro-dos-Santos et al. 1996) con las de los valles de Lluta, Camarones y Azapa de la región norte de Chile (Moraga et al. 2001, 2005) y las momias colombianas (Monsalve et al. 1996); sin embargo, diversos autores avalan la coherencia de esta reagrupación al plantear la hipótesis de un origen amazónico de las poblaciones andinas (Rothhammer y Silva 1989, 1992; Moraga et al. 2001), lo cual será discutido posteriormente .

3.2 3.2. *Análisis de distancias*

En virtud de las consideraciones mencionadas arriba respecto al carácter de las muestras y su no contemporaneidad con la muestra de Arroyo Seco 2, las distancias empleadas en este trabajo son distancias comúnmente empleadas con fines *clasificatorios* al estilo de la taxonomía numérica. Están exentas por lo tanto de cualquier asunción de mecanismos evolutivos implícitos. Del mismo modo debe agregarse que las técnicas empleadas para interpretar las distancias calculadas (dendrogramas, análisis de coordenadas principales) son en este caso herramientas heurísticas. Los dendrogramas no pretenden ser árboles filogenéticos de las poblaciones utilizadas ni del *locus* analizado.

Fueron utilizados dos índices de distancias entre Arroyo Seco 2 y las poblaciones de comparación: La distancia de D_R de Rogers y la distancia D_A de Nei. La distancia de Rogers (1972) es una modificación de la distancia geométrica de modo que el rango de valores de la misma se extienda entre 0 y 1 (y no entre 0 y $\sqrt{2}$, como es el caso cuando se emplea la distancia geométrica con frecuencias génicas). La distancia de Rogers se calcula

$$D_R = \frac{1}{r} \sum_{j=1}^r \sqrt{\sum_{i=1}^{m_j} (x_{ij} - y_{ij})^2} / 2$$

siendo r el número de *loci* estudiados, y x_{ij} e y_{ij} las frecuencias del alelo i en el locus j . Esta medida es una distancia *métrica*, de modo que se satisface la desigualdad triangular, esto es, dadas las distancias d_{ij} , d_{jk} y d_{ik} entre tres poblaciones, cualquiera de estas distancias será menor que la suma y mayor que la diferencia entre las otras dos. La distancia de Rogers no es

proporcional al tiempo evolutivo, y tiende a subestimar la distancia entre poblaciones cuando éstas no comparten alelos.

La distancia D_A de Nei et al. (1983) es una modificación de la distancia de cuerda de Cavalli-Sforza y Edwards (1967) a fin de reducir la influencia de alelos poco frecuentes en la misma. Se calcula

$$D_A = \frac{1}{r} \sum_{j=1}^r (1 - \sum_{i=1}^{m_j} \sqrt{x_{ij} y_{ij}})$$

D_A toma valores entre 0 y 1, siendo $D_A = 1$ cuando las poblaciones comparadas no comparten alelos. Esta distancia no satisface la desigualdad triangular, por lo cual no sería formalmente admisible – como sí lo es la distancia de Rogers – para su empleo en el análisis de coordenadas principales. No obstante sometimos ambas distancias al mismo tratamiento.

3.2.3.2.1. Elaboración y análisis por *bootstrap* de los dendrogramas

Los dendrogramas fueron elaborados por el método de ligamiento entre vecinos o *neighbor-joining* (NJ) (Saitou y Nei 1987). Estrictamente, este tipo de árbol carece de raíz y por lo tanto el formato de presentación más adecuado es el de un diagrama radiado. En este caso se optó, por facilidad de lectura, por un formato clásico.

El *bootstrap* puede definirse como un conjunto de métodos de cálculo de errores estándar e intervalos de confianza de índices a través de procedimientos de remuestreo (Efron y Tibshirani 1993). El *bootstrap* emplea remuestreo repetido con reposición de elementos de la muestra analizada y el subsiguiente recálculo del índice para luego elaborar una distribución empírica de los valores que tomaría el índice en condiciones de muestreo aleatorio. El procedimiento de remuestreo distingue al *bootstrap* del *jackknife* (consistente en eliminar en forma sucesiva un elemento de la muestra) y de la aproximación por permutación (consistente en elaborar la distribución empírica de un índice comparativo de dos o más muestras reasignando al azar elementos entre las mismas, siendo por ende un remuestreo sin reposición). Los procedimientos de remuestreo son robustos en la medida en que no dependen de tantos requisitos de la muestra como las pruebas paramétricas, pero no eluden por ello el requisito básico de representatividad de la muestra.

A nivel de análisis filogenéticos, se emplea remuestreo de sitios (en el caso de análisis

de secuencias) o *loci* (en el caso de análisis de frecuencias génicas) para testar la robustez de la topología de árboles. Al trabajar con un genoma haploide y no recombinante como el mitocondrial, en teoría se está trabajando con un solo *locus*, por lo cual no sería posible realizar un *bootstrap*. Es práctica frecuente sin embargo tomar cada sitio polimórfico como un *locus* (e.g. García et al. 2004) de modo que en el caso que nos ocupa, cada haplogrupo sería un *locus* bialélico, y cada alelo la presencia o ausencia del sitio de restricción o delección diagnóstica. Esta aproximación se utilizó en el presente trabajo, empleando el paquete DISPAN para la distancia D_A , y el entorno R para la distancia D_R . Sin embargo, también se empleó otro procedimiento, consistente en simular el remuestreo de *individuos* dentro de cada población. Este *bootstrap*, al igual que el remuestreo de sitios para la distancia D_R , se realizó mediante secuencias de líneas de comandos en el entorno R versión 2.3.1 (Anexos IIIa y IIIb).

En los análisis filogenéticos basados en distancias, una vez realizado el remuestreo se recalculan las distancias del juego de datos y se reelabora el árbol, extrayéndose luego de repetidos remuestreos la proporción de veces que se repitió cada nodo del árbol original. Otro procedimiento consiste en la elaboración de árboles no a partir de los valores originales sino de los valores remuestreados, con la elaboración de un árbol “consenso” final con una topología ajustada a un criterio de consistencia más o menos estricto. El paquete PHYLIP empleado para la elaboración de árboles consenso en el presente trabajo emplea tres criterios posibles: Consenso estricto (sólo se reproducen los nodos con una consistencia del 100%), mayoritario (sólo se reproducen los nodos con una consistencia mayor al 50%) y mayoritario extendido (la configuración inicial se construye a partir de los nodos mayoritarios, agregándose luego el resto de los nodos en orden decreciente de consistencia siempre y cuando no contradigan bifurcaciones de mayor orden de consistencia). Mientras que en el paquete DISPAN el *bootstrap* muestra la consistencia de los nodos del árbol original, el *output* del entorno R se trató mediante la aproximación de árboles consenso. Los 1000 dendrogramas resultantes del *bootstrap* de loci e individuos para la distancia D_R y de individuos para la distancia D_A fueron analizados mediante el programa CONSENSE del paquete PHYLIP, utilizando el criterio de mayoría extendida a fin de obtener una posición en el árbol para todas las poblaciones analizadas.

3.2.3.2.2. Análisis de coordenadas principales

El Análisis de coordenadas principales (PCoA) consiste en hallar los componentes principales de una matriz de similitudes entre n objetos, sin tener necesariamente los valores originales para ninguna de las p variables utilizadas en el análisis de los n objetos. Los fundamentos de ello se pueden resumir en que la ecuación

$$Ca_i = \lambda_i a_i \quad (1)$$

aplicable a la matriz de covarianzas, siendo a_i el vector propio i correspondiente al i -ésimo valor propio de la matriz de covarianzas C , puede expresarse como

$$X' X a_i / (n-1) = \lambda_i a_i \quad (2)$$

siendo X la matriz de datos originales estandarizados, de modo que el producto de la transpuesta de esta matriz por sí misma se convierte en el numerador de una matriz de covarianzas. A su vez, una matriz S , producto de la multiplicación de la matriz de datos originales y su transpuesta, puede considerarse una matriz de similitudes (*cf.* Manly 1994). Sobre esta base, se sigue que la multiplicación de la ecuación 2 por X nos da

$$S z_i = (n-1) \lambda_i z_i \quad (3)$$

donde $z_i = X a_i$, consistiendo en un vector de largo n que contiene los valores de Z_i para los objetos a comparar. A su vez, si Z es una matriz conteniendo los valores de los p componentes principales para los n objetos (siendo p el número de variables), se sigue que el i -ésimo valor propio de la matriz de similitudes XX' es $(n-1)\lambda_i$ y el vector propio correspondiente nos da los valores del i -ésimo componente principal para los n objetos (Manly 1994). Estos valores pueden ser graficados en un sistema de ejes, representando cada eje un componente a elección.

Las diferencias fundamentales entre el PCoA y el análisis de componentes principales (ACP o PCA) consiste en la escala dada a los componentes, y en el hecho de que el PCoA puede aplicarse a una gama más amplia de datos, si bien esta amplitud no está exenta de complicaciones: El ACP fue desarrollado para datos recogidos en un nivel de medición de intervalo, en tanto que el PCoA, al basarse en una matriz de similitudes, no tiene teóricamente más restricciones que la de que el índice de similitud utilizado se adapte al tipo de datos. No obstante, la matriz de similitudes debe cumplir con ciertas características, y existe la posibilidad de que surjan valores propios negativos (Manly 1994).

En el presente trabajo, las matrices de distancias D_A y D_R tanto de poblaciones

antiguas como modernas fueron sometidos a análisis de coordenadas principales empleando el paquete ADE4 del entorno R. Los valores de cada población para los primeros dos vectores propios fueron graficados en un sistema de ejes a fin de tener una representación en el plano de las relaciones calculadas entre Arroyo Seco 2 y las poblaciones de referencia.

4. Resultados

4.1. Extracción de ADN

De las 23 muestras tomadas, 15 rindieron ADN, de acuerdo a la amplificación control inicial realizada con los cebadores del haplogrupo C (100 pb) (Tabla 16). Sin embargo, a la fecha, se ha podido establecer el haplogrupo de 8 muestras, en tanto que 4 extracciones fueron demostradas y 2 presuntamente contaminadas por ADN del investigador y una muestra permanece a la fecha en una situación indeterminada. Todo esto nos da una tasa de éxito neta de 35%. Tomando los resultados de amplificación piloto, es decir, la presencia o no de ADN humano, hallamos que el tipo de material de base (diente o hueso) influye en forma escasa o nula en el éxito de la extracción, con un *odds-ratio* de 0,68 para la tasa de éxito en diente (3 en 5) respecto a hueso (6 en 12). Sorprendentemente, la masa media de los casos exitosos (0,18 g) es levemente menor que la de los casos fallidos (0,21 g). Esta diferencia no es, sin embargo, estadísticamente significativa ($t = 0,7974$, $df = 14,82$, $p = 0,438$).

Otros posibles factores analizados fueron antigüedad (empleando las unidades cronológicas expuestas en la tabla 11, y excluyendo los individuos carentes de asignación cronológica), sexo y edad del individuo analizado (ya que se ha afirmado que en adultos, la descalcificación de los huesos del individuo con la edad podría influenciar su preservación tras el entierro – Walker et al. 1988). Ninguno de estos factores mostró una asociación significativa con el éxito o fracaso de la extracción. (Antigüedad del entierro: $\chi^2 = 1,0286$, $df = 2$, $p = 0,598$; sexo del individuo: *odds-ratio* = 0,189, $p = 0,282$ por test exacto de Fisher; edad del individuo: test de Kolmogorov-Smirnov, $D = 0,286$, $p = 0,938$)

4.2. Análisis del ADN

4.2.1. Determinación del haplogrupo de los individuos

Las amplificaciones para determinar los haplogrupos daban, en muchos casos, resultado inmediato: En ciertos casos, fue necesario más de un intento para amplificar una muestra (Tabla 16). No fue constatada influencia alguna de la antigüedad de las muestras ni el

largo de los amplicones (máximo 179 pb, mínimo 65 pb) en la eficiencia de la amplificación de los marcadores.

En algunos casos se constató amplificación inespecífica: En las PCR destinadas a amplificar el segmento portador del sitio polimórfico característico del haplogrupo C, y menos frecuentemente en las correspondientes a los haplogrupos A y D, se observaron amplicones de peso aproximado pero menor al esperado y, en contadas

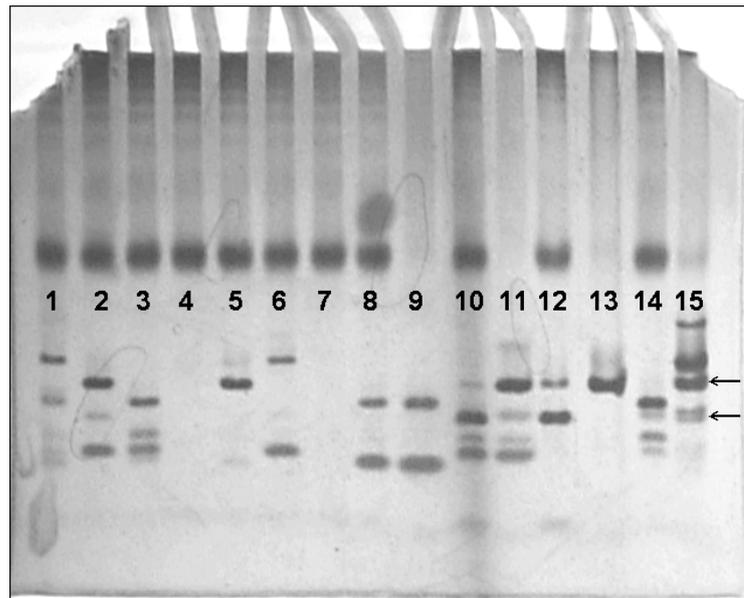


Figura 11: Gel de poliacrilamida de productos de digestión para diagnosticar el haplogrupo C. Los carriles 12 y 13 corresponden a los controles negativo y positivo (sitio de restricción presente y ausente, respectivamente). El carril 15 corresponde al marcador de peso molecular, con fragmentos de 179, 122, 100 pb (flecha), 81 y 75 pb (flecha). Obsérvese la presencia en la mayoría de los carriles de fragmentos mayores de 100 pb, entre 100 y 75 pb y menores a 75 pb.

ocasiones, de peso mayor (Figura 11). Se pretende en un futuro secuenciar dichos amplicones a fin de identificar su naturaleza exacta.

Los resultados y determinaciones se sintetizan en la tabla 16. El detalle del resultado de las extracciones y determinaciones de haplogrupo de cada una de las muestras utilizadas se presenta en el anexo II.

En los casos AS02, AS05 y AS43, se produjeron resultados contradictorios entre marcadores: AS02 y AS43 presentaban la delección de 9 pb correspondiente al haplogrupo B, pero al hacer la digestión del fragmento diagnóstico del haplogrupo D utilizando el par de cebadores D2, en ocasiones se presentaba la ausencia de sitio de corte característica del mismo. Puesto que los haplogrupos fundadores son definidos por patrones de restricción mutuamente excluyentes (Schurr et al. 1990), se puede pensar ya sea en un haplogrupo infrecuente que presentara ambos marcadores, ya en un error de diagnóstico. Sin embargo, se pudo posteriormente observar que la digestión de fragmentos amplificados con el juego de cebadores D2 rendía resultados contradictorios: No sólo falló en digerir fragmentos digeridos

en experimentos previos, sino que falló en digerir el ADN del investigador que presenta el sitio de restricción diagnóstico.

Se desconoce la razón de los inconvenientes generados por el fragmento D2 (diseñado por Handt et al. 1996); se optó sin embargo por descartar momentáneamente los resultados obtenidos con el par de cebadores correspondiente, decisión que afectó a tres muestras: Por un lado, llevó a clasificar a los individuos AS02 y AS43 como haplogrupo B. Por otra parte, AS44, que había dado resultados contradictorios al repetir en el mismo el RFLP del haplogrupo D, pero que no había dado positivo para ningún otro haplogrupo, retuvo no obstante el diagnóstico de haplogrupo D, puesto que había sido diagnosticado con el par de cebadores original. Se asume que las digestiones exitosas en AS44 se deben a la presencia de fragmentos de ADN exógeno que en algunos casos amplificaron con preferencia al ADN del individuo.

Individuo	Unidad cronol.	Tipo Muestra	Fecha extracción	Masa Muestra (g)	ADN	Haplogrupo
AS 01	3	Hueso	09/03/2004	0,12	Sí	S/d
AS 02	3	Hueso	09/03/2004	0,1	Sí	B
AS 04	3	Hueso	27/04/2004	0,11	Sí	B
AS 05	3	Hueso	27/04/2004	0,16	Sí	C***
AS 07	3	Hueso	14/05/2004	0,12	No	
AS 08	3	Hueso	14/05/2004	0,16	No	
AS 14	S/d	Hueso	28/05/2004	0,175	No	
AS 19	2	Hueso	28/05/2004	0,23	No	
AS 20	1	Hueso	25/06/2004	0,32	Sí	C****
AS 22	4	Hueso	13/07/2004	0,25	Sí	K*
AS 23	2	Hueso	13/07/2004	0,28	No	
AS 24	1	Hueso	23/07/2004	0,27	Sí	C****
AS 25	S/d	Hueso	23/07/2004	0,33	No	
AS 27	1	Diente	13/08/2004	0,16	Sí	C
AS 30	S/d	Diente	13/08/2004	0,14	No	
AS 31	2	Hueso	21/09/2004	0,36	Sí	K*
AS 32	1	Hueso	21/09/2004	0,24	Sí	S/d**
AS 36	1	Diente	21/09/2004	0,26	No	
AS 39	1	Diente	05/10/2004	0,21	Sí	K*
AS 40	2	Hueso	05/10/2004	0,3	Sí	S/d**
AS 41	S/d	Hueso	05/10/2004	0,24	Sí	K*
AS 43	S/d	Diente	19/10/2004	0,23	Sí	B
AS 44	S/d	Diente	19/10/2004	0,175	Sí	D

Tabla 16: Resumen de los resultados de extracción y diagnóstico de la muestra de Arroyo Seco 2.

(*) – Muestra contaminada por ADN del investigador de acuerdo a secuencia de HVR I.

(**) – Muestra presuntamente contaminada por ADN del investigador.

(***) – Haplogrupo definido por secuencia de HVR I.

(****) – Haplogrupo definido por RFLP y confirmado por secuencia de HVR I.

AS05 por su parte, fue diagnosticado negativo para el haplogrupo C debido a la presencia del sitio de restricción *HincII* en el sitio 13259; sin embargo, la secuenciación del fragmento M4 mostró la presencia de las mutaciones diagnósticas de este haplogrupo, con lo cual se resolvió diagnosticarlo como haplogrupo C.

En algunos casos de contaminación con ADN moderno, la amplificación se realizó con relativa facilidad; la media de intentos de amplificación es de 2,79 en los casos supuestamente no contaminados, en tanto que es de 1,83 para los casos comprobada o supuestamente contaminados. Sin embargo, esta diferencia no es estadísticamente significativa (test de Kolmogorov-Smirnov, $D = 0,278$, $p = 0,33$), y una observación más detenida muestra que algunos marcadores no pudieron ser amplificados en los casos contaminados.

Muestra	Marcador				Media muestra
	A	B	C	D	
AS 01	1 (1)	3 (2)	1	1 (1)	1,5
AS 02	3 (1)	1 (1)	1	2 (2)	1,75
AS 04	-	3 (2)	1	3 (1)	2,33
AS 05	2 (1)	3 (1)	3	1 (2)	2,25
AS 20	3 (2)	1 (2)	1	2 (1)	1,5
AS 22	3 (1)	1 (1)	-	1 (1)	1,67
AS 24	1 (1)	2 (1)	1	3 (1)	1,75
AS 27	4 (1)	3 (2)	1	-	2,67
AS 31	4 (2)	2 (1)	1	1 (1)	2
AS 32	-	1 (1)	-	1 (1)	1
AS 39	-	3 (1)	3	2 (1)	2,67
AS 40	-	1 (1)	1	1 (1)	1
AS 41	-	2 (1)	1	1 (1)	1,33
AS 43	2 (1)	2 (2)	2	3 (1)	2,25
AS 44	-	4 (2)	2	2 (1)	2,67
<i>Media haplogrupo</i>	2,56	2,13	1,46	1,71	

Tabla 17: Cantidad de experimentos realizados para lograr la primera amplificación de las muestras con los marcadores de RFLP empleados. Las muestras sombreadas corresponden a casos demostrada o presuntamente contaminados por el ADN del investigador.

(-) No se logró amplificar con el marcador.

(1) Amplificó con el juego de cebadores original

(2) Amplificó con el juego de cebadores corto (A2, B2 o D2, véase tabla 12).

4.2.2. Amplificación de los controles de extracción

En paralelo con los análisis de los extractos se realizó el análisis de los controles de extracción. En los casos en que amplificara uno o más de los marcadores empleados para RFLP, se intentó amplificar el segmento 4 de la región hipervariable I a fin de confirmar o descartar que el origen del ADN contaminante fuera el investigador. Los resultados de las amplificaciones se presentan en la tabla 17. Como puede observarse, sólo se pudo descartar en forma terminante la contaminación de los blancos en dos extracciones, en tanto que en cuatro extracciones se pudo identificar el ADN contaminante como procedente del investigador. Los marcadores amplificados por RFLP dieron diagnóstico negativo para el haplogrupo en cuestión, si bien los resultados del haplogrupo D, alcanzados con el par de cebadores D2, como se indicó previamente, deben tomarse con cautela.

Control	A*	B*	C	D*	HVR segm 4	Conclusión
09/03/04	No	Sí	Sí	Sí	No	S/d; presencia de fragmentos \leq 100 pb
27/04/04	Sí	Sí	Sí	Sí	Sí	Investigador
25/06/04	No	No	No	No		No contaminado
13/07/04	No	Sí	No	Sí	Sí	Investigador
23/07/04	No	Sí	Sí	Sí	Sí	Investigador
13/08/04	No	Sí	No	No	No	S/d; presencia de fragmentos \leq 68 pb
21/09/04	No	No	No	Sí	No	S/d; presencia de fragmentos \leq 65 pb
05/10/04	No	No	No	No		No contaminado
19/10/04	No	No	Sí	Sí	Sí	Investigador

Tabla 18: Control de contaminación de los controles de extracción. “Sí” y “No” indica amplificación – o no – del marcador correspondiente al haplogrupo.

(*) Los marcadores de los haplogrupos A, B y D amplificaron únicamente con los juegos de cebadores A2, B2 y D2 (118, 68 y 65 pb respectivamente).

Reciclar insumos queda, por otra parte, fuera de toda cuestión. En las tandas de extracción del 23 de julio y 13 de agosto, se incluyó un segundo blanco de extracción empleando tubos de 15 ml utilizados en extracciones anteriores, reutilizados previo autoclavado y 30 minutos de irradiación UV. Sin excepción, estos blancos amplificaron ADN en las reacciones de PCR.

Sin embargo, desechar todas las muestras de una tanda de extracción sobre la base de la contaminación del blanco de extracción resulta prematuro: En la primera tanda de extracción, que incluyó a los individuos AS01 y AS02, el análisis de los extractos con el marcador del haplogrupo C arrojó que el blanco de extracción amplificaba este segmento, diagnosticándose negativo para el mismo. Ulteriores extracciones fueron detenidas hasta tanto no fueran analizados los individuos AS01 y AS02 para los marcadores indígenas A, B, C y D,

y se comprobaba la autenticidad al menos de uno de ellos. Enseguida, el individuo AS02 dio resultado positivo para el haplogrupo B, en tanto que posteriores intentos de amplificación del control de extracción no tuvieron éxito. Más adelante, el uso de pares de cebadores que amplificaran segmentos de largos menores a 100 bp (B2 y D2) resultaron en producto amplificado en los controles. Habiéndose constatado con relativa seguridad la existencia de ADN endógeno en los extractos, aún con amplificación de los controles de extracción, se prosiguió con las extracciones.

4.2.3. Contaminación de los extractos

Del mismo modo en que la detección de ADN moderno en los controles de extracción no impidió constatar la extracción de ADN auténtico, la ausencia de contaminación en los controles no impidió que se constatará contaminación de los extractos: Aún cuando no logró detectarse contaminación en el extracto del 5 de octubre, hay razones para sospechar que todos los extractos de esa tanda (AS39, AS40 y AS41) están contaminados con ADN moderno (ver Anexo II).

Desde luego, el hecho de que una muestra sea perteneciente a un haplogrupo indígena no excluye por completo la contaminación: En primer lugar puede haber habido contaminación cruzada o “*carryover*”, consistente en el traspaso del ADN de una muestra al extracto de otra, con lo cual un ADN a primera vista endógeno sería el producto de la contaminación con el ADN de otra muestra. Hay tres elementos que nos permiten prácticamente descartar la contaminación cruzada:

- a) No coincidieron nunca dos extractos con el mismo haplogrupo en la misma tanda de extracción (Tabla 16). Los individuos sucesivos diagnosticados como pertenecientes al haplogrupo C fueron sujetos a extracción en tandas distintas, mediadas por la limpieza e irradiación de las de las superficies y los elementos de trabajo. Asimismo, los individuos diagnosticados del mismo haplogrupo lo fueron además en distintos experimentos de PCR. Esto permite descartar la contaminación *carryover* en el correr de las amplificaciones.
- b) La mayor parte de los blancos de extracción contaminados presentaron contaminación con ADN del investigador. En el caso de los que no pudo comprobarse la fuente de la contaminación, los análisis por RFLP dieron resultado negativo para los haplogrupos amerindios amplificados

- c) Los controles contaminados en los que el origen del ADN no pudo identificarse mostraban presencia de fragmentos de ADN en el orden de los 100 pb o menores. Si bien esto podría suponerse contaminación con ADN altamente degradado de posible origen antiguo, no debe olvidarse que el hipoclorito de sodio tiene por efecto fragmentar la molécula de ADN en piezas menores (Willerslev et al. 2004).

En segundo lugar, hay una fuente de contaminación adicional que podría dar por resultado haplogrupos indígenas. Los entierros de Arroyo Seco 2 constituyen una muestra que debido a su buena preservación, su tamaño y su antigüedad documentada con buenos fechados ha sido manipulada por un gran número de investigadores. No hemos de olvidar además que el método de excavación no incluyó precauciones de ningún tipo – siquiera uso de guantes – para evitar la contaminación de las piezas. En todo caso, la probabilidad de contaminación de las muestras con ADN mitocondrial indígena durante la recuperación y el estudio puede considerarse de entre 40 y 50% (en la Plata, 45% de los haplogrupos mitocondriales analizados fueron amerindios, en una muestra que incluyó personal de la Universidad y personas que se atendían en un hospital público – Martínez-Marignac et al. 2004). Sin embargo, en los casos de contaminación de fuente identificada, el ADN contaminante procedía invariablemente del investigador. Con ello se puede asumir que los eventos de contaminación se limitaron a la etapa de extracción del ADN, y no a individuos que hayan manipulado previamente la muestra, cualquiera sea la ascendencia de los mismos.

Considerando que en todos los casos de contaminación identificada (4 casos de 7), el ADN procedía del investigador, se puede afirmar que la principal y quizá única fuente de contaminación fue el ADN del investigador. En todos los casos, su origen se confirmó al leerse una transición T→C en la posición 16311 del segmento 4 de la región hipervariable I (ver Anexo II).

El ADN exógeno pudo además detectarse en algunos extractos junto con el ADN de la muestra, siendo esta la razón por la que en muchos trabajos (incluyendo los realizados sobre muestras de *Homo neanderthalensis* – cf. Krings et al. 1997; Krings et al. 2000; Ovchinnikov et al. 2000) se haya optado por clonar los productos de PCR a fin de aislar el ADN de la muestra de ADN de otro origen. Esto podría explicar los resultados de AS05: Esta muestra dio negativo para el haplogrupo C, pero en la secuencia correspondiente al segmento M4 pueden leerse las mutaciones T→C en la posición 16325 y C→T en la posición 16327, aún con el “ruido” introducido por otras secuencias en el ferograma (Figura 12). Si hay varios

ADN en el extracto, la PCR puede ora amplificar el ADN del individuo, ora ADN contaminante, rindiendo resultados contradictorios. Es por supuesto posible que la presencia de la base C en la secuencia sea producto de un artefacto de PCR, pero siendo que las “mutaciones” artefactuales más frecuentes producidas por *primer jumping* y deaminación oxidativa son G→A o C→T, no parece probable encontrar la presencia de una transición T→C, y menos aún asociada a otra transición (Pääbo et al. 1990). Queda por verse si el individuo en cuestión forma parte de los casos del haplogrupo C “revertidos”, esto es, casos en los cuales se ha recuperado el sitio de restricción ausente en el haplogrupo (Torres et al. 2006) o si se trata de una incoherencia relacionada con la presencia de ADN contaminante en la muestra.

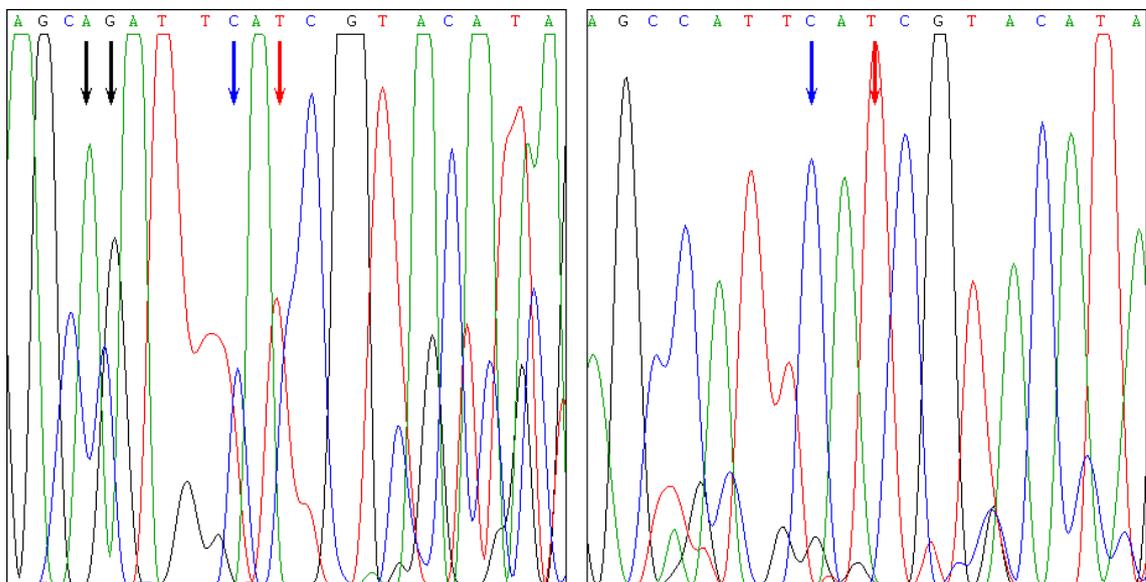


Figura 12: Ferogramas de las posiciones 16318 a 16335 de las secuencias de AS05 (izquierda) y AS20 (derecha), ambos con una transición T→C en la posición 16325 (flecha azul) y una transición C→T en la posición 16327 (flecha roja) Nótese el “ruido” introducido en la posición 16321 por señales A y G – probablemente provenientes de ADN contaminante – solapando la señal de C en el ferograma de AS05 (flechas negras).

En función de los resultados de los análisis de RFLP y secuencias, se alcanzó en síntesis la siguiente composición de haplogrupos en la muestra de Arroyo Seco 2:

Haplogrupo	A		B		C		D	
	N	%	n	%	N	%	n	%
8	0	0%	3	37,5%	4	50%	1	12,5%

Esta muestra, admitidamente pequeña, fue sometida a análisis para servir como base a las hipótesis relacionadas con las afinidades poblacionales de las poblaciones pampeanas del Holoceno Temprano.

4.3. Análisis de variabilidad y afinidad con otras poblaciones

4.3.1. AMOVA

Como puede observarse en las tablas 19 y 20, los AMOVA de todas las reasignaciones a poblaciones modernas y antiguas (ver materiales y métodos, tablas 15a y 15b) son significativos al compararlos con la distribución generada por las permutaciones, por lo que es necesario un análisis más detallado de los resultados. En las reasignaciones a agrupaciones de poblaciones modernas (Tabla 19), el agregado de la muestra de Arroyo Seco 2 a cualquiera de ellas no tuvo mayor impacto en la variación intra- e intergrupar, probablemente en virtud del tamaño muestral comparativamente pequeño. Los casos en que los componentes de varianza fueron menos modificados por el agregado de Arroyo Seco 2 fueron las reasignaciones 3, 4 y 5, correspondientes a los grupos chaqueños por un lado y por el otro, los grupos del tronco lingüístico andino a un lado y otro de la cordillera de los Andes. Debe destacarse no obstante que mantener a Arroyo Seco 2 en una agrupación aparte, si bien tuvo una significación estadística menor, mantuvo virtualmente intactos los componentes de varianza. Sobre esta base es razonable suponer que, sin ignorar la afinidad de Arroyo Seco 2 con los grupos araucanos, la distancia temporal respecto a los mismos hace de Arroyo Seco 2 una población con diferencias notorias respecto a las poblaciones indígenas actuales del Cono Sur.

Subdivisión	Componentes de varianza		Porcentajes de varianza	Índices de fijación	$p <$
1 – Aymara Atacameño	Intergrupala	0,0983 σ^2_a	24,91	Φ_{CT} : 0,2491	,000
	Interpoblacional	0,0266 σ^2_b	6,74	Φ_{SC} : 0,0897	,000
	Intrapoblacional	0,2697 σ^2_c	68,35	Φ_{ST} : 0,3165	,000
2 – Guarani (ARG) Guarani (BR)	Intergrupala	0,0939 σ^2_a	23,85	Φ_{CT} : 0,2385	,001
	Interpoblacional	0,0305 σ^2_b	7,73	Φ_{SC} : 0,1015	,000
	Intrapoblacional	0,2697 σ^2_c	68,42	Φ_{ST} : 0,3158	,000
3 – Chorote Mataco Pilaga Toba	Intergrupala	0,0985 σ^2_a	24,96	Φ_{CT} : 0,2496	,000
	Interpoblacional	0,0264 σ^2_b	6,68	Φ_{SC} : 0,0890	,000
	Intrapoblacional	0,2697 σ^2_c	68,36	Φ_{ST} : 0,3164	,000
4 – Mapuche (ARG) Tehuelche	Intergrupala	0,0986 σ^2_a	24,99	Φ_{CT} : 0,2499	,001
	Interpoblacional	0,0264 σ^2_b	6,68	Φ_{SC} : 0,0891	,000
	Intrapoblacional	0,2697 σ^2_c	68,33	Φ_{ST} : 0,3167	,000
5 – Mapuche (CH) Pehuenche Huilliche	Intergrupala	0,0991 σ^2_a	25,12	Φ_{CT} : 0,2512	,000
	Interpoblacional	0,0258 σ^2_b	6,53	Φ_{SC} : 0,0873	,000
	Intrapoblacional	0,2697 σ^2_c	68,35	Φ_{ST} : 0,3165	,000
Ninguna	Intergrupala	0,0986 σ^2_a	25,01	Φ_{CT} : 0,2501	,004
	Interpoblacional	0,0260 σ^2_b	6,59	Φ_{SC} : 0,0879	,000
	Intrapoblacional	0,2697 σ^2_c	68,39	Φ_{ST} : 0,3161	,000

Tabla 19: Componentes, porcentajes de varianza e índices Φ resultantes del AMOVA de la reasignación de la muestra de Arroyo Seco 2 a cada una de las agrupaciones modernas descritas en la tabla 15a. Compárese los índices Φ de las reasignaciones con los que aparecen en la tabla 15a.

Subdivisión	Componentes de varianza		% de varianza	Índices de fijación	$p <$
1 – Amazonia Norte Chile Colombia	Intergrupala	0,0889 σ^2_a	22,93	Φ_{CT} : 0,2293	,000
	Interpoblacional	0,0079 σ^2_b	2,05	Φ_{SC} : 0,0266	,000
	Intrapoblacional	0,2909 σ^2_c	75,01	Φ_{ST} : 0,2499	,000
2 – Uruguay E	Intergrupala	0,0827 σ^2_a	21,49	Φ_{CT} : 0,2149	,002
	Interpoblacional	0,0112 σ^2_b	2,91	Φ_{SC} : 0,0371	,000
	Intrapoblacional	0,2909 σ^2_c	75,60	Φ_{ST} : 0,244	,000
3 – Noroeste argentino	Intergrupala	0,0821 σ^2_a	21,24	Φ_{CT} : 0,2124	,004
	Interpoblacional	0,0135 σ^2_b	3,50	Φ_{SC} : 0,0445	,000
	Intrapoblacional	0,2909 σ^2_c	75,26	Φ_{ST} : 0,2474	,000
4 – Tierra del Fuego Patagonia	Intergrupala	0,0946 σ^2_a	24,52	Φ_{CT} : 0,2452	,000
	Interpoblacional	0,0004 σ^2_b	0,10	Φ_{SC} : 0,0014	,000
	Intrapoblacional	0,2909 σ^2_c	75,38	Φ_{ST} : 0,2462	,000
5 – Taíno Ciboney	Intergrupala	0,0876 σ^2_a	22,68	Φ_{CT} : 0,2268	,000
	Interpoblacional	0,0076 σ^2_b	1,96	Φ_{SC} : 0,0254	,000
	Intrapoblacional	0,2909 σ^2_c	75,35	Φ_{ST} : 0,2465	,000
Ninguna	Intergrupala	0,0911 σ^2_a	23,64	Φ_{CT} : 0,2364	,001
	Interpoblacional	0,0033 σ^2_b	0,87	Φ_{SC} : 0,0113	,000
	Intrapoblacional	0,2909 σ^2_c	75,50	Φ_{ST} : 0,245	,000

Tabla 20: Componentes, porcentajes de varianza e índices Φ resultantes del AMOVA de la reasignación de la muestra de Arroyo Seco 2 a cada una de las agrupaciones antiguas descritas en la tabla 15b. Compárese los índices Φ de las reasignaciones con los que aparecen en la tabla 15b.

La disrupción resultante de la reasignación de Arroyo Seco 2 a las reagrupaciones de muestras antiguas (Tabla 20) es más notable. De las reasignaciones realizadas, llama la atención la asignación de la muestra a la agrupación 4 donde, al contrario del aumento constatado en la variación intragrupal de las demás reasignaciones, la variación se reduce prácticamente al 10% ($\Phi_{SC} = 0,0014$, comparado con el Φ_{SC} original de 0,0117) manteniéndose prácticamente intacta la variación intergrupala. La agrupación 4 incluye muestras arqueológicas prehistóricas e históricas de Tierra del Fuego y Patagonia, constatándose una relativa coherencia respecto a lo observado con muestras modernas.

4.3.2. Análisis de distancias

4.3.2.1. Poblaciones modernas

4.3.2.1.1. D_A de Nei

Los dendrogramas de las muestras de Arroyo Seco 2 y modernas analizadas por la distancia D_A se presentan en las figuras 13 y 14. Ambas modalidades de *bootstrap* presentaron una consistencia importante en cuanto a topología, si bien claramente el soporte de los nodos arrojado por el *bootstrap* de “*loci*” es menor en todos los casos. Dos *clusters* se destacan consistentemente en ambos análisis:

- Un *cluster* consistente en Arroyo Seco 2 junto con los Tehuelches y los grupos de habla araucana de Chile;
- Un *cluster* conteniendo los grupos Mataco, Aymara y Atacameño.

En forma menos consistente se agrupan las muestras guaraníes en la vecindad de tobas y mapuches argentinos. Los grupos Pilaga y Chorote o Chorotí, chaqueños, se encuentran en la periferia, aunque en diferentes jerarquías de ramificación según el análisis empleado.

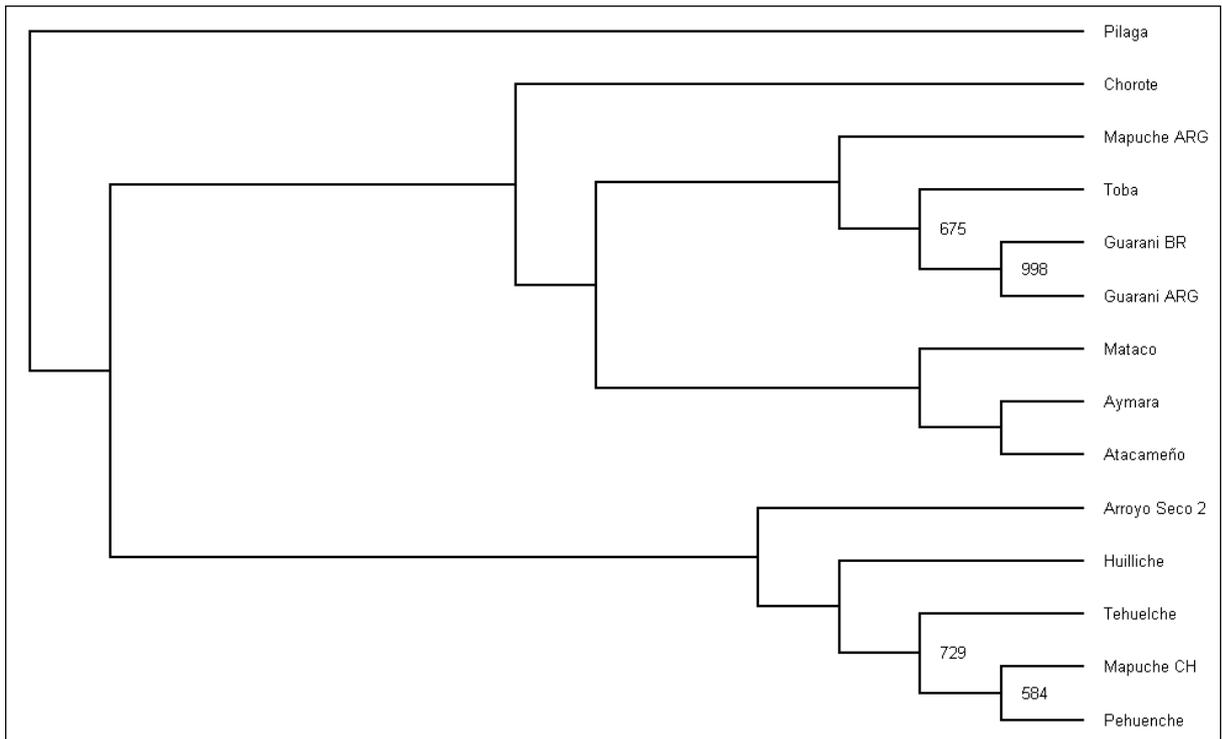


Figura 13: Árbol elaborado por el método *neighbor-joining* sobre distancias D_A de Arroyo Seco 2 y poblaciones modernas. La consistencia del árbol fue testada mediante *bootstrap* de individuos de cada muestra. Los nodos con un soporte igual o mayor que 50% presentan la cantidad de remuestreos en un total de 1000 en que se repitió la bifurcación.

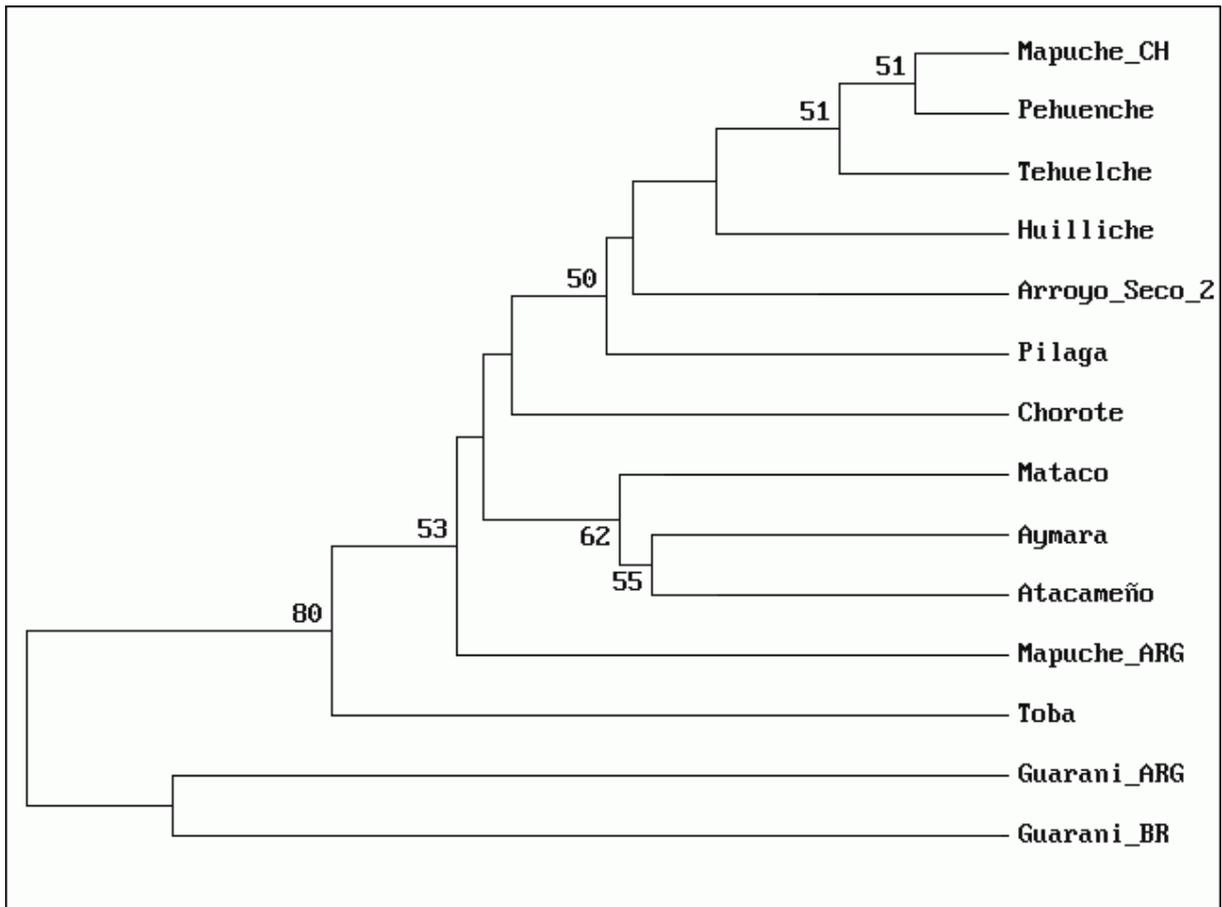


Figura 14: Árbol elaborado por el método *neighbor-joining* sobre distancias D_A de Arroyo Seco 2 y poblaciones modernas. La consistencia del árbol fue testada mediante *bootstrap* de sitios de “*loci*”. Los nodos con un soporte igual o mayor que 50% presentan el porcentaje en un total de 1000 remuestreos en que se repitió la bifurcación.

Pasando este análisis al plano mediante el análisis de coordenadas principales, los tres primeros ejes responden por el 73% de la variación total, con los siguientes valores propios:

	Eje 1	Eje 2	Eje 3
Valor propio	1,120	0,519	0,207
%Varianza	44%	21%	8%
%Varianza acum.	44%	65%	73%

El eje 1 está asociado a la frecuencia del haplogrupo A, en tanto que el eje 2 está relacionado con la frecuencia del haplogrupo B. Al graficar los valores de las poblaciones para estos ejes (Figura 15) se distinguen dos conjuntos: Uno con altos valores en el eje 2 (Atacameño, Matabo y Aymara) y otro con valores negativos para el eje 1 y valores entre $-0,1$

y 0,1 para el eje 2 (Pilagas, Chorotes, Huilliches y Mapuches argentinos). Más alejados y relativamente dispersos se encuentran Arroyo Seco 2, Tehuelches, Pehuenches, Mapuches chilenos y Guaraníes.

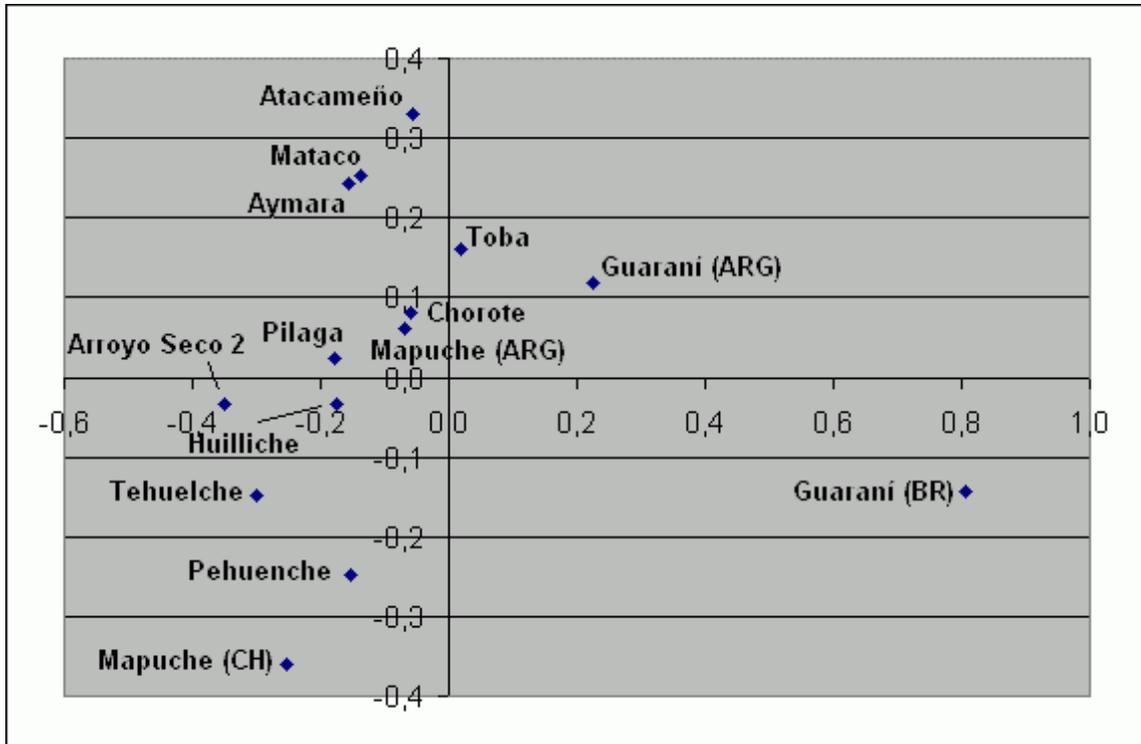


Figura 15: Representación de los dos primeros ejes del análisis de coordenadas principales de la matriz de distancias D_A entre las poblaciones de Arroyo Seco 2 y modernas.

4.3.2.1.2. D_R de Rogers

Los dendrogramas de las muestras de Arroyo Seco 2 y modernas analizadas por la distancia D_R se presentan en las figuras 16 y 17. Las distancias D_R resultan en patrones de agrupación que en general se visualizan en los árboles elaborados a partir de ambas modalidades de *bootstrap*, aunque con menor soporte en el remuestreo de individuos. Se constatan dos grandes *clusters*:

- Un *cluster* agrupa a Matacos, Atacameños y Aymaras, incluyendo en sus cercanías a Tobas y Guaraníes;
- El otro *cluster* incluye Pilagas, Huilliches, Tehuelches, Mapuches chilenos y Pehuenches.

Aparte de estos dos grandes grupos, el comportamiento de Arroyo Seco 2, Chorotes y Mapuches argentinos es variable. Los primeros dos tienden a agrupar juntos, cerca del segundo *cluster*, aunque en forma poco robusta. El tercero agrupa en forma errática entre el primer y segundo *cluster*, según el patrón utilizado de remuestreo.

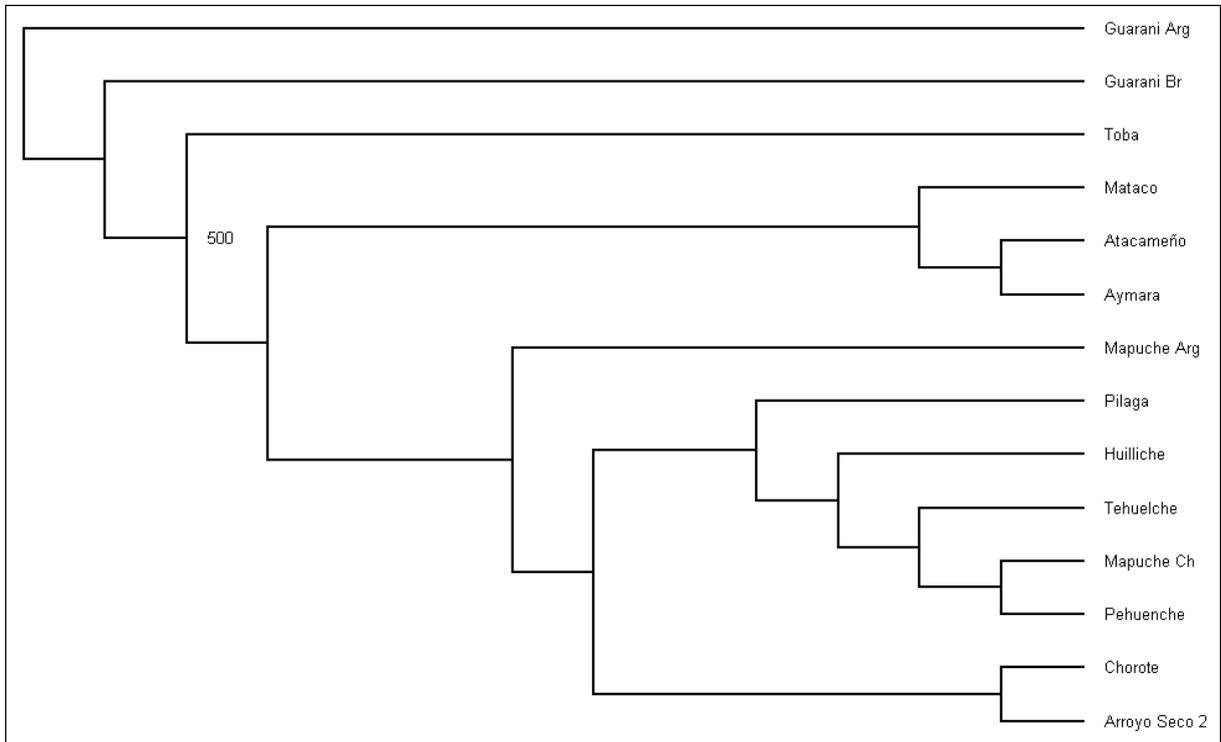


Figura 16: Árbol elaborado por el método *neighbor-joining* sobre distancias D_R de Arroyo Seco 2 y poblaciones modernas. La consistencia del árbol fue testada mediante *bootstrap* de individuos de cada muestra. Los nodos con un soporte igual o mayor que 50% presentan la cantidad de remuestreos en un total de 1000 en que se repitió la bifurcación.

En el análisis de coordenadas principales, los tres primeros ejes responden por el 95% de la variación total, con los siguientes valores propios:

	Eje 1	Eje 2	Eje 3
Valor propio	0,650	0,410	0,129
% Varianza	52%	33%	10%
% Varianza acum.	52%	85%	95%

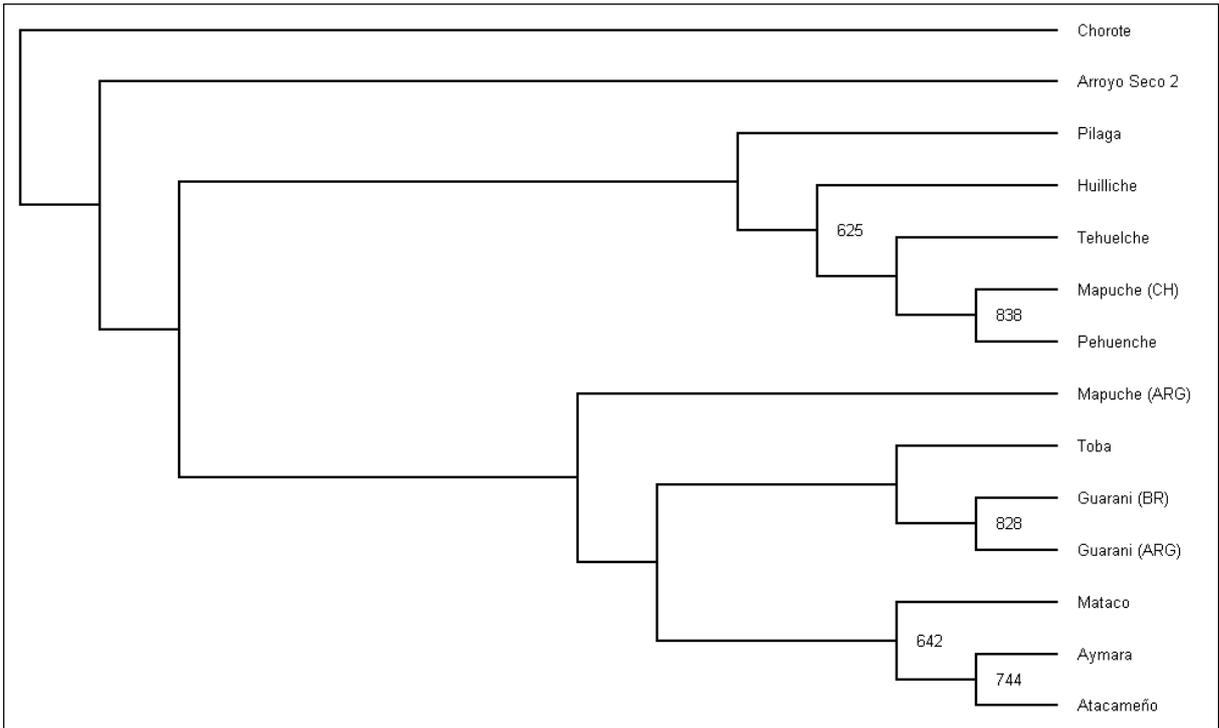


Figura 17: Árbol elaborado por el método *neighbor-joining* sobre distancias D_R de Arroyo Seco 2 y poblaciones modernas. La consistencia del árbol fue testada mediante *bootstrap* de "loci". Los nodos con un soporte igual o mayor que 50% presentan la cantidad de remuestreos en un total de 1000 en que se repitió la bifurcación..

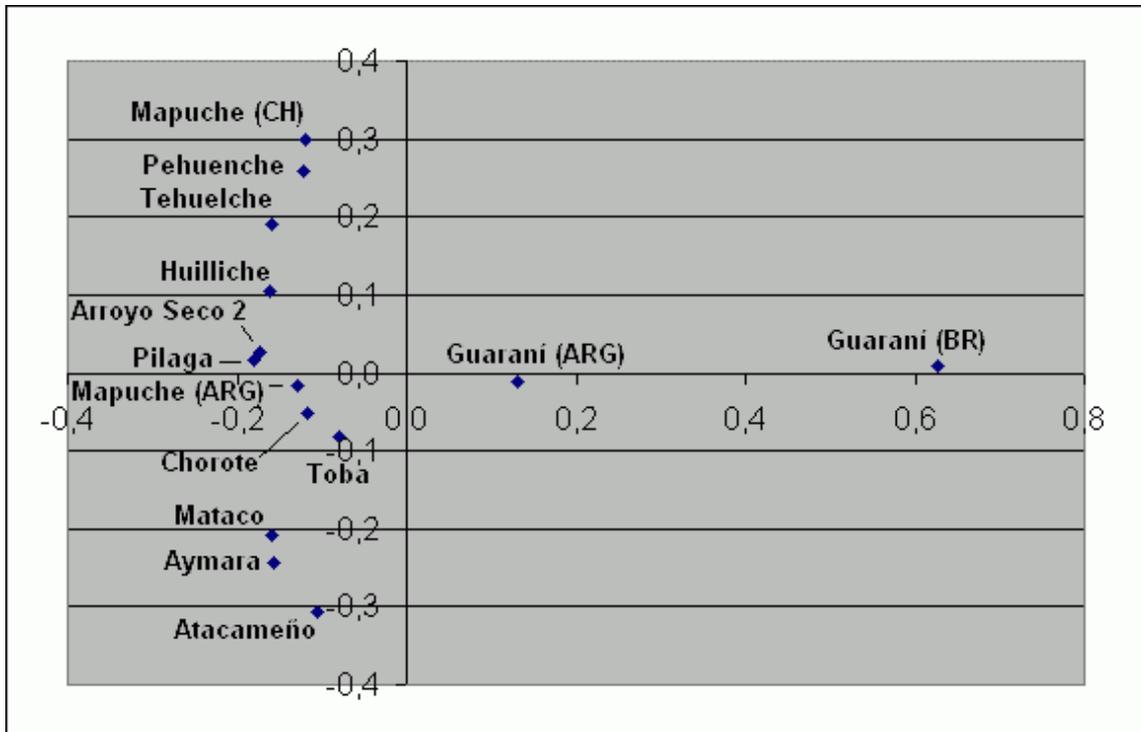


Figura 18: Representación de los dos primeros ejes del análisis de coordenadas principales de la matriz de distancias D_R entre las poblaciones de Arroyo Seco 2 y modernas.

Una vez más, los dos primeros ejes, que responden por el 85% de la variación, se asocian el eje 1 con las frecuencias del haplogrupo A y el eje 2 con las frecuencias del haplogrupo B. Observando la figura 18 se constatan tres agrupaciones: Una consistente en Mataco, Aymara y Atacameño, con valores negativos en ambos ejes. Con valores negativos para el eje 1 y en el entorno de -0,1 a 0,1 para el eje 2 se agrupan luego Arroyo Seco 2, Pilagas, Chorotes, Mapuches argentinos y Tobas. Con valores más altos en el eje 2 tienden a agruparse los grupos chilenos de habla araucana y los Tehuelches. Por último y claramente diferenciados entre sí y respecto a los demás grupos se encuentran los grupos Guaraníes.

Al analizar la correlación entre distancias genéticas y distancias geográficas entre las muestras, se constata una correlación positiva significativa (Distancias D_A : $r = 0,368$, $p < 0,002$; distancias D_R : $r = 0,396$, $p < 0,002$; la significación fue establecida mediante test de Mantel con 500 permutaciones). Esta correlación positiva es consistente con patrones de aislamiento por distancia continuos o del tipo *stepping-stone* uni- y bidimensionales (*cf.* Ramachandran et al. 2005 para una aplicación reciente de este modelo en poblaciones humanas). Lo que no puede resolverse dada la resolución de los datos constituye el núcleo mismo del debate en torno al poblamiento de América, esto es, si las diferencias poblacionales responden a una microevolución posterior al poblamiento por un contingente homogéneo, o si por el contrario hay un cierto grado de variación “heredada” de las poblaciones asiáticas que dieran origen a las poblaciones americanas actuales mediante varios eventos migratorios.

El análisis de correlaciones entre las distancias geográficas y genéticas de Arroyo Seco 2 con las poblaciones modernas utilizadas da correlaciones no significativas (Distancias D_A : $r = 0,159$, $p > 0,6$; distancias D_R : $r = 0,199$, $p > 0,5$). La razón de ello se encuentra probablemente en la distancia temporal entre la muestra analizada y las de referencia, y los varios eventos que alteraron tanto la localización de las poblaciones como sus características en este lapso, eventos que se discutirán más adelante.

4.3.2.2. Poblaciones antiguas

4.3.2.2.1. D_A de Nei

Los dendrogramas a partir de las distancias D_A de Nei et al. (1983) entre poblaciones antiguas, testados mediante *bootstrap* de individuos (Figura 19) y de “loci” (Figura 20) mostraron los siguientes patrones de agrupación:

- Un *cluster* robusto agrupando las muestras fueguinas y patagónicas;
- Otro *cluster*, con idéntico patrón de agrupación en ambos dendrogramas, aunque con distinto soporte estadístico, en el siguiente orden: Amazonia, agrupado con Colombia, en segundo orden la muestra de los valles de Lluta, Camarones y Azapa del norte chileno, en tercer término la muestra del este uruguayo y las muestras de Salta y Arroyo Seco 2 en cuarto y quinto término respectivamente.

En una posición intermedia en ambos dendrogramas se encuentran las muestras de las Antillas, si bien el árbol generado por *bootstrap* de individuos los sitúa más cerca del segundo *cluster* mencionado que del primero.

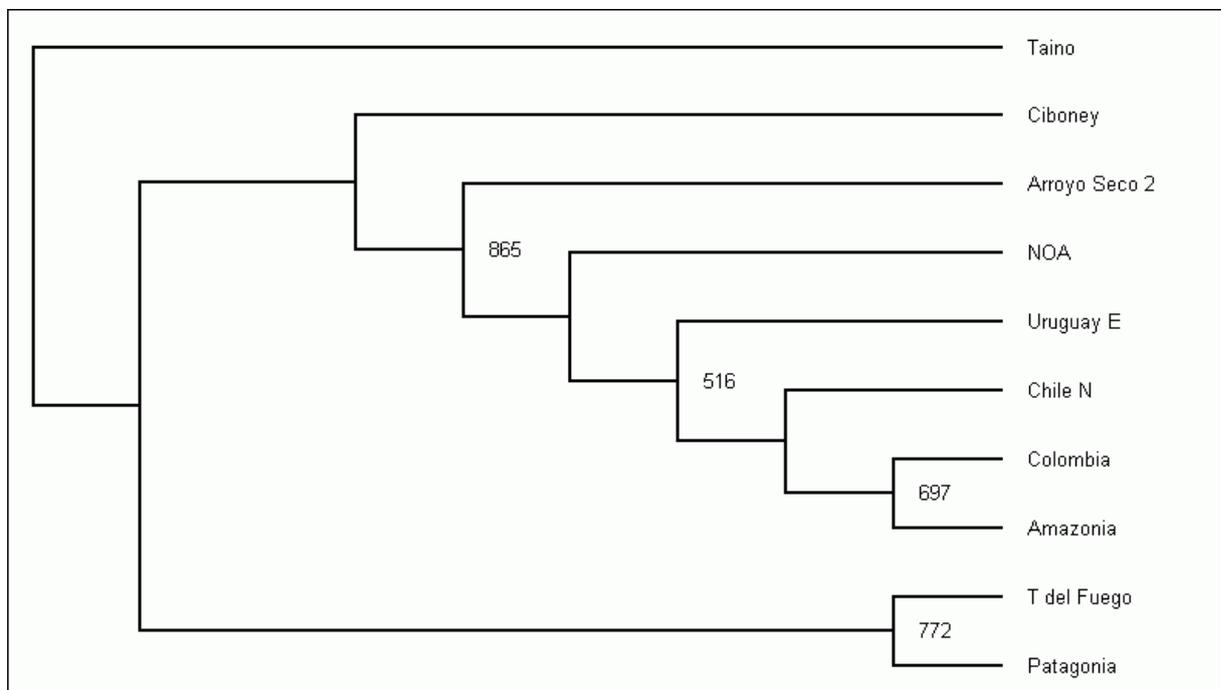


Figura 19: Árbol elaborado por el método *neighbor-joining* sobre distancias D_A de Arroyo Seco 2 y poblaciones antiguas. La consistencia del árbol fue testada mediante *bootstrap* de individuos. Los nodos con un soporte igual o mayor que 50% presentan el número de remuestreos en un total de 1000 en que se repitió la bifurcación.

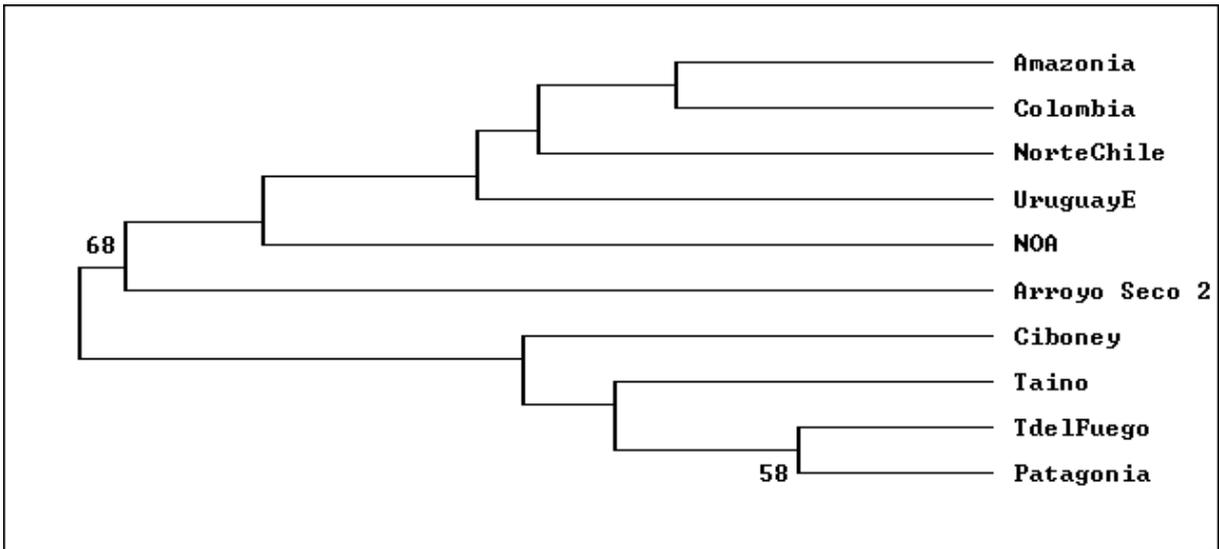


Figura 20: Árbol elaborado por el método *neighbor-joining* sobre distancias D_A de Arroyo Seco 2 y poblaciones antiguas. La consistencia del árbol fue testada mediante *bootstrap* de “*loci*”. Los nodos con un soporte igual o mayor que 50% presentan el porcentaje en un total de 1000 remuestreos en que se repitió la bifurcación.

En el análisis de coordenadas principales, los primeros tres ejes dan cuenta del 85% de la variación:

	Eje 1	Eje 2	Eje 3
Valor propio	1,495	0,756	0,298
%Varianza	50%	25%	10%
%Varianza acum.	50%	75%	85%

El primer eje está asociado con las frecuencias de los haplogrupos A y D, en tanto que el segundo eje está asociado con las frecuencias de los haplogrupos B y C. Al observar la representación gráfica de las muestras en los ejes primero y segundo (Figura 21) se visualizan dos agrupaciones, una de Fuego-Patagonia y las Antillas, y otra incluyendo Amazonia, Colombia, el este uruguayo y el norte de Chile. En posiciones intermedias entre estas agrupaciones se encuentran Arroyo Seco 2 y el noroeste argentino.

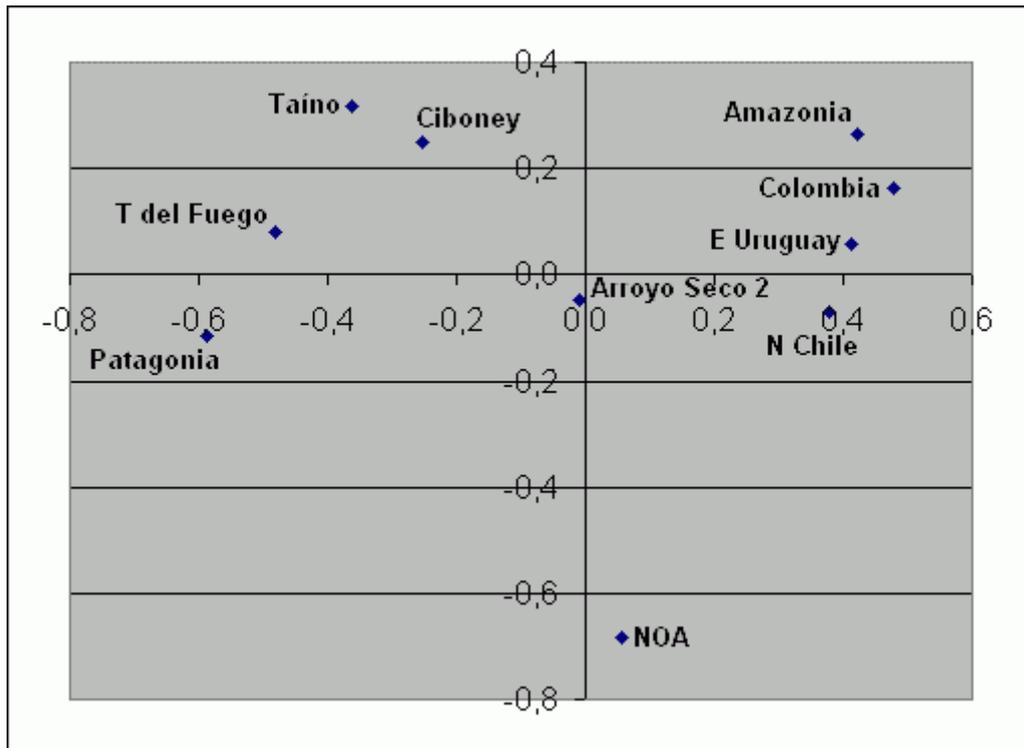


Figura 21: Representación de los dos primeros ejes del análisis de coordenadas principales de la matriz de distancias D_A entre las poblaciones de Arroyo Seco 2 y antiguas.

4.3.2.2.2. D_R de Rogers

Empleando la distancia D_R , los dendrogramas testados mediante ambos tipos de *bootstrap* repiten con claridad un *cluster* consistente en las muestras del las Antillas junto con las de Fuego-Patagonia, cada par en un grupo aparte pero unidos en una misma rama. Siguiendo este *cluster* claro, se van agregando las ramas correspondientes al noroeste argentino, Arroyo Seco 2 y el este uruguayo. Hay diferencias en la claridad con la que agrupan las muestras del norte de Chile, Amazonia y Colombia, si bien se ven juntas las tres poblaciones en ambos dendrogramas.

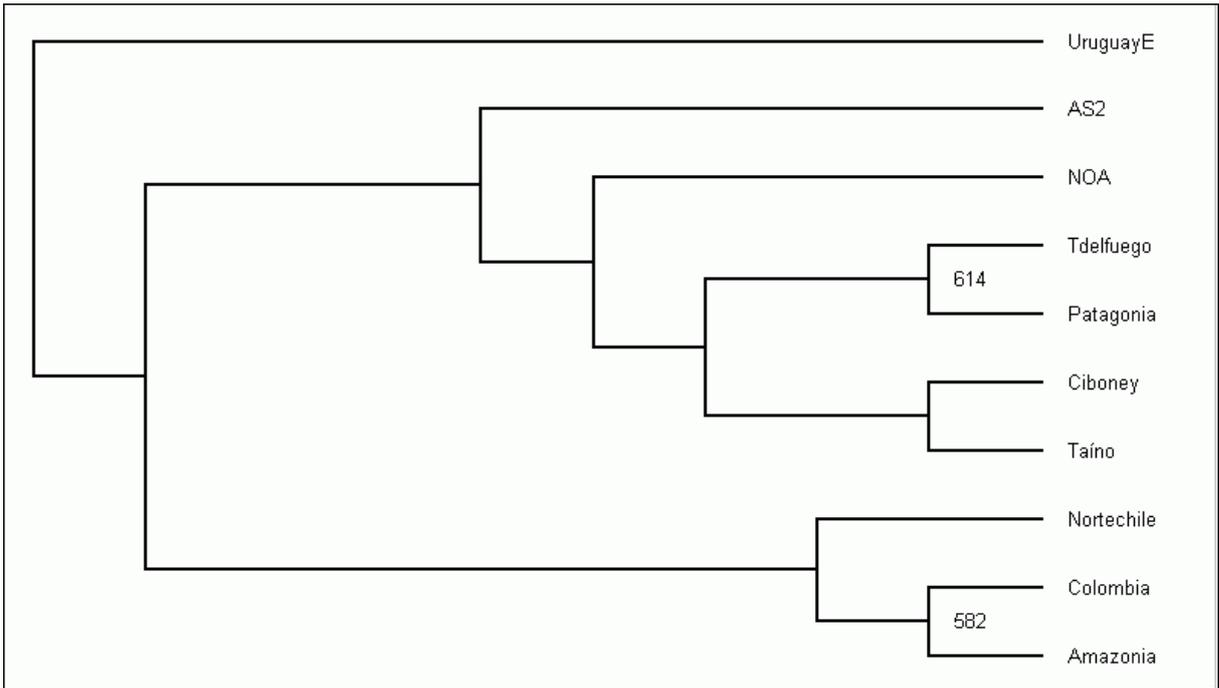


Figura 22: Árbol elaborado por el método *neighbor-joining* sobre distancias D_R de Arroyo Seco 2 y poblaciones antiguas. La consistencia del árbol fue testada mediante *bootstrap* de individuos. Los nodos con un soporte igual o mayor que 50% presentan el número de remuestreos en un total de 1000 en que se repitió la bifurcación.

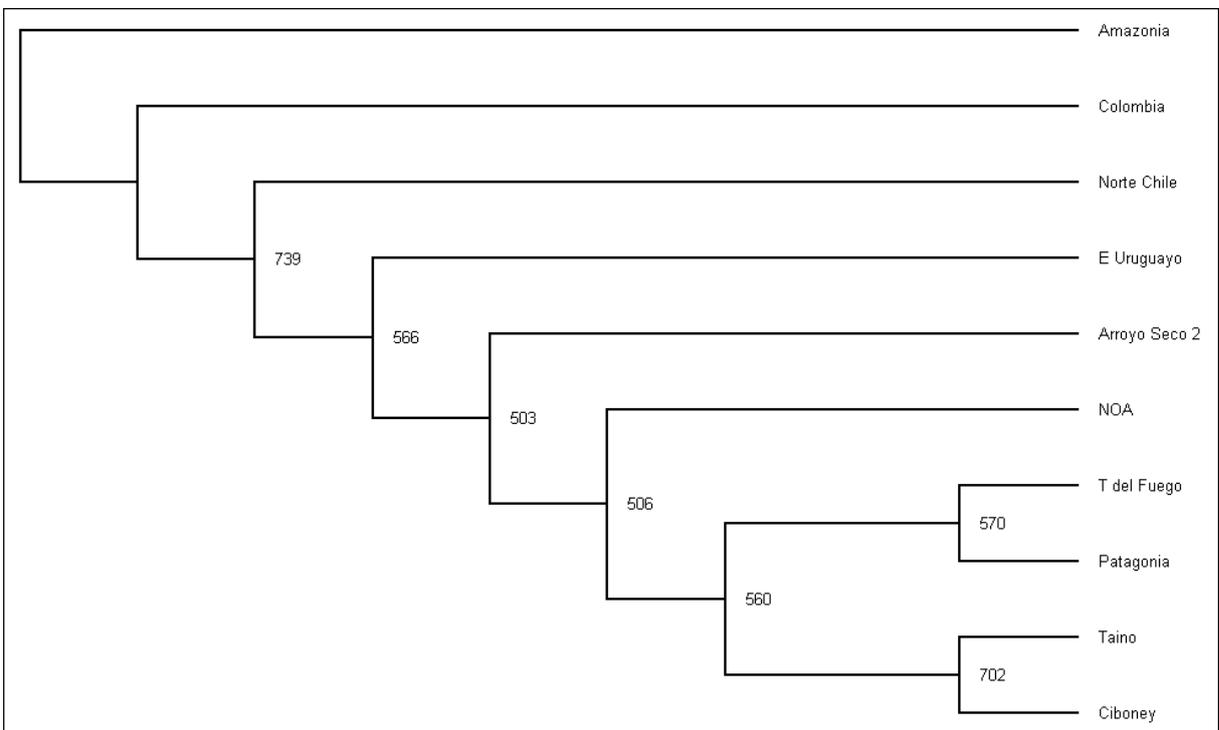


Figura 23: Árbol elaborado por el método *neighbor-joining* sobre distancias D_R de Arroyo Seco 2 y poblaciones antiguas. La consistencia del árbol fue analizada mediante *bootstrap* de "loci". Los nodos con un soporte igual o mayor que 50% presentan el número de remuestreos en un total de 1000 en que se repitió la bifurcación.

Los tres primeros ejes del análisis de coordenadas principales de las muestras antiguas empleando la distancia D_R , dan cuenta de nada menos que del 99%. El primer eje está asociado con las frecuencias de los haplogrupos A y C, en tanto que el segundo eje está asociado con las frecuencias de los haplogrupos B y D.

	Eje 1	Eje 2	Eje 3
Valor propio	0,457	0,289	0,145
%Varianza	51%	32%	16%
%Varianza acum.	51%	83%	99%

En la figura 24 pueden verse dos agrupaciones: Una incluye a Tierra del Fuego y Patagonia, y la otra agrupa a Amazonia, Colombia, el norte de Chile y el noroeste argentino. Arroyo Seco 2 se encuentra en una posición separada pero tendiente a aproximarse a las muestras de Fuego-Patagonia, en tanto que el este uruguayo y las Antillas se encuentran separadas entre sí y respecto a los demás grupos, con la excepción de la muestra de República Dominicana que se encuentra en vecindad lejana del *cluster* amazónico-andino.

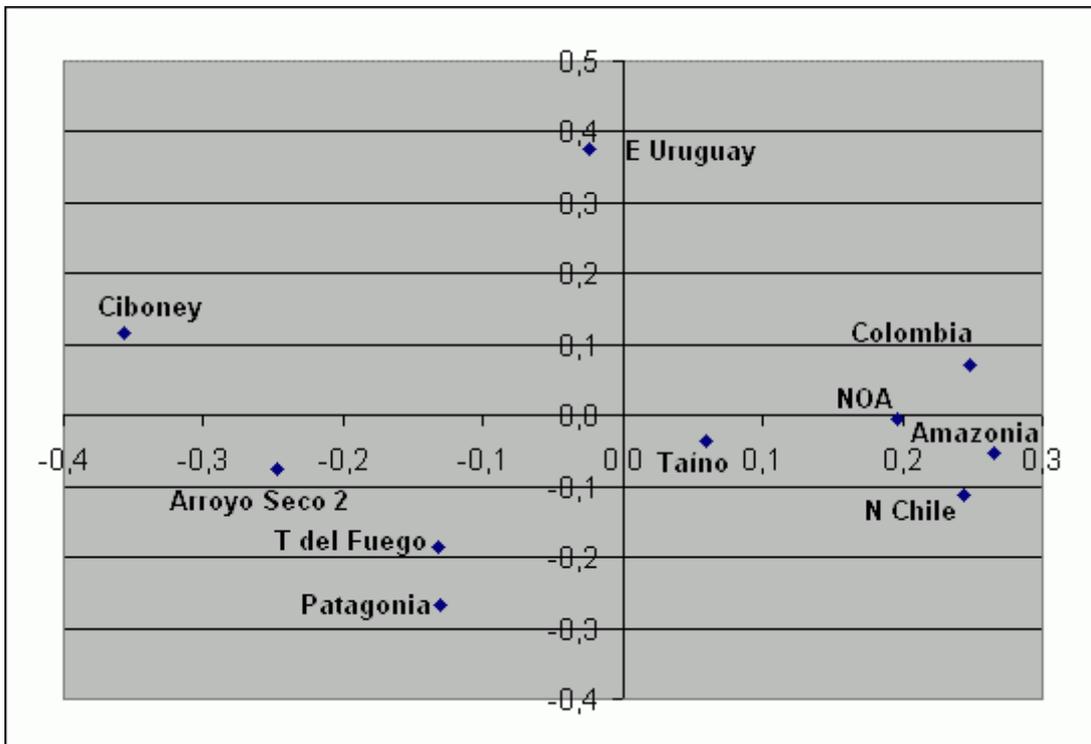


Figura 24: Representación de los dos primeros ejes del análisis de coordenadas principales de la matriz de distancias D_R entre las poblaciones de Arroyo Seco 2 y antiguas.

El test de Mantel entre las matrices de distancias genéticas y geográficas entre las muestras arroja resultados no significativos (Distancias D_A : $r = -0,138$, $p > 0,8$; distancias D_R : $r = -0,058$, $p > 0,6$; la significación fue establecida mediante 500 permutaciones). Las razones para ello se encuentran sin lugar a dudas tanto en consideraciones de muestreo como de la dispersión temporal de las muestras, que lleva a diferencias cronológicas en algunos casos de 7000 años. Rigen por ello las mismas consideraciones hechas en el análisis de correlaciones entre Arroyo Seco 2 y las muestras modernas, esto es, que la separación temporal implica una serie de eventos cuyo panorama está al día de hoy lejos de completo, y que implica la modificación tanto de la localización de las poblaciones como de sus características.

5. Discusión

5.1. Extracción y análisis de ADN

5.1.1. Tasa de éxito de las extracciones

El índice de éxito de la extracción de ADN de la muestra de Arroyo Seco 2 fue de 35%, lo cual se encuentra en el extremo inferior de las tasas de éxito registradas en diversos trabajos realizados sobre muestras americanas (Tabla 21). Sin embargo, ordenando las muestras de la tabla 21 de acuerdo al punto medio de los rangos de antigüedades, y tomando de Arroyo Seco 2 los primeros tres eventos de inhumación (los más representados en la muestra utilizada en este trabajo; la extracción del único individuo asignado al Holoceno tardío – AS22 – resultó contaminada) se constata una correlación negativa entre antigüedad y tasa de éxito ($r_s = -0,563$, $p = 0,036$), es decir, a mayor antigüedad, menor éxito. Es importante enfatizar que la muestra de Arroyo Seco 2, con restos fechados en 7800 años, se cuenta entre las más antiguas estudiadas en América, habiendo muy pocos análisis de ADN de restos humanos más antiguos en nuestro continente (*i.e.* la muestra de Windover, Florida, EE.UU., con fechados máximos de 8100 A.P. – Hauswirth et al. 1994; un individuo de Pyramid Lake, Nevada, EE.UU., datado en 9200 años A.P. – Kaestle y Smith 2001). A nivel mundial, si bien se ha analizado ADN no humano de entre 10.000 y 100.000 años de antigüedad (incluyendo restos de mamut, oso de las cavernas, caballo, bisonte y bacterias conservadas en *permafrost* – véase una lista actualizada en Willerslev y Cooper 2005), los análisis de ADN de restos de humanos u homínidos correspondientes al Pleistoceno tardío y Holoceno temprano son escasos, agotándose la lista en restos de *Homo neandertalensis* europeos (30.000-45.000 años A.P. – Krings et al. 1997, 2000; Ovchinnikov et al. 2000; Schmitz et al. 2002; Lalueza-Fox et al. 2005), *Homo sapiens* europeos (24.000 años A.P. – Caramelli et al. 2003) y los muy debatidos *Homo sapiens* australianos (aproximadamente 60.000 años A.P. – Adcock et al. 2001).

Ocho muestras de 23, es decir una tercera parte, dieron resultado negativo a la hora de confirmar la presencia de ADN en el extracto. La razón de los primeros fracasos fue aparentemente un problema relacionado con los reactivos utilizados, pues el recambio de los mismos, tras el fracaso de dos tandas de extracción, dio por resultado que volviera a haber extracciones exitosas. Como fue mencionado en resultados, no hay una clara influencia del material utilizado (hueso o dentina) ni la masa de partida en el éxito de las extracciones, si

bien es nuestra impresión que en general el hueso rinde mejores resultados, en parte porque en la pieza dentaria, la cantidad de polvo extraíble es limitada. Como puede verse en resultados, otros factores como antigüedad del entierro y sexo y edad del individuo no se asocian en forma significativa con el éxito o fracaso de la extracción.

Trabajo	Individuos de partida	Extracciones exitosas	Cronología aproximada (años A.P.)	% éxito
Monsalve et al. 1996	6	6	1500-100	100
Parr et al. 1996	47	28	1550-600	60
Ribeiro-Dos-Santos et al. 1996	26	18	4000-500	69
Lalueva et al. 1997	75	60	5000-100	80
Stone y Stoneking 1998	152	108	650	71
González-Oliver et al. 2001	28	25	1350-400	89
Kaestle y Smith 2001	48	29	9200-350	60
Lalueva-Fox et al. 2001	27	18	1300-300	67
Moraga et al. 2001	42	32	2900-600	76
Lalueva-Fox et al. 2003	47	15	4700-1600	32
García-Bour et al. 2004	60	24	5000-100	40
Dejean et al. 2005	21	16	1500-1000	76
Moraga et al. 2005	83	61	3900-500	73
Presente trabajo*	16	6	7800-6300	37,5

Tabla 21: Tasa de éxito en diversos trabajos sobre muestras antiguas americanas. Por extracciones exitosas se entiende muestras de las que se extrajo ADN endógeno con una determinación de haplogrupo. No se incluyen trabajos en los cuales las muestras analizadas no rindieron ADN amplificable (e.g. Merriwether et al. 2000).

(*) Si bien la tasa de éxito de Arroyo Seco 2 es de 8 en 23 (35%), se incluyen en esta tabla las muestras del período mayoritariamente representado (7800-6300 A.P.), excluyéndose la única muestra del Holoceno tardío (AS22) y las carentes a la fecha de asignación cronológica (AS8, AS14, AS25, AS30, AS41, AS43 y AS44)

5.2. Análisis de variabilidad y afinidad con otras poblaciones

5.2.1. Poblaciones modernas

Las discrepancias observadas entre los patrones de agrupación de Arroyo Seco 2 con las poblaciones modernas utilizadas para comparación se pueden atribuir, en parte, a los métodos mismos: El AMOVA (Tabla 19) fue realizado sobre una matriz de distancias euclídeas entre las poblaciones generada por el mismo *software*, lo cual explica la relativa coincidencia de los patrones “preferidos” en el AMOVA con las agrupaciones observadas en los dendrogramas derivados de las distancias D_R (Figuras 16 y 17), que constituyen una adaptación de la distancia euclídea clásica. Más allá de estas discrepancias, sin embargo, hay algunas agrupaciones que se pueden observar repetidas por todos los métodos: La muestra de

Arroyo Seco 2 presenta evidentes afinidades con las poblaciones de la Patagonia y las poblaciones de habla araucana de Chile. Asimismo, y en segunda instancia, algunos grupos del Chaco, especialmente Pilagás y Chorotes, aparecen como próximos a Arroyo Seco 2 en cuanto a su composición genética.

Para discutir las agrupaciones observadas de Arroyo Seco 2, es conveniente hacer un repaso del panorama indígena de la Argentina hacia la época del descubrimiento, especialmente en lo que respecta a las áreas culturales ocupadas por los grupos empleados para comparación en este trabajo (Pampa, Patagonia y Chaco).

Pampa: La identidad étnica de los ocupantes de la Pampa al momento del contacto con el europeo es tema de discusión, y el tema se hace especialmente complejo considerando que al momento de las primeras crónicas, el área estaba siendo sujeto a procesos de cambio cultural relacionado con el contacto hacia el sur con grupos Tehuelches y hacia el oeste con grupos Araucanos, sumada a la introducción del ganado con la llegada del europeo, que cambiaron radicalmente el estilo de vida de los indígenas de la Pampa. Ottonello y Lorandi (en Hernández 1992) afirman que en el área pampeana, hacia el siglo XVIII, sobrevivían los siguientes grupos:

- Pampa, Puelche, Carayhet o Pampa Carayhet, en el norte de la Provincia de Buenos Aires;
- Serranos o Pampas Serranos (llamados también Chechehet, Tehuelhet, Leuvuche o Dihuihet), en la zona de Tandil y la Ventana en la actual provincia de Buenos Aires, y las serranías hasta la actual provincia de Río Negro;
- Giocos y Chiquillanes en el Sur de Mendoza;
- Pampa cordobeses, cuyanos o Picunche;
- Ranqueles;
- Tehuelches septentrionales o Gunun a ken, que habitaron el norte de la Patagonia y la Pampa, ocupando las actuales provincias de Río Negro, La Pampa y el Sur de la provincia de Buenos Aires;
- Pampa salineros o Llamalmache.

De estos grupos, se considera que los Picunche y los Pampa salineros provenían del actual territorio de Chile.

Patagonia: Canals Frau (en Hernández 1992) caracteriza a los habitantes de la Patagonia como Patagones del norte, o Puelche-guénaken, y los Patagones del sur, Chónik o

Tshon. Los patagones del sur estarían conformados por tres grupos distinguibles desde el punto de vista lingüístico, de los cuales uno, los Tehuelche, alcanzaron el territorio pampeano en tiempos históricos. Los Tehuelche pueden dividirse en meridionales (Aonik'enk) y los septentrionales (Páyneken o Gunun a ken), estos últimos ya mencionados al referirnos a la Pampa.

Se calcula que la penetración del pueblo Mapuche en el actual territorio argentino, proceso denominado “araucanización” (Los Mapuche – “Gente de la Tierra” en su propia lengua – fueron llamados Araucanos por los conquistadores desde los primeros contactos), habría comenzado antes del siglo XVI. Un pueblo de horticultores asentados en el sur chileno habría comenzado en una fecha imprecisa el cruce de la cordillera hacia la Patagonia septentrional, expandiéndose luego por la Patagonia y la Pampa. Los Tehuelches septentrionales habrían sido los primeros en ser sometidos por los Mapuche, en tanto que es claro que los Pampas serranos estaban bajo la influencia de los Mapuches hacia el siglo XVIII (Hernández 1992). Por añadidura, la etnografía de la Pampa y la Patagonia describe hacia el siglo XVIII una serie de parcialidades fuertemente araucanizadas: Pewenche, Puelche, Furiloche, Chaziche, Leufuche, Rankulche, Pikunche y Wiliche. Los actuales Mapuches de la Argentina son descendientes de estas parcialidades, que fueron derrotadas militarmente recién en 1885 en las campañas dirigidas por el general Julio Argentino Roca (Hernández 2003).

Los Mapuche eran pueblos horticultores, pero la adopción del caballo introdujo importantes cambios en la economía y movilidad de los grupos de la Pampa y la Patagonia, y como se verá, también en los grupos chaqueños. La conquista de Chile se convirtió en un proceso largo y sangriento en parte por la adopción hacia mediados del siglo XVI del caballo por parte de los Mapuche chilenos. Éstos, al igual que los Pampas (que domesticaron a la prole de los caballos abandonados por la expedición del primer fundador de Buenos Aires, Pedro de Mendoza) y más tarde los Tehuelche se convirtieron en expertos jinetes, empleando al caballo como vehículo, alimento y elemento de combate (Palermo 2000).

Chaco: El Chaco es una extensa llanura que podría verse como la prolongación hacia el norte de las planicies pampeanas. Limita con los ríos Paraguay y Paraná al este, con los últimos desprendimientos del Planalto brasileño al norte, al oeste con la precordillera de los Andes y al sur con las sierras de Córdoba. Se considera que el poblamiento de esta región habría sido tardío en virtud de la anegación del territorio hasta el octavo milenio antes del presente (considérese que las primeras evidencias de poblamiento de la Patagonia se remontan al duodécimo milenio A.P.). Las culturas presentes en la región previamente a la llegada del

—

europeo se han clasificado – de un modo a veces arbitrario – a partir de los grupos lingüísticos conocidos en tiempos recientes en tres tipos: Los chaquenses típicos, integrados por los grupos pertenecientes a las familias lingüísticas Mbayá-Guaycurú y Mataco-Mataguayo; las culturas “andinizadas” de la familia Lule-Vilela y las culturas amazónicas de habla Tupí-Guaraní y Arawak. Nos centraremos en los chaquenses típicos, pero es claro a partir del panorama observado en tiempos históricos que el Chaco es, tanto cultural como geográficamente, una zona de “transición” entre la zona amazónica y la pampeana, capaz de recibir además influencias andinas (Demarchi et al. 2001b).

La familia lingüística *Mbayá-Guaycurú* está representada por los grupos Guaycurú, Abipón (actualmente extintos), Pilagá, Toba y Mocoví. En general eran grupos cazadores-recolectores con horticultura incipiente (Hernández 1992) que adoptaron rápidamente el caballo, volviéndose hábiles jinetes. En cuanto a la familia lingüística *Mataco-Mataguayo*, estaba integrada por diez grupos de los cuales perviven únicamente los Wichí (o Mataco), los Chorote y los Chulupí. Eran también cazadores-recolectores, y si bien hay noticias de que adoptaron el caballo hacia el siglo XVIII, no se destaca este hecho en las crónicas de la época pues no lo habrían empleado como instrumento de combate en la forma que sí lo hicieron las parcialidades de habla Guaycurú, que hostigaron permanentemente los asentamientos coloniales.

Los grupos pampeano-patagónicos por un lado, y los chaqueños por el otro, fueron los últimos grupos en ser sometidos al dominio del criollo. El Chaco y el sur del territorio argentino constituían el “desierto” que debía ser ocupado por poblaciones civilizadas. La lucha contra el indígena la había empezado el colono desde su llegada, pero la concluyó el criollo en dos series de campañas militares. En el sur, la férrea resistencia de varias parcialidades araucanizadas fue finalmente quebrada en 1885, en tanto que en el Chaco, las campañas organizadas a partir de la década de 1880 y concluidas en 1911 encontraron menor resistencia. La “incorporación” y el “disciplinamiento” (caracterizado por eventos de represión y matanzas) fueron características de la conquista del Chaco, en tanto que el “exterminio” fue la marca de la conquista del sur del país (Hernández 1992).

Como puede observarse en la tabla 22, en la actualidad las parcialidades indígenas de la Pampa y la Patagonia son principalmente Mapuche, y en la zona fronteriza sur de la cordillera de los Andes no hay modo de distinguir Mapuche “argentinos” de “chilenos” (éstos últimos suman en la actualidad unas 600.000 personas). Por otro lado los Tehuelche que aún se identifican como tales son meridionales (Aoni'ken) y la mayoría de ellos hablan la lengua

Mapuche (Hernández 1992). En el Chaco por otra parte, aún se distinguen varias parcialidades distintas, a veces viviendo unas a pocos kilómetros de otras.

En virtud de los cambios sufridos por los grupos de la Pampa y la Patagonia a lo largo de tiempos históricos y prehistóricos, la agrupación consistente de la muestra de Arroyo Seco 2 con los grupos chilenos y Tehuelche actuales – excluyendo

<i>Pueblo indígena</i>	<i>Población</i>
Chaco	
Wichí	24000
Chorote	1200
Chulupí	2800
Toba	39000
Mocoví	9800
Pilagá	27700
Chiriguano, Chané	2600
Pampa-Patagonia	
Mapuche	36700
Tehuelche	500

Tabla 22: Grupos indígenas censados en el Chaco y la Pampa-Patagonia en 1977. Modificado de Hernández (1992).

en forma permanente a la muestra Mapuche de la Argentina (Figuras 13, 14, 16 y 17) – es intrigante. En primer lugar llama la atención la distancia de la muestra Mapuche de Argentina respecto de las demás muestras mencionadas: Los actuales Mapuches del territorio argentino debieran tener una composición genética semejante a la de los grupos chilenos (lo cual se ha visto, por ejemplo, en que Mapuche argentinos y chilenos comparten frecuencias muy bajas de DI*A [Goicoechea et al. 2000] y frecuencias generalmente altas de L*M, especialmente del haplotipo L*Ms [Carnese et al. 1996]). En este sentido, vale la pena mencionar que el análisis de varios polimorfismos proteicos dan una clara semejanza de Tehuelches y Mapuches argentinos, debido al prolongado contacto entre los mismos (Goicoechea et al. 2001) o un posible origen común. Sin embargo, este patrón de semejanza se rompe al momento de analizar las frecuencias de haplogrupos de ADNmt indígenas en ambas muestras: Los Mapuche chilenos presentan ausencia de haplogrupo A y baja frecuencia de haplogrupo B (circunstancia que los asemeja a los otros grupos chilenos y a los grupos de Tierra del Fuego), en tanto que en Mapuche argentinos el haplogrupo A está presente y el haplogrupo B es el más frecuente (*cf.* tabla 13). Una posible explicación de esta radical diferencia podría hallarse en el mismo proceso de araucanización: Es posible que en el transcurso del sometimiento de parcialidades indígenas de la Patagonia y la Pampa, la incorporación de mujeres de las parcialidades sometidas haya sido una práctica corriente. Si ese fuera el caso, el ADN mitocondrial de los Mapuche de la Argentina sería el reflejo distorsionado de las frecuencias de haplogrupos presentes en los grupos pampeanos y patagónicos antes de ser araucanizados.

En segundo lugar está la cuestión de la semejanza de la muestra de Arroyo Seco 2 en primer término con los grupos chilenos, que representan en teoría un *stock* andino tanto desde el punto de vista morfológico (*cf.* Imbelloni 1957) como de acuerdo al análisis de grupos

sanguíneos (Luiselli et al. 2000) y ADN mitocondrial (Lewis et al. 2005; Cabana et al. 2006). Una posible explicación de estas semejanzas sería que los habitantes del Sur Chile deban en forma previa su composición genética a un tronco común, del que Arroyo Seco 2 sería representante. En síntesis, la pertenencia de Mapuches – tanto argentinos como chilenos – Huilliches, Pehuenches y Tehuelches a un tronco aborigen común explicaría las semejanzas halladas a nivel de polimorfismos proteicos, y la divergencia a nivel de ADN mitocondrial de los Mapuche argentinos se explicaría por flujo génico durante la invasión de éstos al actual territorio argentino.

Sin embargo, esta explicación deja pendientes dos problemas. En primer lugar están las diferencias observadas a nivel craneométrico entre Mapuches y Tehuelches: Los primeros serían ándidos y los segundos pámpidos *sensu* Imbelloni (1957). Si se acepta el origen de ambos pueblos en un *stock* común, se debe atribuir las diferencias a nivel de morfología craneana como adaptaciones post-poblamiento. En segundo lugar, si Arroyo Seco 2, Mapuches chilenos y Tehuelches representan un mismo contingente, ¿por qué las poblaciones pampeanas posteriores a Arroyo Seco 2 y anteriores a la araucanización tendrían frecuencias de haplogrupos distintas al punto de modificar las de los Mapuches invasores? Para explicarlo habría que postular una sustitución poblacional entre el Holoceno Temprano y el Holoceno Tardío: Sobre la base de datos craneométricos y análisis de series de fechados de carbono 14, Barrientos y Pérez (2002) postulan un reemplazo de poblaciones en sudeste pampeano en algún punto entre 6.000 y 3.500 A.P, reemplazo debido a un evento de contracción / extinción de la población local, seguido por una expansión / dispersión por otra población hacia el área. Lo que no se ha establecido es de dónde provendría esa población, pero si asumimos que las poblaciones pampeanas sometidas e incorporadas por la araucanización tendrían frecuencias de haplogrupos semejantes a las observadas en los Mapuches de la Argentina, y sabiendo que los grupos del sur del continente carecen de los haplogrupos A y B, hemos de aceptar regiones al norte de la Pampa como zona de origen para los nuevos pobladores.

En cuanto a los grupos chaqueños, si bien presentan una agrupación con Arroyo Seco 2 de acuerdo con el AMOVA (Tabla 19), al momento del análisis de distancias este patrón se desdibuja: Pilagás y Chorotes se aproximan en forma más consistente con Arroyo Seco 2, en tanto que Matacos y Tobas agrupan lejos entre sí y con diferentes afinidades: Los Tobas tienden a aproximarse a los grupos Guaraníes, y los Matacos hacia el *cluster* Aymara-Atacameño. Este patrón se vuelve aún más desconcertante si tomamos en cuenta la historia y afinidades culturales de estos grupos (ver *supra*): Tobas y Pilagás pertenecen a la familia

Guaycurú, y Matacos y Chorotes pertenecen a la rama Mataco-Mataguayo. Tras la llegada del europeo, los grupos Guaycurú adoptaron con rapidez el caballo y aumentaron su movilidad y belicosidad, cosa que no hicieron los grupos Mataco. Dejean et al. (2004), empleando loci de microsátélites, constataron una baja diferenciación interpoblacional debida probablemente a una historia de alto flujo génico entre los distintos grupos chaqueños, atribuida en parte al cambio en el patrón de movilidad observado en tiempos históricos.

Ni la afinidad lingüística ni la baja diferenciación interpoblacional se constatan en nuestro análisis de distancias, ya que puede verse que ambos grupos lingüísticos chaqueños quedan fragmentados y con afinidades dispersas. Goicochea et al. (2001) constatan que, en efecto, al analizar el ADN mitocondrial, las afinidades evidentes al analizar distancias entre grupos chaqueños mediante proteínas se desdibujan, y lo explican por las características de sensibilidad a factores estocásticos del ADNmt. Esto quedaría de hecho confirmado por Cabana et al. (2006) que al ampliar su muestra para contrastar la hipótesis propuesta por Demarchi et al. (2001b) constatan que el la diversidad intrapoblacional y la diferenciación interpoblacional se encuentran dentro de los rangos de otros grupos de poblaciones de América del Sur.

Por último, debe mencionarse que las afinidades de Arroyo Seco 2 con las poblaciones modernas mencionadas (chilenas, pampeano-patagónicas y chaqueñas) no tiene soporte firme: Al utilizar las distancias D_A de Nei (Figuras 13-15), el *cluster* de Arroyo Seco 2 con los grupos Araucanos y Tehuelche, visible al emplear D_R (Figuras 16-18), se desdibuja, a la vez que en los dendrogramas el soporte por *bootstrap* es menor al 50%. Esto indica a las claras que las relaciones observadas son sugestivas, pero no son consistentes al grado de poder afirmar en forma terminante una continuidad entre Arroyo Seco 2 y los grupos Araucanos, mucho menos con los Chaqueños.

5.2.2. Poblaciones antiguas

Ambas distancias, la D_A de Nei y la D_R de Rogers, dan cuenta de dos *clusters* consistentes, robustos y netamente separados: Uno, consistente en Tierra del Fuego y Patagonia, y otro, que une las muestras del Norte de Chile con la Amazonia y Colombia. La unión de Tierra del Fuego y Patagonia, con un soporte por *bootstrap* mayor al 50% en todos los casos (Figuras 19, 20, 22 y 23), puede atribuirse sin duda a la ausencia en ambas poblaciones de los haplogrupos A y B, ausencia que llevó a Lalueza et al. (1997) a afirmar

que estos indígenas descienden de una primera oleada de poblamiento, y que la clina latitudinal observada en los haplogrupos A y B se debe al aporte por flujo génico de estos haplogrupos por oleadas posteriores. Analizando secuencias de los haplogrupos C y D del sur chileno, Moraga et al. (2000) no encuentran mutaciones exclusivas en muestras fueguinas, por lo que consideran que la fijación de los haplogrupos C y D en Tierra del Fuego se debería a un efecto fundador durante el poblamiento del sur del continente. El trabajo de García-Bour et al. (2005) empleando microsátélites del cromosoma Y tiende sin embargo a reafirmar la visión de Lalueza et al. (1997).

Desde el punto de vista morfológico, los fueguinos presentan una serie de particularidades que han llevado a la discusión en cuanto a la posibilidad de que sean representantes de un contingente poblacional separado del que pobló el resto del continente. Mientras que Marcellino y Colantonio (1994) consideran sobre análisis craneométricos que esto es improbable, González-José (2003) concluye que los grupos fuego-patagónicos estarían vinculados a poblaciones “paleoamericanas” no mongolizadas que habrían ingresado a través del Estrecho de Bering por vía costera (aprox. 13.500 A.P.) antes que los contingentes de morfología claramente mongoloide. La existencia de esta oleada paleoamericana fue planteada también por Powell y Neves (1999) y Powell et al. (1999) sobre la base de análisis morfológicos de los restos más antiguos del continente americano (aprox. 10.000 A.P.).

El *cluster* que agrupa a las poblaciones del norte chileno con muestras provenientes del extremo norte del subcontinente (Colombia y Amazonia) (véase especialmente figuras 19, 20 y 22) tiene un soporte cercano al 50%; asociaciones similares llevaron a Moraga et al. (2001, 2005) a afirmar, en consonancia con Rothhammer y Silva (1989, 1992), que las poblaciones andinas serían el resultado del desplazamiento hacia el oeste del mismo contingente que alcanzó la actual floresta amazónica. Debe destacarse que, restándole apoyo a la hipótesis mencionada, análisis de polimorfismos proteicos y grupos sanguíneos (Luiselli et al. 2000), ADN mitocondrial (Fuselli et al. 2003; Lewis et al. 2004) y cromosoma Y (Tarazona-Santos et al. 2001), coinciden en señalar a la región andina como una zona con alto flujo génico intrarregional y en gran medida separado del resto de América del Sur a lo largo de la mayor parte de su historia.

Alrededor de y entre estos dos racimos definidos, cinco poblaciones forman dos agrupaciones con un comportamiento variable según el análisis aplicado, pero con algunas regularidades. El primer grupo es el par de poblaciones de las Antillas (Taíno y Ciboney), cuyos orígenes pueden rastrearse en teoría al norte del subcontinente sudamericano.

Curiosamente, sin embargo, todos los métodos excepto el análisis de *clusters* por D_A empleando *bootstrap* de individuos (Figura 19) y el análisis de coordenadas principales usando D_R (Figura 24) tienden a aproximar a estas poblaciones a las muestras de Tierra del Fuego y Patagonia. No es el objetivo de este trabajo discutir este fenómeno, si bien podemos arriesgar que sería un agrupamiento espurio generado por fenómenos microevolutivos locales en ambos pares de poblaciones.

Sí es de importancia para este trabajo discutir el comportamiento de las tres poblaciones restantes: El noroeste argentino (NOA), Arroyo Seco 2 y el grupo de muestras individuos hasta ahora denominadas como “Uruguay este”, que corresponden a restos enterrados en los “cerritos de indios” de esa región. Mientras que las distancias D_A de Nei tienden a agrupar estas muestras más cerca del *cluster* amazónico/andino que del fuego-patagónico (Figuras 19-21), las distancias D_R las aproximan – en forma bastante errática y en general inconsistente – a este último *cluster* (Figuras 22-24). Esta inconsistencia se puede atribuir muy probablemente al manejo dado primero por los análisis de distancias y luego por el procedimiento de agrupación a las frecuencias génicas: La muestra del este uruguayo tiene frecuencias intermedias de los haplogrupos B y C y bajas frecuencias de los haplogrupos A y D. Según el método, dada la ausencia de una semejanza clara, tiende a ser agrupada o bien con un extremo, o bien con el otro del subcontinente.

Ahora bien, ¿cómo se explican estas frecuencias en el marco de la información que nos brindan las otras agrupaciones? Podríamos atribuir las frecuencias intermedias al efecto de la latitud, que pone a las poblaciones de Arroyo Seco 2, Uruguay y el noroeste argentino en una posición intermedia respecto a los dos contingentes mencionados y resultando por lo tanto una amalgama de ambos. Esto es plausible en el caso de las muestras del NOA y del este uruguayo, que son del Holoceno tardío, para cuando el continente ya llevaba largo tiempo poblado. Para Arroyo Seco 2, que se remonta al Holoceno temprano, es de todos modos posible pensar en un panorama semejante si consideramos que la región Pampeana ya llevaría 3000 años poblada, con indicios de contacto de larga distancia con el actual territorio uruguayo (Flegenheimer et al. 2001).

Sin embargo, las relaciones observadas entre las tres poblaciones antiguas “intermedias” sobre la base de haplogrupos mitocondriales son difíciles de conciliar con los análisis disponibles a nivel craneométrico para el Cono Sur. Los trabajos de Marcellino y Colantonio (1994, 1995) y Colantonio y Marcellino (1995), basados en colecciones arqueológicas provenientes de Argentina, Bolivia y Chile, establecen, utilizando distancias de

Mahalanobis para diversas variables métricas, cuatro agrupaciones consistentes para las poblaciones prehistóricas empleadas: 1) noroeste argentino y norte chileno; 2) la franja central de la Argentina, que se agrupa con muestras prehispánicas inmediatas de Perú y Bolivia; 3) la Pampa, la Patagonia y el Delta del río Paraná; 4) Tierra del Fuego. Si los tomamos por su distancia con las muestras de Tierra del Fuego, las agrupaciones más semejantes son, en igual medida, la franja central y Pampa-Patagonia-Delta, y la más alejada es la del NOA y Chile. Marcellino y Colantonio (1995) observan que se distinguen claramente tres perfiles a partir de estas agrupaciones: Uno “andino” (agrupación 1), uno “patagónico” (agrupación 3, con el cual la agrupación 2 tendría afinidades) y uno “fueguino” (agrupación 4), con una fuerte asociación entre las agrupaciones y su proximidad geográfica. Destacan que estos resultados serían “expresivos de un proceso biológico dinámico que no alcanzó a desdibujar – entrecruzamientos mediante – las diferencias morfológicas que pudieron caracterizar a los distintos grupos humanos primigenios en el subcontinente” (Marcellino y Colantonio 1995: 25).

Los datos de estos análisis, junto con los de Neves y Cocilovo (1989), fueron empleados para analizar las afinidades de las poblaciones prehistóricas del Uruguay. Ese análisis arrojó entre otras cosas, afinidad entre las poblaciones prehistóricas del este uruguayo con las representadas en los restos asociados con sambaquíes en el Brasil, y una relación menos precisa pero no por ello menos notable de los restos recuperados en la región Oeste con los provenientes de sitios con túmulos del Delta del Paraná (Bertoni et al. 2000). Este trabajo enfatiza que las similitudes en la morfología no necesariamente reflejan relaciones genéticas (origen común o flujo génico) ya que si bien pueden deberse a un origen biológico común, también pueden corresponder a contactos entre grupos, y, en el caso específico de la morfología, a adaptaciones ambientales, por lo que denotarían distancias geográficas.

En síntesis, mientras que los análisis craneométricos dan cuenta de un patrón claramente geográfico y en ocasiones relacionado con modos de subsistencia, los análisis basados en frecuencias de haplogrupos mitocondriales dan otra pauta de agrupación menos clara, y donde Arroyo Seco 2 no agrupa en forma consistente con ninguna de las agrupaciones más robustas. Esta inconsistencia puede deberse a que si Arroyo Seco 2 se encontrara en una zona de tránsito, los individuos allí recuperados reflejarían ese carácter y el flujo génico existente en el área, tanto con grupos del sur como, eventualmente, del norte, tanto del Uruguay como de la zona andina. Igualmente, algunos agrupamientos pueden deberse a un problema de método: La relación de Arroyo Seco 2 con el NOA, visible en los dendrogramas,

es sorprendente en virtud de que las frecuencias de haplogrupos del NOA (Tabla 14) no se asemejan prácticamente en nada a las de Arroyo Seco 2. Es posible que esta relación sea un efecto del intento del algoritmo del método *neighbor-joining* por ubicar en el árbol a poblaciones que no se agrupan claramente con ningún *cluster*, y de hecho, los análisis de coordenadas principales coinciden en alejar a ambas poblaciones entre sí.

En cuanto a las relaciones ambiguas de Arroyo Seco 2 con el *cluster* amazónico/andino y con el *cluster* Tierra del Fuego-Patagonia, tanto los estudios craneométricos tipológicos como los basados en análisis multivariados refuerzan la agrupación observada en el AMOVA (Tabla 20) y a empleando la distancia D_R (Figuras 22, 23 y 24) con las poblaciones del sur del subcontinente sudamericano: Arroyo Seco 2 se encuentra en la región atribuida por Imbelloni (1938, 1957) a los grupos “pámpidos”, que habrían llegado hasta la Patagonia. El perfil “patagónico” de Marcellino y Colantonio (1995) por su parte presenta mayor semejanza con Tierra del Fuego que con los demás perfiles. Si aceptamos que Tierra del Fuego y Patagonia fueron pobladas por grupos llegados en forma temprana al continente americano, Arroyo Seco 2 representaría una población pampeana temprana emparentada con estos grupos, y ya posiblemente influenciada por las oleadas poblacionales posteriores. Si por otra parte rechazamos con Moraga et al. (2000) esta visión, la composición genética de Arroyo Seco 2 representa fenómenos relacionados con la variabilidad de las poblaciones americanas, variabilidad que habría empezado a desarrollarse *a posteriori* del poblamiento.

6. Conclusiones y perspectivas

A continuación se enumeran las principales conclusiones de la presente tesis:

1. En el presente trabajo se logró extraer ADN amplificable y secuenciable a partir de muestras de una antigüedad que alcanza los 7800 años (el individuo AS24, cuyo diagnóstico por RFLP como perteneciente al haplogrupo C fue corroborado por el análisis de secuencias, tiene un fechado de 7800 +/- 115 años A.P. [AA-9046]). La tasa de éxito alcanzada (35%) está dentro de lo esperado dada la antigüedad de la muestra, y la información que arroja el ADN extraído nos brinda un primer acercamiento a una población del Holoceno temprano de la región Pampeana.
2. En lo que respecta a la composición genética de la población prehistórica de Arroyo Seco 2 y las conclusiones derivadas de las comparaciones realizadas con la misma, es de rigor prologar las mismas con la advertencia de que, en virtud de el tamaño de la muestra, su representatividad puede ser discutida. Por lo tanto, deberá ahondarse en estas conclusiones utilizándolas como hipótesis a contrastar. Hecha esta advertencia, destacaremos los resultados más notables, las conclusiones provisionales que de ellos se derivan y los posibles caminos futuros.
 - 2.1. En primer lugar, en lo referente a las relaciones de continuidad-discontinuidad de las poblaciones pampeanas del Holoceno temprano con las que poblaron la región pampeana en tiempos históricos, se constata que éstas son posibles representantes de un *stock* común que habría poblado el Cono Sur sudamericano en el límite Pleistoceno-Holoceno, y de las cuales en principio serían descendientes las poblaciones del sur de Chile, Tierra del Fuego, Patagonia y posiblemente el actual Chaco argentino, caracterizadas por la baja o nula frecuencia del haplogrupo A y frecuencias bajas a intermedias del haplogrupo B. Es prematuro hacer suposición alguna sobre si estas poblaciones representan o no un contingente poblacional discreto (una “oleada”, si utilizamos la terminología empleada en el debate aún no resuelto sobre el poblamiento del Nuevo Mundo), pero es claro que hacia el norte, las poblaciones de la actual floresta amazónica tienen una composición claramente diferente.
 - 2.2. Las poblaciones que ocuparon la región Pampeana y reemplazaron a la población original durante el Holoceno tardío probablemente se habrían originado al norte de la

misma, y habrían luego pasado por diversas interacciones con grupos procedentes del sur (“tehuelchización”) y de más allá de la cordillera (“araucanización”) más o menos en la época de la llegada del europeo. Las frecuencias de los cuatro haplogrupos de las poblaciones pampeanas del Holoceno tardío serían intermedias, y es posible que restos de la composición genética de estos grupos se mantengan hasta hoy día en los Mapuche de la República Argentina. Una comparación de secuencias de las regiones hipervariables del ADNmt de Mapuche chilenos, Mapuche argentinos y Arroyo Seco 2 sería extremadamente informativa a este respecto.

2.3. La situación del noroeste argentino, el Chaco y el territorio uruguayo respecto a Arroyo Seco 2 por un lado y las diversas poblaciones sudamericanas analizadas por el otro es, por otra parte, menos clara. A nivel de ADNmt, las poblaciones chaqueñas se “fragmentan” hallando afinidad con distintos agrupamientos de poblaciones y el noroeste argentino y el este uruguayo no agrupan en forma consistente ni con el norte ni con el sur del subcontinente. Esto puede deberse a la situación “limítrofe” de estos territorios respecto al resto del continente, no ocupando zonas limitadas por barreras geográficas y por ello capaces de recibir aportes virtualmente de cualquier dirección. También en este caso, el análisis detallado de secuencias de HVR y, en el caso de las muestras del noroeste argentino y Uruguay, un aumento en el tamaño muestral pueden aclarar esta incertidumbre.

* * *

La información potencial contenida en los ADN extraídos del sitio Arroyo Seco 2 es extremadamente valiosa, en la medida en que representa una imagen de la composición genética de las poblaciones Pampeanas del Holoceno temprano. Una mejor capitalización de esta información implicaría: a) La ampliación de la muestra, repitiendo la extracción de ADN de los casos fallidos; b) un análisis de mayor resolución del ADN mitocondrial, incluyendo secuencias de la región hipervariable y eventualmente regiones informativas del resto del genoma mitocondrial; c) la ampliación del análisis a *loci* nucleares, constituyendo el análisis exitoso de diversos microsatélites del cromosoma Y (García-Bour et al. 2004; Dejean et al. 2005) una perspectiva prometedora al respecto.

A partir de este “estándar” de una población posiblemente representante de las primeras en América, hay una serie de interrogantes que pueden ser respondidas:

- A nivel del proceso de poblamiento del continente americano, la diversidad que arrojen las secuencias de HVR y eventuales loci nucleares pueden dar cuenta de la tasa de diversidad de las poblaciones tempranas a aproximadamente cuatro milenios de su llegada, y con ello de la intensidad del cuello de botella postulado para el pasaje de los primeros pobladores a través del Estrecho de Bering.
- A nivel de la dinámica poblacional de la región Pampeana, el análisis de otras muestras antiguas de la región, pertenecientes a momentos posteriores, nos permitiría ahondar en la dinámica poblacional de la misma y dar respuesta a hipótesis de reemplazo planteadas sobre datos provenientes de otras fuentes (fechados, análisis morfológicos y datos paleoclimáticos).
- A nivel de las relaciones con regiones vecinas, se incluye la información potencial que brinda Arroyo Seco 2 para la historia de las poblaciones del territorio uruguayo: Los análisis morfológicos de restos prehistóricos uruguayos (Bertoni et al. 2000) dan lugar a la posibilidad de que las poblaciones prehistóricas del este y el oeste del Uruguay sean descendientes de distintos troncos poblacionales. La continuación de los análisis en curso de restos prehistóricos uruguayos (la mayoría del Holoceno tardío) y su comparación con una muestra del Holoceno temprano permitiría trazar un panorama de la historia poblacional de nuestro territorio, en tanto caracterización de su continuidad o cambio a lo largo del tiempo y sus afinidades con regiones vecinas.

Referencias citadas

Adcock GJ, Dennis ES, Eastal S, Huttley GA, Jermin LS, Peacock WJ, Thorne A (2001) Mitochondrial DNA sequences in ancient Australians: Implications for modern human origins. *Proc Natl Acad Sci USA* 98:537-542

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Rose BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290:457-465

Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23:147

Austin JJ, Ross AJ, Smith AB, Fortey RA, Thomas RH (1997) Problems of reproducibility – does geologically ancient DNA survive in amber-preserved insects? *Proc R Soc Lond B* 264:467-474

Bailliet G, Rothhammer F, Carnese FR, Bravi CM, Bianchi NO (1994) Founder Mitochondrial Haplotypes in Amerindian Populations. *Am J Hum Genet* 54:27-33

Bandelt HJ, Herrnstadt C, Yao YG, Kong QP, Kivisild T, Rengo C, Scozzari R, Richards M, Villems R, Macaulay V, Howell N, Torroni A, Zhang YP (2003) Identification of Native American Founder mtDNAs Through the Analysis of Complete mtDNA Sequences: Some Caveats. *Ann Hum Genet* 67:512–524

Baron H, Hummel S, Herrmann B (1996) Mycobacterium tuberculosis Complex DNA in Ancient Human Bones. *J Archaeol Sci* 23:667-671

Barrientos G (1997) Nutrición y dieta de las poblaciones aborígenes prehispánicas de la región pampeana. Tesis de Doctorado inédita, Universidad de La Plata, La Plata

Barrientos G, Pérez SI (2002) Was there a population replacement during the late middle Holocene in the southeastern Pampas of Argentina? Discussing its archaeological evidence and its paleoecological basis. En: Neme G, Gil A, Zárata M (eds) *Quaternary International, Volumen especial*

Bennett LB, Shriver MD, Bowcock AM (1998) Markers and Methods for Reconstructing Modern Human History. *DNA Sequence - The Journal of Sequencing and Mapping* 8:329-341

Berón MA, Politis GG (1997) Arqueología pampeana en la década de los '90. Estado de las investigaciones y perspectivas. En: Berón MA, Politis GG (eds) *Arqueología pampeana en la década de los '90*. Museo de Historia Natural de San Rafael / INCUAPA - UNCPBA, Mendoza / Olavarría, pp 7-32

Bertoni B, Figueiro G, Cabanas G, McDonough JE, Bluteau C, Merriwether D, Sans M (2004) Primeras secuencias de ADN mitocondrial de indígenas prehistóricos del Uruguay. En: Bevide L, Barreto I, Curbelo C (eds) *X Congreso Uruguayo de Arqueología: La Arqueología Uruguaya ante los desafíos del nuevo siglo*. Asociación Uruguaya de Arqueología, Montevideo

Bertoni B, Portas M, Sans M (2000) Relaciones Morfológicas de las Poblaciones prehistóricas del Uruguay: Análisis de Restos Esqueletarios Humanos. En: Durán Coirolo A, Bracco Boksar R (eds) *Arqueología de las Tierras Bajas*. Ministerio de Educación y Cultura, Montevideo, pp 369-384

- Bonatto SL, Salzano FM (1997a) Diversity and Age of the Four Major mtDNA Haplogroups, and Their Implications for the Peopling of the New World. *Am J Hum Genet* 61:1413-1423
- Bonatto SL, Salzano FM (1997b) A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. *Proc Natl Acad Sci USA* 94:1866-1871
- Boom R, Sol CJA, Salimans MMM, Jansen CL, Wertheim van Dillen PME, van der Nordaa J (1990) Rapid and simple method for purification of nucleic acids. *J Clin Microbiol* 28:495-503
- Brown MD, Hosseini SH, Torroni A, Bandelt HJ, Allen JC, Schurr TG, Scozzari R, Cruciani F, Wallace DC (1998) mtDNA Haplogroup X: An Ancient Link between Europe/Western Asia and North America? *Am J Hum Genet* 63:1852-1861
- Brown WM (1980) Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc Natl Acad Sci USA* 77:3605-3609
- Cabana GS, Merriwether DA, Hunley K, Demarchi DA (2006) Is the genetic structure of Gran Chaco populations unique? Interregional perspectives on native South American mitochondrial DNA variation. *Am J Phys Anthropol* 131:108-119
- Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. *Nature* 325:31-36
- Cano RJ, Poinar HN, Pieniazek NJ, Acra A, Poinar Jr. GO (1993) Amplification and sequencing of DNA from a 120-135-million-year-old weevil. *Nature* 363:536-538
- Caramelli D, Lalueza-Fox C, Vernesi C, Lari M, Casoli A, Mallegni F, Chiarelli B, Dupanloup I, Bertranpetit J, Barbujani G, Bertorelle G (2003) Evidence for a genetic discontinuity between Neandertals and 24,000-year-old anatomically modern Europeans. *Proc Natl Acad Sci USA* 100:6593-6597
- Carlyle SW, Parr RL, Hayes MG, O'Rourke DH (2000) Context of Maternal Lineages in the Greater Southwest. *Am J Phys Anthropol* 113:85-101
- Carnese FR, Caratini AL, Goicoechea AS, Weimer TA, Franco MHL, Hutz MH, Salzano FM (1996) Demography and blood genetics of Argentinian Mapuche Indians. *Int J Anthropol* 11:33-42
- Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic Analysis: Models and Estimation Procedures. *Evolution* 21:550-570
- Cockerham CC (1973) Analyses of gene frequencies. *Genetics* 74:679-700
- Colantonio SE, Marcellino AJ (1995) Acerca de la taxonomía aborigen del cono sur americano: región central de Argentina. Un análisis multivariado. En: *Estudios de antropología biológica*, vol. V. Universidad Nacional Autónoma de México, México, pp 161-182
- Comas J (1974) *Antropología de los pueblos iberoamericanos*. Editorial Labor, Barcelona
- Cooper A, Mourier-Chauviré C, Chambers GK, von Haeseler A, Wilson AC, Pääbo S (1992) Independent origins of New Zealand moas and kiwis. *Proc Natl Acad Sci USA* 89:8741-8744
- Crawford MH (1998) *The origins of Native Americans: evidence from anthropological genetics*. Cambridge University Press, Cambridge.

- Dejean CB, Crouau-Roy B, Goicoechea AS, Avena SA, Carnese FR (2004) Genetic variability in Amerindian populations of Northern Argentina. *Genet Mol Biol* 27:489-495
- Dejean CB, Keyser-Tracqui C, Crubézy E, Blandin P, Dugoujon J-M, Ludes B, Avena S, Goicoechea AS, Carnese FR (2005) ADN antiguo en restos esqueléticos precolombinos de Pampa Grande, Prov. de Salta, Argentina. *Revista Argentina de Antropología Biológica* 7:58 (Resumen)
- Demarchi DA, Panzetta-Dutari GM, Colantonio SE, Marcellino AJ (2001a) Absence of the 9-bp deletion of mitochondrial DNA in pre-Hispanic inhabitants of Argentina. *Hum Biol* 73:575-582
- Demarchi DA, Panzetta-Dutari GM, Motran CC, López de Basualdo MA, Marcellino AJ (2001b) Mitochondrial DNA Haplogroups in Amerindian Populations From the Gran Chaco. *Am J Phys Anthropol* 115:199-203
- Denaro M, Blanc H, Johnson MJ, Chen KH, Wilmsen E, Cavalli-Sforza LL, Wallace DC (1981) Ethnic variation in *Hpa* I endonuclease cleavage patterns of human mitochondrial DNA. *Proc Natl Acad Sci USA* 78:5768-5772
- DeSalle R, Grimaldi D (1994) Very old DNA. *Curr Opin Genet Dev* 4:810-815
- Dillehay TD (1997) *Monte Verde: A Late Pleistocene Settlement in Chile*. Smithsonian Institution Press, Washington DC
- Dixon EJ (2002) How and When Did People First Come to North America? *Athena Review* 3:23-27. <http://www.athenapub.com/10Dixon.htm>
- Easton RD, Merriwether DA, Crews DE, Ferrell RE (1996) mtDNA variation in the Yanomami: evidence for additional New World founding lineages. *Am J Hum Genet* 59:213-225
- Efron B, Tibshirani RJ (1993) *An introduction to the bootstrap*. Chapman & Hall / CRC, Boca Raton
- Eshleman JA, Malhi RS, Smith DG (2003) Mitochondrial DNA Studies of Native Americans: Conceptions and Misconceptions of the Population Prehistory of the Americas. *Evol Anthropol* 12:7-18
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplogroups: applications to human mitochondrial DNA restriction data. *Genetics* 131:479-491
- Eyre-Walker A, Smith NH, Smith JM (1999) How clonal are human mitochondria? *Proc R Soc Lond B* 266:477-483
- Felsenstein J (2004) PHYLIP (Phylogeny Inference Package) versión 3.6. Distribuido por el autor. Department of Genome Sciences, University of Washington, Seattle. <http://evolution.gs.washington.edu/phylip.html>
- Fiedel SJ (1996) *Prehistoria de América*. Crítica, Barcelona
- Fix AG (2005) Rapid Deployment of the Five Founding Amerind mtDNA Haplogroups Via Coastal and Riverine Colonization. *Am J Phys Anthropol* 128:430-436
- Flegenheimer N, Bayón C, Baeza J, Femenías J, Valente M (2001) Relaciones tempranas (vínculos tempranos) entre grupos de la región pampeana y Uruguay. Trabajo presentado en el X Congreso Nacional de Arqueología Uruguaya, Montevideo. Montevideo

- Forster P, Harding R, Torroni A, Bandelt HJ (1996) Origin and evolution of Native American mtDNA variation: a reappraisal. *Am J Hum Genet* 59:935-945
- Fuselli S, Tarazona-Santos E, Dupanloup I, Soto A, Luiselli D, Pettener D (2003) Mitochondrial DNA Diversity in South America and the Genetic History of Andean Highlanders. *Mol Biol Evol* 20:1682-1691
- García F, Moraga M, Vera S, Enríquez H, Llop E, Ocampo C, Aspillaga E, Rothhammer F (2004) Origen y microdiferenciación de la población humana del Archipiélago de Chiloé. *Rev Chil Hist Nat* 77:539-546
- García-Bour J, Pérez-Pérez A, Álvarez S, Fernández E, López-Parra AM, Arroyo-Pardo E, Turbón D (2004) Early Population Differentiation in Extinct Aborigines From Tierra del Fuego-Patagonia: Ancient mtDNA Sequences and Y-Chromosome STR Characterization. *Am J Phys Anthropol* 123:361-370
- Garzón L (2002) La degradación natural del ADN: el concepto de edad térmica. *Rev Esp Antrop Biol* 23:65-76
- Goicoechea AS, Carnese FR, Caratini AL, Avena S, Salaberry M, Salzano FM (2000) Demography, genetic diversity, and population relationships among Argentinean Mapuche Indians. *Genet Mol Biol* 23:513-518
- Goicoechea AS, Carnese FR, Dejean C, Avena SA, Weimer TA, Franco MHL, S.M. Callegari-Jacques, Estalote AC, Simões MLMS, Palatnik M, Salzano FM (2001) Genetic Relationships Between Amerindian Populations of Argentina. *Am J Phys Anthropol* 115:133-143
- González-José R (2003) El poblamiento de la Patagonia. Análisis de la variación craneofacial en el contexto del poblamiento americano. Tesis de doctorado inédita, Universidad de Barcelona, Barcelona
- González-Oliver A, Márquez-Morfín L, Jiménez JC, Torre-Blanco A (2001) Founding Amerindian Mitochondrial DNA Lineages in Ancient Maya From Xcaret, Quintana Roo. *Am J Phys Anthropol* 116:230-235
- Gordon RG (ed.) (2005) *Ethnologue: Languages of the World*, Fifteenth edition. SIL International, Dallas. Versión *online*: <http://www.ethnologue.com/>
- Götherström A, Collins MJ, Angerbjörn A, Lidén K (2002) Bone preservation and DNA amplification. *Archaeometry* 44:395-404
- Gravlee CC, Bernard HR, Leonard WR (2003a) Heredity, Environment, and Cranial Form: A Reanalysis of Boas's Immigrant Data. *Am Anthropol* 105:125-138
- Gravlee CC, Bernard HR, Leonard WR (2003b) Boas's Changes In Bodily Form: The Immigrant Study, Cranial Plasticity, And Boas's Physical Anthropology. *Am Anthropol* 105:326-332
- Greenberg JH, Turner II CG, Zegura SL (1986) The settlement of the Americas: A comparison of the linguistic, dental end genetic evidence. *Curr Anthropol* 4:477-497
- Hagelberg E (1994) Ancient DNA Studies. *Evol Anthropol* 2:199-207
- Hagelberg E, Goldman N, Liò P, Whelan S, Schiefenhövel W, Clegg JB, K BD (1999) Evidence for mitochondrial DNA recombination in a human population of island Melanesia. *Proc R Soc Lond B* 266:485-492

- Handt O, Krings M, Ward RH, Paabo S (1996) The retrieval of ancient human DNA sequences. *Am J Hum Genet* 59:368-376
- Hänni C, Brousseau T, Laudet V, Stehelin D (1995) Isopropanol precipitation removes PCR inhibitors from ancient bone extracts. *Nucleic Acids Res* 23:881-882
- Hauswirth WW, Dickel CD, Rowold DJ, Hauswirth MA (1994) Inter- and intrapopulation studies of ancient humans. *Experientia* 50:585-591
- Hernández I (1992) *Los indios de Argentina*. MAPFRE, Madrid
- Hernández I (2003) *Autonomía o ciudadanía incompleta: El Pueblo Mapuche en Chile y Argentina*. Pehuén Editores, Santiago de Chile
- Heyer E, Zietkiewicz E, Rochowski A, Yotova V, Puymirat J, Labuda D (2001) Phylogenetic and Familial Estimates of Mitochondrial Substitution Rates: Study of Control Region Mutations in Deep-Rooting Pedigrees. *Am J Hum Genet* 69:1113-1126
- Hofreiter M, Jaenicke V, Serre D, von Haeseler A, Pääbo S (2001) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Res* 29:4793-4799
- Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima K (1993) Peopling of the Americas, founded by four major lineages of mitochondrial DNA. *Mol Biol Evol* 10:23-47
- Höss M, Pääbo S (1993) DNA extraction from Pleistocene bones by a silica-based purification method. *Nucleic Acids Res* 21:3913-3914
- Imbelloni J (1938) Tabla clasificatoria de los indios, regiones biológicas y grupos humanos de América. *Physis* 21:229-249
- Imbelloni J (1957) *De historia primitiva de América. Los grupos raciales aborígenes*. Facultad de Historia y Letras, Serie Historia de América, N° 1:71-88, Buenos Aires
- Johnson MJ, Wallace DC, Ferris DC, Rattazzi MC, Cavalli-Sforza LL (1983) Radiation of Human Mitochondrial DNA Types Analyzed by Restriction Endonuclease Cleavage Patterns. *J Mol Evol* 19:255-271
- Jones M (2003) Ancient DNA in pre-Columbian archaeology: a review. *J Archaeol Sci* 30:629-635
- Kaestle FA, Horsburgh KA (2002) Ancient DNA in Anthropology: Methods, Applications, and Ethics. *Yrbk Phys Anthropol* 45:92-130
- Kaestle FA, Smith DG (2001) Ancient Mitochondrial DNA Evidence for Prehistoric Population Movement: The Numic Expansion. *Am J Phys Anthropol* 115:1-12
- Karafet TM, Zegura SL, Posukh O, Osipova L, Bergen A, Long J, Goldman D, Klitz W, Harihara S, de Knijff P, Wiebe V, Griffiths RC, Templeton AR, Hammer MF (1999) Ancestral Asian source(s) of new world Y-chromosome founder haplotypes. *Am J Hum Genet* 64:817-831
- Kemp BM, Smith DG (2005) Use of bleach to eliminate contaminating DNA from the surface of bones and teeth. *Forensic Sci Int* 154:53-61

- Krings M, Capelli C, Tschentscher F, Geisert H, Meyer S, von Haeseler A, Grossschmidt K, Possnert G, Paunovic M, Pääbo S (2000) A view of Neandertal genetic diversity. *Nat Genet* 26:144-146
- Krings M, Stone AC, Schmitz RW, Krainitzki H, Stoneking M, Pääbo S (1997) Neandertal DNA Sequences and the Origin of Modern Humans. *Cell* 90:19-30
- Lahn BT, Pearson NM, Jégalian K (2001) The Human Y Chromosome, in the Light of Evolution. *Nat Rev Genet* 2:207-216
- Lalueza C, Pérez-Pérez A, Prats E, Cornudella L, Turbón D (1997) Lack of founding Amerindian mitochondrial DNA lineages in extinct Aborigines from Tierra del Fuego-Patagonia. *Hum Mol Genet* 6:41-46
- Lalueza-Fox C, Gilbert MTP, Martínez-Fuentes AJ, Calafell F, Bertranpetit J (2003) Mitochondrial DNA from Pre-Columbian Ciboneys from Cuba and the Prehistoric Colonization of the Caribbean. *Am J Phys Anthropol* 121:97-108
- Lalueza-Fox C, Luna-Calderón F, Calafell F, Morera B, Bertranpetit J (2001) MtDNA from extinct Tainos and the peopling of the Caribbean. *Ann. Hum. Genet.* 65:137-151
- Lalueza-Fox C, Sampietro ML, Caramelli D, Puder Y, Lari M, Calafell F, Martínez-Maza C, Bastir M, Fortea J, de la Rasilla M, Bertranpetit J, Rosas A (2005) Neandertal Evolutionary Genetics: Mitochondrial DNA Data from the Iberian Peninsula. *Mol Biol Evol* 22:1077-1081
- Lell JT, Sukernik RI, Starikovskova YB, Su B, Lin L, Schurr TG, Underhill PA, Wallace DC (2002) The dual origin and Siberian affinities of Native American Y-chromosomes. *Am J Hum Genet* 70:192-206
- Lewis CM, Tito RY, Lizárraga B, Stone AC (2004) Land, Language, and Loci: mtDNA in Native Americans and the Genetic History of Peru. *Am J Phys Anthropol* 127:351-360
- Lindahl T (1993a) Instability and decay of the primary structure of DNA. *Nature* 362:709-715
- Lindahl T (1993b) Recovery of antediluvian DNA. *Nature* 365:700
- Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell JE (2000) *Molecular Cell Biology*. W H Freeman, New York
- Luiselli D, Simoni L, Tarazona-Santos E, Pastor S, Pettener D (2000) Genetic Structure of Quechua-Speakers of the Central Andes and Geographic Patterns of Gene Frequencies in South Amerindian Populations. *Am J Phys Anthropol* 113:5-17
- Macauley V, Richards M, Sykes B (1999) Mitochondrial DNA recombination - no need to panic. *Proc R Soc Lond B* 266:2037-2039
- MacHugh DE, Edwards CJ, Bailey JF, Bancroft DR, Bradley DG (2000) The Extraction and Analysis of Ancient DNA From Bone and Teeth: a Survey of Current Methodologies. *Ancient Biomol* 3:81-102
- Malhi RS, Eshleman JA, Greenberg JA, Weiss DA, Schultz Shook BA, Kaestle FA, Lorenz JG, Kemp BM, Johnson JR, Smith DG (2002) The Structure of Diversity within New World Mitochondrial DNA Haplogroups: Implications for the Prehistory of North America. *Am J Hum Genet* 70:905-919
- Malhi RS, Smith DG (2002) Brief Communication: Haplogroup X Confirmed in Prehistoric North America. *Am J Phys Anthropol* 119:84-86

- Manly BFJ (1994) *Multivariate Statistical Methods: A Primer*. Second Edition. Chapman & Hall / CRC, London / New York
- Marcellino AJ, Colantonio SE (1994) Relaciones morfológicas de los aborígenes prehispánicos del territorio argentino. IX: Los fueguinos. *Paleoetnologica* 7:51-71
- Marcellino AJ, Colantonio SE (1995) Relaciones morfológicas de los aborígenes prehispánicos del territorio argentino. VII: La región serrana de Córdoba. *Scripta Ethnologica, Supplementa* 13:17-28
- Marota I, Basile C, Ubaldi M, Rollo F (2002) DNA decay rate in papyri and human remains from Egyptian archaeological sites. *Am J Phys Anthropol* 117:310-318
- Martínez Marignac VL, Bertoni B, Parra EJ, Bianchi NO (2004) Characterization of Admixture in an Urban Sample from Buenos Aires, Argentina, Using Uniparentally and Biparentally Inherited Genetic Markers. *Hum Biol* 76:543-557
- Massone M, Prieto A (2004) Evaluación de la modalidad cultural Fell 1 en Magallanes. *Chungará (Arica)* 36:303-315
- Meltzer DJ, Grayson DK, Ardila G, Barker AW, Dincauze DF, Vance Haynes C, Mena F, Nunez L, Stanford DJ (1997) On the Pleistocene antiquity of Monte Verde, southern Chile. *Am Antiq* 62:659-663
- Merriwether DA, Cabana GS, Reed, DM (2000) Kennewick Man Ancient DNA Analysis: Final Report Submitted to the Department of the Interior, National Park Service. http://www.cr.nps.gov/archeology/Kennewick/Merriwether_Cabana.htm
- Merriwether DA, Rothhammer F, Ferrell RE (1995) Distribution of the four founding lineage haplotypes in native Americans suggests a single wave of migration for the New World. *Am J Phys Anthropol* 98:411-430
- Monsalve MV, Cárdenas F, Guhl F, Delaney AD, Devine DV (1996) Phylogenetic analysis of mtDNA lineages in South American mummies. *Ann Hum Genet* 60:293-303
- Moraga M, Aspillaga E, Santoro C, Standen V, Carvallo P, Rothhammer F (2001) Análisis de ADN mitocondrial en momias del norte de Chile avala hipótesis de origen amazónico de poblaciones andinas. *Rev Chil Hist Nat* 74:719-726
- Moraga M, Santoro CM, Standen VG, Carvallo P, Rothhammer F (2005) Microevolution in prehistoric Andean populations: Chronologic mtDNA variation in the desert valleys of northern Chile. *Am J Phys Anthropol* 127:170-181
- Moraga ML, Rocco P, Miquel JF, Nervi F Llop E, Chakraborty R, Rothhammer F, Carvallo P (2000) Mitochondrial DNA Polymorphisms in Chilean Aboriginal Populations: Implications for the Peopling of the Southern Cone of the Continent. *Am J Phys Anthropol* 113:19-29
- Mullis KB, Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Met Enzymol* 155:335-350
- Mullis KB, Faloona FA, Scharf S, Saiki RK, Horn G, Erlich HA (1986) Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* 51:263-273

- Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. *J Mol Evol* 19:153-170
- Neves WA, Cocilovo JA (1989) Componentes craneofuncionales y microdiferenciación de las poblaciones prehistóricas del litoral centro-sur de Brasil. *Ciência e Cultura* 41:1071-1085
- Nicholas KB, Nicholas HB, Deerfield DW (1997) GeneDoc: Analysis and Visualization of Genetic Variation. *EMBNEW.NEWS* 4:14. <http://www.psc.edu/biomed/genedoc>
- O'Rourke DH, Hayes MG, Carlyle SW (2000a) Ancient DNA Studies in Physical Anthropology. *Annu Rev Anthropol* 29:217-242
- O'Rourke DH, Hayes MG, Carlyle SW (2000b) Spatial and Temporal Stability of mtDNA Haplogroup Frequencies in Native North America. *Hum. Biol.* 72:15-34
- Ota T (1993) *Dispan: Genetic distance and phylogenetic analysis*. Pennsylvania State University, University Park, Pennsylvania. <ftp://ftp.bio.indiana.edu/molbio/ibmpc/dispan.zip>
- Ovchinnikov IV, Götherström A, Romanova GP, Kharitonov VM, Lidén K, Goodwin W (2000) Molecular analysis of Neanderthal DNA from the northern Caucasus. *Nature* 404:490-493
- Pääbo S (1989) Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proc Natl Acad Sci USA* 86:6196-6200
- Pääbo S, Gifford JA, Wilson AC (1988) Mitochondrial DNA sequences from a 7000-year old brain. *Nucleic Acids Res* 16:9775-9787
- Pääbo S, Irwin DM, Wilson AC (1990) DNA Damage Promotes Jumping between Templates during Enzymatic Amplification. *J Biol Chem* 265:4718-4721
- Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, Rohland N, Kuch M, Krause J, Vigilant L, Hofreiter M (2004) Genetic Analyses from Ancient DNA. *Annu Rev Genet* 38:645-679
- Pagano S, Sans M, Pimenoff V, Cantera AM, Alvarez JC, Lorente JA, Peco JM, Mones P, Sajantila A (2005) Assessment of HV1 and HV2 mtDNA variation for forensic purposes in an Uruguayan population sample. *J Forensic Sci* 50:1239-1242
- Palermo MA (2000) A través de la frontera. Economía y sociedad indígenas desde el tiempo colonial hasta el siglo XIX. En: Tarragó MN (ed) *Nueva Historia Argentina*. Tomo I: Los pueblos originarios y la conquista. Editorial Sudamericana, Buenos Aires, pp 343-382
- Parr RL, Carlyle SW, O'Rourke DH (1996) Ancient DNA analysis of remnant Amerindians of the Great Salt Lake Wetlands. *Am J Phys Anthropol* 99:507-518
- Pena SDJ, Santos FR, Bianchi NO, Bravi CM, Carnese FR, Rothhammer F, Gerelsaikhan T, Munkhtuja B, Oyunsuren T (1995) A major founder Y-chromosome haplotype in Amerindians. *Nat Genet* 11:15-16
- Poinar GO (1993) Recovery of Antediluvian DNA: Reply to Lindahl. *Nature* 365:700
- Poinar GO, Poinar HN, Cano RJ (1994) DNA from Amber Inclusions. En: Herrmann B, Hummel S (eds) *Ancient DNA Recovery and Analysis of Genetic Material from Paleontological, Archaeological, Museum, Medical and Forensic Specimens*. Springer-Verlag, New York, pp 92-103

- Poinar HN, Hofreiter M, Spaulding WG, Martin PS, Stankiewicz BA, Bland H, Evershed RP, Possnert G, Pääbo S (1998) Molecular coproscopy: dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. *Science* 281:402-406
- Poinar HN, Stankiewicz BA (1999) Protein preservation and DNA retrieval from ancient tissues. *Proc Natl Acad Sci USA* 96:8426-8431
- Politis GG (1989) ¿Quién mató al Megaterio? *Ciencia Hoy* 1:26-35
- Politis GG (2000) Los cazadores de las llanuras. En: Tarragó MN (ed) *Nueva Historia Argentina. Tomo I: Los pueblos originarios y la conquista*. Editorial Sudamericana, Buenos Aires, pp 63-103
- Politis GG, Madrid PEE (2001) Arqueología Pampeana: Estado actual y perspectivas. En: Berberían EE, Nielsen AE (eds) *Historia Argentina Prehispánica, tomo II*. Editorial Brujas, Córdoba, pp 737-814
- Politis GG, Messineo PG, Kaufmann CA (2004) El poblamiento temprano de las llanuras pampeanas de Argentina y Uruguay. *Complutum* 15:207-224
- Powell J, Neves WA, Ozolins E, Pucciarelli HM (1999) Afinidades biológicas extra-continetales de los esqueletos más antiguos de América: implicaciones para el poblamiento del Nuevo Mundo. *Antropología física latinoamericana* 2:7-22
- Powell JF, Neves WA (1999) Craniofacial morphology of the first Americans: Pattern and process in the peopling of the New World. *Am J Phys Anthropol Suppl* 29:153-188
- R Development Core Team (2006) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <http://www.R-project.org>
- Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, Cavalli-Sforza LL (2005) Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proc Natl Acad Sci USA* 102:15942-15947
- Ribeiro-Dos-Santos AKC, Santos SEB, Machado AL, Guapindaia V, Zago MA (1996) Heterogeneity of Mitochondrial DNA Haplotypes in Pre-Columbian Natives of the Amazon Region. *Am J Phys Anthropol* 101:29-37
- Rivet P (1943) *Les origines de l' homme américain*. Gallimard, Paris
- Rogan PK, Salvo JJ (1990) Study of Nucleic Acids Isolated From Ancient Remains. *Yrbk Phys Anthropol* 33:195-214
- Rogers JS (1972) Measures of genetic similarity and genetic distance. In: *Studies in Genetics VII*. University of Texas Publication 7213, Austin, pp 145-153
- Rothhammer F, Silva C (1989) Peopling Andean South America. *Am J Phys Anthropol* 78:403-410
- Rothhammer F, Silva C (1992) Gene geography of South America: testing models of populations displacement based on archaeological evidence. *Am. J. Phys. Anthropol.* 89:441-446
- Ruiz-Linares A, Ortiz-Barrientos D, Figuerola M, Mesa N, Munera JG, Bedoya G, Velez ID, Garcia LF, Perez-Lezaun A, Bertranpetit J, Feldman MW, Goldstein DB (1999) Microsatellites provide evidence for Y chromosome diversity among the founders of the New World. *Proc Natl Acad Sci USA* 96:6312-6317

- Saitou N, Nei M (1987) The *neighbor-joining* method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Sans M (2004) Estudio de los restos del cacique Vaimaca Perú. Facultad de Humanidades y Ciencias de la Educación, Sección Antropología Biológica.
<http://www.fhuce.edu.uy/antrop/cursos/abiol/ifvaimaca.pdf>
- Sans M, Figueiro G (2005) Continuidad indígena en la población uruguaya actual: ¿Guaraníes o Charrúas? XI Congreso Nacional de Arqueología Uruguay, Salto
- Sans M, Merriwether DA, Hidalgo PC, Bentancor N, Weimer TA, Franco MH, Alvarez I, Kemp BM, Salzano FM (2006) Population structure and admixture in Cerro Largo, Uruguay, based on blood markers and mitochondrial DNA polymorphisms. *Am J Hum Biol* 18:513-522
- Schmitz RW, Serre D, Bonani G, Feine S, Hillgruber F, Krainitzki H, Pääbo S, Smith FH (2002) The Neandertal type-site revisited: interdisciplinary investigations of skeletal remains from the Neander Valley, Germany. *Proc Natl Acad Sci USA* 99:13342-13347
- Schneider S, Roessli D, Excoffier L (2000) *Arlequin ver 2.000: a software package for population genetics data analysis*. University of Geneva, Genetics and Biometry Laboratory, Geneva
- Schurr TG, Ballinger SW, Gan Y-Y, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, Weiss KM, Wallace DC (1990) Amerindian Mitochondrial DNAs have Rare Asian Mutations at High Frequencies, Suggesting They Derived from Four Primary Maternal Lineages. *Am J Hum Genet* 46:613-623
- Schurr TG, Sukernik RI, Starikovskaya YB, Wallace DC (1999) Mitochondrial DNA Variation in Koryaks and Itel'men: Population Replacement in the Okhotsk Sea-Bering Sea Region During the Neolithic. *Am J Phys Anthropol* 108:1-39
- Schurr TG, Sherry ST (2004) Mitochondrial DNA and Y Chromosome Diversity and the Peopling of the Americas: Evolutionary and Demographic Evidence. *Am J Hum Biol* 16:420-439
- Silva WA, Bonatto SL, Holanda AJ, Ribeiro-dos-Santos AK, Paixão BM, Goldman GH, Abe-Sandes K, Rodriguez-Delfin L, Barbosa M, Paço-Larson ML, Petzl-Erler ML, Valente V, Santos SEB, Zago MA (2002) Mitochondrial Genome Diversity of Native Americans Supports a Single Early Entry of Founder Populations into America. *Am J Hum Genet* 71:187-192
- Smith DG, Malhi RS, Eshleman J, Lorenz JG, Kaestle FA (1999) Distribution of mtDNA Haplogroup X Among Native North Americans. *Am J Phys Anthropol* 110:271-284
- Sparks CS, Jantz RL (2002) A reassessment of human cranial plasticity: Boas revisited. *Proc Natl Acad Sci USA* 99:14636-14639
- Spencer M, Howe CJ (2004) Authenticity of Ancient-DNA Results: A Statistical Approach. *Am J Hum Genet* 75:240-250
- Stone AC, Stoneking M (1993) Ancient DNA From a Pre-Columbian Amerindian Population. *Am J Phys Anthropol* 92:463-471

- Stone AC, Stoneking M (1998) mtDNA Analysis of a Prehistoric Oneota Population: Implications for the Peopling of the New World. *Am J Hum Genet* 62:1153-1170
- Stone AC, Stoneking M (1999) Analysis of Ancient DNA from a prehistoric Amerindian Cemetery. *Phil Trans R Soc Lond B* 354:153-159
- Stoneking M (2000) Hypervariable Sites in the mtDNA Control Region Are Mutational Hotspots. *Am J Hum Genet* 67:1029-1032
- Suárez R, López JM (2003) Archaeology of the Pleistocene-Holocene transition in Uruguay: an overview. *Quaternary International* 109-110:65-76
- Surovell TA (2003) Simulating Coastal Migration in New World Colonization. *Curr Anthropol* 44:580-591
- Tarazona-Santos E, Carvalho-Silva DR, Pettener D, Luiselli D, De Stefano GF, Martínez Labarga C, Rickards O, Tyler-Smith C, Pena SDJ, Santos FR (2001) Genetic Differentiation in South Amerindians Is Related to Environmental and Cultural Diversity: Evidence from the Y Chromosome. *Am J Hum Genet* 68:1485-1496
- Tarazona-Santos E, Santos FR (2002) The peopling of the Americas: a second major migration? *Am J Hum Genet* 70:1377-1381
- Torres MM, Bravi CM, Bortolini MC, Duque C, Callegari-Jacques S, Ortiz D, Bedoya G, Groot de Restrepo H, Ruiz-Linares A (2006) A revertant of the major founder Native American haplogroup C common in populations from northern South America. *Am J Hum Biol* 18:59-65
- Torrioni A, Neel JV, Barrantes R, Schurr TG, Wallace DC (1994) A mitochondrial DNA "clock" for the Amerinds and its implications for timing their entry into North America. *Proc Natl Acad Sci USA* 91:1158-1162
- Torrioni A, Schurr TG, Yang C-C, Szathmary EJE, Williams RC, Schanfield MS, Troup GA, Knowler WC, Lawrence DN, Weiss KM, Wallace DC (1992) Native American Mitochondrial DNA Analysis Indicates That the Amerind and the Nadene Populations Were Founded by Two Independent Migrations. *Genetics* 130:153-162
- Underhill PA, Jin L, Zeman R (1996) A pre-Columbian Y-chromosome-specific transition and its implications for human evolutionary history. *Proc Natl Acad Sci USA* 93:196-200
- Walker PL, Johnson JR, Lambert PM (1988) Age and sex biases in the preservation of human skeletal remains. *Am J Phys Anthropol* 76:183-188
- Wallace DC, Garrison K, Knowler WC (1985) Dramatic Founder Effects in Amerindian Mitochondrial DNAs. *Am J Phys Anthropol* 68:149-155
- Wallace DC, Lott MT (2003) MITOMAP: A Human Mitochondrial Genome Database. <http://www.mitomap.org>
- Ward RH, Frazier BL, Dew-Jager K, Pääbo S (1991) Extensive mitochondrial diversity within a single Amerindian Tribe. *Proc Natl Acad Sci USA* 88:8720-8724

Willerslev E, Cooper A (2005) Ancient DNA. *Proc R Soc Lond B* 272:3-16

Willerslev E, Hansen AJ, Poinar HN (2004) Isolation of nucleic acids and cultures from fossil ice and permafrost. *Trends Ecol Evol* 19:141-147

Wiuf C (2001) Recombination in Human Mitochondrial DNA? *Genetics* 159:749-756

Yang DY, Watt K (2005) Contamination controls when preparing archaeological remains for ancient DNA analysis. *J Archaeol Sci* 32:331-336

Zegura SL, Karafet TM, Zhivotovsky LA, Hammer MF (2004) High-Resolution SNPs and Microsatellite Haplotypes Point to a Single, Recent Entry of Native American Y Chromosomes into the Americas. *Mol Biol Evol* 21:164-175

Zink A, Nerlich AG (2003) Molecular analyses of the "Pharaos:" Feasibility of molecular studies in ancient Egyptian material. *Am J Phys Anthropol* 121:109-111

ANEXOS

ANEXO I

Protocolo de extracción de ADN utilizado en el presente trabajo (basado en Baron et al. 1996)

DESCALCIFICACIÓN

Echar hasta 0,3 g de polvo en tubos etiquetados de 15 ml.

Agregar 2 ml 0.5M EDTA (pH 8.3).

Sellar con *parafilm*. Dejar agitando a temperatura ambiente por 72-96 horas.

Centrifugar 1 minuto a 700 rpm.

DIGESTIÓN CON PROTEINASA K*

Agregar 1.925 ml de agua.

Agregar 75 µl de Proteinasa K 20 mg/ml.

Agitar en *vortex* e incubar rotando 1.5 h a 60°C.

EXTRACCIÓN DEL ADN

Agregar un volumen de fenol/cloroformo/isoamil alcohol (~4 ml) a las muestras.

Mezclar lentamente por 10 minutos.

Centrifugar 3 minutos a 1500 rpm.

Transferir la fase acuosa de cada muestra a una alícuota de p/c/i (Normalmente la fase superior; agregar fenol o quedarse con ambas fases en caso de duda).**

Mezclar lentamente por 10 minutos.

Centrifugar 3 minutos a 1500 rpm.

Transferir la fase acuosa de cada muestra a una alícuota de cloroformo.

Mezclar lentamente por 10 minutos.

Centrifugar 3 minutos a 1500 rpm.

PRECIPITACIÓN DEL ADN

Transferir la fase acuosa de cada muestra a una alícuota de 80 µl de acetato de sodio y 3,3 ml de isopropanol (el pH del acetato de sodio ha de ser reducido a 4-4,5 con ácido clorhídrico; pH de la mezcla: 6,5-7,0)***

Mezclar por inversión durante 2 minutos.

Agregar 10 µl de suspensión de sílice (*glassmilk*).

Rotar lentamente durante 20 minutos a 55 °C.

Centrifugar 3 minutos a 700 rpm. Desechar líquido con cuidado, y volver a centrifugar si el *pellet* de *glassmilk* se desprende.

1) Agregar 4 ml de etanol 80% helado. Agitar en *vortex*.

2) Centrifugar 3 minutos a 700 rpm.

* A diferencia de otras proteasas, la proteinasa K no es inhibida por el EDTA utilizado en la descalcificación.

** Siendo un ácido, el ADN permanece disuelto en la fase acuosa, en tanto que otros compuestos generalmente hallados en el tejido y el sedimento permanecen en la fase orgánica.

*** Hänni et al. (1995) sugieren el uso de isopropanol en la primera precipitación dado que tiene una supuesta mayor selectividad por el ADN. Asimismo, Poinar et al. (1998) indican que el isopropanol es más eficiente a la hora de quitar el soluto de color amarronado atribuido a productos de Maillard, y que inhiben fuertemente la PCR.

3) Volcar etanol con cuidado de que el *glassmilk* permanezca.
Repetir 1-3 una vez.

Colocar los tubos de lado y dejar secar en incubadora a 60 °C por 2 horas, o hasta que se haya evaporado el etanol.

ELUCIÓN

Agregar 125 µl TE. Agitar bien en *vortex*.

Centrifugar a 750 rpm. Transferir a un tubo etiquetado sin el *glassmilk*.****

**** La suspensión de sílice es un paso adicional en la eliminación de coextractos potencialmente inhibidores de la reacción de PCR, pero es a su vez un inhibidor.

ANEXO II

Detalle de la determinación de haplogrupo de los individuos de la muestra de Arroyo Seco 2

EXTRACCIÓN 1: 9 de marzo de 2004

ARROYO SECO 01

Índice promedio de éxito: 1,75

Haplogrupo A: Negativo.

Haplogrupo B: Negativo.

Haplogrupo C: Negativo.

Haplogrupo D: Resultados positivos (1) y negativos (1).

En virtud del fracaso repetido en el intento de amplificar la sección M4 de la región hipervariable, el haplogrupo de este individuo permanece incierto.

ARROYO SECO 02

Índice promedio de éxito: 2,75

Haplogrupo A: Negativo.

Haplogrupo B: Positivo.

Haplogrupo C: Negativo.

Haplogrupo D: Resultados positivos (1) y negativos (3).

Los intentos repetidos de amplificar la sección M4 de la región hipervariable fueron infructuosos; se le asigna el haplogrupo B.

EXTRACCIÓN 2: 27 de abril de 2004

ARROYO SECO 04

Índice promedio de éxito: 3,33

Haplogrupo A: No amplificó.

Haplogrupo B: Positivo.

Haplogrupo C: Negativo.

Haplogrupo D: Negativo.

Se resuelve clasificarlo como haplogrupo B.

ARROYO SECO 05

Índice promedio de éxito: 3,25

Haplogrupo A: Negativo.

Haplogrupo B: Negativo.

Haplogrupo C: Negativo.

Haplogrupo D: Resultados positivos (1) y negativos (2).

La secuenciación del tramo M4 (16288–16355) de la región hipervariable I reveló la presencia de una transición T → C en la posición 16325 y una transición C → T en la posición 16327 característicos del haplogrupo C. Se resuelve asignar a AS05 el haplogrupo C.

EXTRACCIÓN 3 (14 de mayo de 2004; AS07 y AS08) y **EXTRACCIÓN 4** (28 de mayo de 2004; AS14 y AS19): No rindieron ADN. Se reanudó la actividad tras un recambio general de reactivos.

EXTRACCIÓN 5: 25 de junio de 2004

ARROYO SECO 20

Índice promedio de éxito: 2,5

Haplogrupo A: Negativo.

Haplogrupo B: Negativo.

Haplogrupo C: Positivo.

Haplogrupo D: Negativo.

Se resuelve clasificarlo como haplogrupo C. Por añadidura, la secuenciación del tramo M4 (16288–16355) de la región hipervariable I reveló la presencia de una transición T → C en la posición 16325 y una transición C → T en la posición 16327 característicos del haplogrupo C.

EXTRACCIÓN 6: 13 de julio de 2004

ARROYO SECO 22

Índice promedio de éxito: 1,67

Haplogrupo A: Negativo.

Haplogrupo B: Negativo.

Haplogrupo C: No amplificó.

Haplogrupo D: Negativo.

La secuenciación del tramo M4 (16288–16355) de la región hipervariable I reveló la presencia de una transición T → C en la posición 16311, presente en el investigador y característica de algunos haplogrupos europeos y africanos. Por lo tanto, se considera que la extracción de AS22 fue contaminada por ADN moderno.

ARROYO SECO 23

No rindió ADN.

EXTRACCIÓN 7: 23 de julio de 2004

ARROYO SECO 24

Índice promedio de éxito: 1,75

Haplogrupo A: Negativo.

Haplogrupo B: Negativo.

Haplogrupo C: Positivo.

Haplogrupo D: Negativo.

Se resuelve clasificarlo como haplogrupo C. Por añadidura, la secuenciación del tramo M4 (16288–16355) de la región hipervariable I reveló la presencia de una transición T → C en la posición 16325 y una transición C → T en la posición 16327 característicos del haplogrupo C.

ARROYO SECO 25

Haplogrupo C: Negativo.

Ulteriores intentos de amplificar el marcador del haplogrupo C, así como todos los intentos de amplificar los demás marcadores fracasaron. Se considera que se trató de una amplificación producto de contaminación esporádica, y se concluye que la muestra no rindió ADN.

EXTRACCIÓN 8: 13 de agosto de 2004

ARROYO SECO 27

Índice promedio de éxito: 3,67

Haplogrupo A: Negativo.

Haplogrupo B: Negativo.

Haplogrupo C: Positivo.

Haplogrupo D: No amplificó.

Se resuelve clasificarlo como haplogrupo C.

ARROYO SECO 30

No rindió ADN.

EXTRACCIÓN 9: 21 de setiembre de 2004

ARROYO SECO 31

Índice promedio de éxito: 2,75

Haplogrupo A: Negativo.

Haplogrupo B: Negativo.

Haplogrupo C: Negativo.

Haplogrupo D: Negativo.

La secuenciación del tramo M4 (16288–16355) de la región hipervariable I reveló la presencia de una transición T → C en la posición 16311, presente en el investigador y característica de algunos haplogrupos europeos y africanos. Por lo tanto, se considera que la extracción de AS31 fue contaminada por ADN moderno.

ARROYO SECO 32

Índice promedio de éxito: 1

Haplogrupo A: No amplificó.

Haplogrupo B: Negativo.

Haplogrupo C: No amplificó.

Haplogrupo D: Negativo.

En virtud del fracaso repetido en el intento de amplificar la sección M4 de la región hipervariable, el haplogrupo de este individuo permanece incierto. Sin embargo, en vista de la contaminación observada en AS31, se asume que AS32 puede haber sido contaminada por ADN moderno.

ARROYO SECO 36

No rindió ADN.

EXTRACCIÓN 10: 5 de octubre de 2004

ARROYO SECO 39

Índice promedio de éxito: 2,67

Haplogrupo A: No amplificó.

Haplogrupo B: Negativo.

Haplogrupo C: Negativo.

Haplogrupo D: Negativo.

La secuenciación del tramo M4 (16288–16355) de la región hipervariable I reveló la presencia de una transición T → C en la posición 16311, presente en el investigador y característica de algunos haplogrupos europeos y africanos. Por lo tanto, se considera que la extracción de AS39 fue contaminada por ADN moderno.

ARROYO SECO 40

Índice promedio de éxito: 1

Haplogrupo A: No amplificó.

Haplogrupo B: Negativo.

Haplogrupo C: Negativo.

Haplogrupo D: Negativo.

En virtud del fracaso repetido en el intento de amplificar la sección M4 de la región hipervariable, el haplogrupo de este individuo permanece incierto. Sin embargo, en vista de la contaminación observada en AS39 y AS41, se asume que AS40 puede haber sido contaminada por ADN moderno.

ARROYO SECO 41

Índice promedio de éxito: 1,33

Haplogrupo A: No amplificó.

Haplogrupo B: Negativo.

Haplogrupo C: Negativo.

Haplogrupo D: Negativo.

La secuenciación del tramo M4 (16288–16355) de la región hipervariable I reveló la presencia de una transición T → C en la posición 16311, presente en el investigador y característica de algunos haplogrupos europeos y africanos. Por lo tanto, se considera que la extracción de AS41 fue contaminada por ADN moderno.

EXTRACCIÓN 11: 19 de octubre de 2004

ARROYO SECO 43

Índice promedio de éxito: 3

Haplogrupo A: Negativo.

Haplogrupo B: Positivo.

Haplogrupo C: Negativo.

Haplogrupo D: Resultados positivos (2) y negativos (3).

Se resuelve asignar a AS43 el haplogrupo B.

ARROYO SECO 44:

Índice promedio de éxito: 3,67

Haplogrupo A: No amplificó.

Haplogrupo B: Negativo.

Haplogrupo C: Negativo.

Haplogrupo D: Positivo en tres ocasiones y negativo en otras tres. Se resuelve asignar a AS44 el haplogrupo D.

ANEXO III

Detalle de los procedimientos de remuestreo implementados con el paquete R

2.3.1

En todos los casos, se describe la rutina implementada con poblaciones antiguas; la rutina implementada con poblaciones modernas difiere únicamente en la definición y número de las muestras.

- a. *Bootstrap* de frecuencias de los marcadores A, B, C y D para testar la consistencia de los dendrogramas elaborados a partir de las distancias D_R de Rogers

```
# Bootstrap de "loci" y dendrogramas por NJ de distancias de Rogers de
AS2 y Antiguos

#"Dataset" es un vector ficticio a partir del cual se va a hacer el
remuestreo de haplos cual loci
Dataset<-c(1,2,3,4)

#Definición de "Haplos", una lista de frecuencias de A, B, C, D en las
poblaciones Nortechile, Amazonia,
#Tierradelfuego, Patagonia, NOA, Esteuruguayo, Colombia, Dominicana,
Cuba y Pampa, respectivamente

Haplos<-NULL
Haplos<-as.list(Haplos)

Haplos[[1]]<-
c(0.342,0.455,0.000,0.000,0.060,0.143,0.500,0.000,0.070,0.000)
Haplos[[2]]<-
c(0.368,0.091,0.000,0.000,0.500,0.429,0.167,0.000,0.000,0.375)
Haplos[[3]]<-
c(0.250,0.364,0.485,0.267,0.000,0.429,0.333,0.750,0.600,0.500)
Haplos[[4]]<-
c(0.039,0.091,0.515,0.733,0.440,0.000,0.000,0.250,0.330,0.125)

library(ade4)
library(ape)
library(cluster)

#Definición de objetos
distr Rogers<-NULL
treerogers<-NULL

#Conversión de los objetos en listas
distr Rogers<-list(distr Rogers)
treerogers<-list(treerogers)

#Bucle: Bootstrap, cálculo de distancias y elaboración de árboles
for(i in 1:1000)
{
#Bootstrap: Definición de cada "remuestra" de loci
muestra<- sample(Dataset, length(Dataset), replace = TRUE)
```

```

#Armado de los vectores de "loci" remuestreados y sus frecuencias
complementarias
j<-muestra[1]
k<-muestra[2]
l<-muestra[3]
m<-muestra[4]

a1<-Haplos[[j]]
a2<-1-a1
b1<-Haplos[[k]]
b2<-1-b1
c1<-Haplos[[l]]
c2<-1-c1
d1<-Haplos[[m]]
d2<-1-d1

#Elaboración del marco de datos a partir de los nuevos vectores
frecs <- data.frame(L1.1=a1, L1.2=a2, L2.1=b1, L2.2=b2, L3.1=c1,
L3.2=c2, L4.1=d1, L4.2=d2)

#Conversión de tabla de frecuencias a objeto tipo genet
frecs<-freq2genet(frecs)

#Cálculo de distancias
distograms[[i]]<-dist.genet(frecs, method = 4, diag = TRUE, upper =
TRUE)

#Conversión de las matrices de distancias i en objetos del tipo dist
distograms[[i]]<-as.dist(distograms[[i]])

#Chequeo de distancias infinitas en cada matriz (i) y, si no las
hubiere, elaboración de NJ
infinitosrogers<-is.infinite(distograms[[i]])
ifelse(any(infinitosrogers=='TRUE'), treerogers[[i]]<-treerogers[[i-
1]], treerogers[[i]]<-nj(distograms[[i]])

#Guardado
write.tree(treerogers[[i]], file="rogersantloci.txt", append = TRUE)

#Cierre bucle
}

```

- b. *Bootstrap* de individuos de las muestras comparativas para testar la consistencia estadística de los dendrogramas elaborados a partir de las distancias D_A y D_R

1. D_A

```
# Bootstrap de dendrogramas por NJ de DA de Nei de AS2 y Antiguos
#Definición de las muestras
Nortechile<-
c("A","A","A","A","A","A","A","A","A","A","A","A","A","A","A","A","A",
"A","A","A","A","A","A","A","A","A","A","B","B","B","B","B","B","B","B",
"B","B","B","B","B","B","B","B","B","B","B","B","B","B","B","B","B",
"B","B","C","C","C","C","C","C","C","C","C","C","C","C","C","C","C",
"C","C","C","D","D","D")
Amazonia<-c("A","A","A","A","A","B","C","C","C","C")
Tierradelfuego<-
c("C","C","C","C","C","C","C","C","C","C","C","C","C","C","C","C","C",
"C","C","C","C","C","C","C","C","C","C","C","C","C","C","C","D","D",
"D","D","D","D","D","D","D","D","D","D","D","D","D","D","D","D",
"D","D","D","D","D","D","D","D","D","D","D","D","D","D","D","D")
Patagonia<-
c("C","C","C","C","D","D","D","D","D","D","D","D","D","D","D")
NOA<-
c("A","B","B","B","B","B","B","B","B","B","D","D","D","D","D","D")
Esteuruguayo<-c("A","B","B","B","C","C","C")
Colombiana<-c("A","A","A","B","C","C")
Dominicana<-
c("C","C","C","C","C","C","C","C","C","C","C","C","C","C","C","C",
"C","D","D","D","D","D","D")
Cuba<-c("A","C","C","C","C","C","C","C","C","C","D","D","D","D","D")
Pampa<-c("B","B","B","C","C","C","C","D")

library(ade4)
library(ape)
library(cluster)

#Definición de objetos
distanei<-NULL
treeanei<-NULL

#Conversión de los objetos en listas
distanei<-list(distanei)
treeanei<-list(treeanei)

#Bucle: Bootstrap, cálculo de distancias y elaboración de árboles
for(i in 1:1000)
{
#Bootstrap
muestral<- sample(Nortechile, length(Nortechile), replace = TRUE)
muestrala<-subset(muestral, muestral=='A')
muestralb<-subset(muestral, muestral=='B')
muestralc<-subset(muestral, muestral=='C')
muestrald<-subset(muestral, muestral=='D')
muestra2<- sample(Amazonia, length(Amazonia), replace = TRUE)
muestra2a<-subset(muestra2, muestra2=='A')
muestra2b<-subset(muestra2, muestra2=='B')
muestra2c<-subset(muestra2, muestra2=='C')
muestra2d<-subset(muestra2, muestra2=='D')
```

```
muestra3<- sample(Tierradelfuego, length(Tierradelfuego), replace =
TRUE)
muestra3a<-subset(muestra3, muestra3=='A')
muestra3b<-subset(muestra3, muestra3=='B')
muestra3c<-subset(muestra3, muestra3=='C')
muestra3d<-subset(muestra3, muestra3=='D')
muestra4<- sample(Patagonia, length(Patagonia), replace = TRUE)
muestra4a<-subset(muestra4, muestra4=='A')
muestra4b<-subset(muestra4, muestra4=='B')
muestra4c<-subset(muestra4, muestra4=='C')
muestra4d<-subset(muestra4, muestra4=='D')
muestra5<- sample(NOA, length(NOA), replace = TRUE)
muestra5a<-subset(muestra5, muestra5=='A')
muestra5b<-subset(muestra5, muestra5=='B')
muestra5c<-subset(muestra5, muestra5=='C')
muestra5d<-subset(muestra5, muestra5=='D')
muestra6<- sample(Esteuruguayo, length(Esteuruguayo), replace = TRUE)
muestra6a<-subset(muestra6, muestra6=='A')
muestra6b<-subset(muestra6, muestra6=='B')
muestra6c<-subset(muestra6, muestra6=='C')
muestra6d<-subset(muestra6, muestra6=='D')
muestra7<- sample(Colombia, length(Colombia), replace = TRUE)
muestra7a<-subset(muestra7, muestra7=='A')
muestra7b<-subset(muestra7, muestra7=='B')
muestra7c<-subset(muestra7, muestra7=='C')
muestra7d<-subset(muestra7, muestra7=='D')
muestra8<- sample(Dominicana, length(Dominicana), replace = TRUE)
muestra8a<-subset(muestra8, muestra8=='A')
muestra8b<-subset(muestra8, muestra8=='B')
muestra8c<-subset(muestra8, muestra8=='C')
muestra8d<-subset(muestra8, muestra8=='D')
muestra9<- sample(Cuba, length(Cuba), replace = TRUE)
muestra9a<-subset(muestra9, muestra9=='A')
muestra9b<-subset(muestra9, muestra9=='B')
muestra9c<-subset(muestra9, muestra9=='C')
muestra9d<-subset(muestra9, muestra9=='D')
muestra10<- sample(Pampa, length(Pampa), replace = TRUE)
muestra10a<-subset(muestra10, muestra10=='A')
muestra10b<-subset(muestra10, muestra10=='B')
muestra10c<-subset(muestra10, muestra10=='C')
muestra10d<-subset(muestra10, muestra10=='D')
```

```

#Armado de los vectores de frecuencias absolutas remuestreadas
a<-
c(length(muestrala),length(muestra2a),length(muestra3a),length(muestra
4a),length(muestra5a),length(muestra6a),length(muestra7a),length(muestr
a8a),length(muestra9a),length(muestra10a))
b<-
c(length(muestralb),length(muestra2b),length(muestra3b),length(muestra
4b),length(muestra5b),length(muestra6b),length(muestra7b),length(muestr
a8b),length(muestra9b),length(muestra10b))
c<-
c(length(muestralc),length(muestra2c),length(muestra3c),length(muestra
4c),length(muestra5c),length(muestra6c),length(muestra7c),length(muestr
a8c),length(muestra9c),length(muestra10c))
d<-
c(length(muestrald),length(muestra2d),length(muestra3d),length(muestra
4d),length(muestra5d),length(muestra6d),length(muestra7d),length(muestr
a8d),length(muestra9d),length(muestra10d))

#Armado de los vectores de frecuencias RELATIVAS remuestreadas
sizes<-
c(length(muestra1),length(muestra2),length(muestra3),length(muestra4),
length(muestra5),length(muestra6),length(muestra7),length(muestra8),le
ngth(muestra9),length(muestra10))
fra<-a/sizes
frb<-b/sizes
frc<-c/sizes
frd<-d/sizes

#Creación de la matriz para albergar las distancias
distanei[[i]]<-matrix(data = NA, nrow = 10, ncol = 10, byrow = FALSE)

#Cálculo de distancia A de Nei (un poco complicado, pero sirve para
salir del paso)
for(j in 1:10)
{
pob1<-c(fra[j],frb[j],frc[j],frd[j])
  for(k in 1:10)
  {
pob2<-c(fra[k],frb[k],frc[k],frd[k])
danei<-1-sum(sqrt(pob1*pob2))
distanei[[i]][j,k]<-danei
distanei[[i]][k,j]<-danei
}
}

#Conversión de la matriz de distancias en objeto del tipo dist
distanei[[i]]<-as.dist(distanei[[i]])

#Elaboración de árboles NJ
treeanei[[i]]<-nj(distanei[[i]])

#Guardado
write.tree(treeanei[[i]], file="antnjnea2.txt", append = TRUE)

#Cierre bucle
}

```

2. D_R

```
# Bootstrap de dendrogramas por NJ de distancias de Rogers entre AS2 y
Antiguos

#Definición de las muestras
Nortechile<-
c("A","A","A","A","A","A","A","A","A","A","A","A","A","A","A","A","A",
"A","A","A","A","A","A","A","A","A","A","B","B","B","B","B","B","B","B","B",
"B","B","B","B","B","B","B","B","B","B","B","B","B","B","B","B","B","B",
"B","B","C","C","C","C","C","C","C","C","C","C","C","C","C","C","C","C",
"C","C","C","C","D","D","D")
Amazonia<-c("A","A","A","A","A","B","C","C","C","C")
Tierradelfuego<-
c("C","C","C","C","C","C","C","C","C","C","C","C","C","C","C","C","C",
"C","C","C","C","C","C","C","C","C","C","C","C","C","C","C","D","D",
"D","D","D","D","D","D","D","D","D","D","D","D","D","D","D","D","D",
"D","D","D","D","D","D","D","D","D","D","D","D","D","D","D","D","D")
Patagonia<-
c("C","C","C","C","D","D","D","D","D","D","D","D","D","D","D","D")
NOA<-
c("A","B","B","B","B","B","B","B","B","B","D","D","D","D","D","D","D")
Esteuruguayo<-c("A","B","B","B","C","C","C")
Colombia<-c("A","A","A","B","C","C")
Dominicana<-
c("C","C","C","C","C","C","C","C","C","C","C","C","C","C","C","C","C",
"C","D","D","D","D","D","D","D")
Cuba<-c("A","C","C","C","C","C","C","C","C","C","C","C","D","D","D","D","D")
Pampa<-c("B","B","B","C","C","C","C","C","D")

library(ade4)
library(ape)
library(cluster)

#Definición de objetos
distr Rogers<-NULL
treerogers<-NULL

#Conversión de los objetos en listas
distr Rogers<-list(distr Rogers)
treerogers<-list(treerogers)

#Bucle: Bootstrap, cálculo de distancias y elaboración de árboles
for(i in 1:1000)
{
#Bootstrap
muestra1<- sample(Nortechile, length(Nortechile), replace = TRUE)
muestra1a<-subset(muestra1, muestra1=='A')
muestra1b<-subset(muestra1, muestra1=='B')
muestra1c<-subset(muestra1, muestra1=='C')
muestra1d<-subset(muestra1, muestra1=='D')
muestra2<- sample(Amazonia, length(Amazonia), replace = TRUE)
muestra2a<-subset(muestra2, muestra2=='A')
muestra2b<-subset(muestra2, muestra2=='B')
muestra2c<-subset(muestra2, muestra2=='C')
muestra2d<-subset(muestra2, muestra2=='D')
```

```
muestra3<- sample(Tierradelfuego, length(Tierradelfuego), replace =
TRUE)
muestra3a<-subset(muestra3, muestra3=='A')
muestra3b<-subset(muestra3, muestra3=='B')
muestra3c<-subset(muestra3, muestra3=='C')
muestra3d<-subset(muestra3, muestra3=='D')
muestra4<- sample(Patagonia, length(Patagonia), replace = TRUE)
muestra4a<-subset(muestra4, muestra4=='A')
muestra4b<-subset(muestra4, muestra4=='B')
muestra4c<-subset(muestra4, muestra4=='C')
muestra4d<-subset(muestra4, muestra4=='D')
muestra5<- sample(NOA, length(NOA), replace = TRUE)
muestra5a<-subset(muestra5, muestra5=='A')
muestra5b<-subset(muestra5, muestra5=='B')
muestra5c<-subset(muestra5, muestra5=='C')
muestra5d<-subset(muestra5, muestra5=='D')
muestra6<- sample(Esteuruguayo, length(Esteuruguayo), replace = TRUE)
muestra6a<-subset(muestra6, muestra6=='A')
muestra6b<-subset(muestra6, muestra6=='B')
muestra6c<-subset(muestra6, muestra6=='C')
muestra6d<-subset(muestra6, muestra6=='D')
muestra7<- sample(Colombia, length(Colombia), replace = TRUE)
muestra7a<-subset(muestra7, muestra7=='A')
muestra7b<-subset(muestra7, muestra7=='B')
muestra7c<-subset(muestra7, muestra7=='C')
muestra7d<-subset(muestra7, muestra7=='D')
muestra8<- sample(Dominicana, length(Dominicana), replace = TRUE)
muestra8a<-subset(muestra8, muestra8=='A')
muestra8b<-subset(muestra8, muestra8=='B')
muestra8c<-subset(muestra8, muestra8=='C')
muestra8d<-subset(muestra8, muestra8=='D')
muestra9<- sample(Cuba, length(Cuba), replace = TRUE)
muestra9a<-subset(muestra9, muestra9=='A')
muestra9b<-subset(muestra9, muestra9=='B')
muestra9c<-subset(muestra9, muestra9=='C')
muestra9d<-subset(muestra9, muestra9=='D')
muestra10<- sample(Pampa, length(Pampa), replace = TRUE)
muestra10a<-subset(muestra10, muestra10=='A')
muestra10b<-subset(muestra10, muestra10=='B')
muestra10c<-subset(muestra10, muestra10=='C')
muestra10d<-subset(muestra10, muestra10=='D')
```

```

#Armado de los vectores de frecuencias absolutas remuestreadas
a<-
c(length(muestra1a),length(muestra2a),length(muestra3a),length(muestra4a
),length(muestra5a),length(muestra6a),length(muestra7a),length(muestra8a
),length(muestra9a),length(muestra10a))
b<-
c(length(muestra1b),length(muestra2b),length(muestra3b),length(muestra4b
),length(muestra5b),length(muestra6b),length(muestra7b),length(muestra8b
),length(muestra9b),length(muestra10b))
c<-
c(length(muestra1c),length(muestra2c),length(muestra3c),length(muestra4c
),length(muestra5c),length(muestra6c),length(muestra7c),length(muestra8c
),length(muestra9c),length(muestra10c))
d<-
c(length(muestra1d),length(muestra2d),length(muestra3d),length(muestra4d
),length(muestra5d),length(muestra6d),length(muestra7d),length(muestra8d
),length(muestra9d),length(muestra10d))

#Elaboración del marco de datos a partir de los vectores
freqs <- data.frame(L1.A=a, L1.B=b, L1.C=c, L1.D=d)

#Conversión de tabla de frecuencias a objeto tipo genet
freqs<-count2genet(freqs)

#Cálculo de distancias
distograms[[i]]<-dist.genet(freqs, method = 4, diag = TRUE, upper =
TRUE)

#Conversión de las matrices de distancias i en objetos del tipo dist
distograms[[i]]<-as.dist(distograms[[i]])

#Chequeo de distancias infinitas en cada matriz (i) y, si no las
hubiere, elaboración de NJ
infinitogramas<-is.infinite(distograms[[i]])
ifelse(any(infinitogramas=='TRUE'), treeogramas[[i]]<-treeogramas[[i-1]],
treeogramas[[i]]<-nj(distograms[[i]]))

#Guardado
write.tree(treeogramas[[i]], file="antnjogramas.txt", append = TRUE)

#Cierre bucle
}

```