



UNIVERSITAT DE BARCELONA



**Variació genètica i evolució d'elements *Alu* recents
en poblacions humanes.
Inferències biodemogràfiques i filogeogràfiques**

**Genetic variation and evolution of recent *Alu* elements in human populations.
Biodemographic and philogeographic inferences**

Memòria presentada per

Emili González Pérez

per optar al grau de
Doctor per la Universitat de Barcelona

Dirigida pel Dr. Pedro Moral Castrillo, Professor Titular d'Antropologia Física
de la Unitat d'Antropologia del Departament de Biologia Animal
de la Universitat de Barcelona.

Programa de Doctorat d'*Antropologia Biològica*, bienni 1999-2001.
Departament de Biologia Animal – Facultat de Biologia

Pedro Moral Castrillo
Director

Clara García Moro
Tutora

Emili González Pérez
Doctorand

Annexos

La ciencia no me interesa.
Ignora el sueño, el azar, la risa,
el sentimiento y la contradicción,
cosas que me son preciosas.

Luis Buñuel

ANNEX 1

Localització geogràfica i poblacions del Mediterrani

Les mostres utilitzades en aquest estudi corresponen a un total de 2205 individus provinents de 24 poblacions humanes significatives de la regió mediterrània, amb una especial rellevància de la part occidental. La selecció de mostres poblacionals s'ha portat a terme seguint els següents objectius principals: per una banda, la descripció i caracterització de la diversitat genètica per aquest tipus de marcadors a les poblacions mediterrànies; en segon lloc, la inferència de patrons poblacionals actuals i del passat d'aquestes poblacions tan importants per a la comprensió de l'evolució humana recent i de l'expansió de les civilitzacions modernes; i finalment, trobar noves evidències per desxifrar algunes de les incògnites encara existents sobre el paper de les poblacions berbers en l'evolució i diferenciació de les poblacions humanes autòctones del Nord d'Àfrica i de la resta de poblacions mediterrànies.

Les poblacions seleccionades per a l'anàlisi es poden classificar en 5 grups principals, en funció bàsicament del seu origen geogràfic:

1) Poblacions del Mediterrani Occidental Europeu:

Inclou 5 mostres poblacionals de la península Ibèrica, incloent-hi mostres del nord, centre i sud peninsular, així com representants particularment interessants des del punt de vista antropològic com ara els bascos o la població del Vall de Pas a Cantàbria. També inclou una població extrapeninsular del Sud de França.

2) Poblacions del Mediterrani Oriental:

Entre aquestes destaquen poblacions de l'Orient Mediterrani de la costa nord, com Grecs i Turcs, l'illa de Creta i alguna població particular de l'Orient Mediterrani del Nord d'Àfrica com els berbers de l'Oasi de Siwa a Egipte.

3) Poblacions del Nord-Oest d'Àfrica:

Aquesta regió, pretén ser especialment ben caracteritzada sobretot pel que fa al Magrib més occidental. Així destaca la bona representació tant berber com arabòfona de diferents regions del Marroc (des de les muntanyes del Rif fins a diferents valls de l'Alt Atlas, passant per poblacions berberòfones de l'Atlas Mitjà). En aquest grup, també s'inclou una població Mozabite d'Algèria.

4) Poblacions insulars del Mediterrani Occidental:

En aquest apartat s'inclouen poblacions de les principals illes del Mediterrani Occidental: Mallorca, Còrsega, Sardenya i Sicília. En general, amb subgrups de les diferents regions de les illes, com la part interior i costanera, o part occidental i oriental.

5) Finalment, s'inclouen dues poblacions externes a la regió circum-mediterrània que és el motiu d'estudi principal d'aquest treball, i amb un component genètic en principi clarament diferenciat. La funció d'aquests grups poblacionals és la de referents externs i comparatius per a l'anàlisi global. Es tracta d'una població centreuropea d'Alemanya i una població sud-sahariana de Costa d'Ivori.

La recol·lecció de mostres poblacionals utilitzades en el posterior anàlisi genètic, es va dur a terme seguint els protocols usuals i estrictes, i en base a criteris de tipus antropològic i històric ben establerts. Aquests criteris tenen com a objectiu garantir la representativitat antropològica de la mostra com a població homogènia i característica de la regió geogràfica estudiada:

- a) Els individus que conformen les diferents mostres van ser seleccionats a partir de donants de sang, aparentment sans, als quals se'ls va demanar el consentiment formal per a l'estudi genètic i poblacional, mitjançant una entrevista personal.
- b) Els individus no podien estar emparentats entre ells, com a mínim fins al quart grau de parentiu.
- c) El conjunt de persones mostrejades per a cada població i els seus avantpassats han d'haver nascut en la zona d'interès, incloent-hi almenys les tres darreres generacions (fins als quatre avis).

- d) Les mostres i poblacions seleccionades pertanyen a zones rurals. D'aquesta manera s'intenta evitar l'efecte d'incloure individus provinents de migracions recents cap a les grans ciutats que s'han produït en el darrer segle.

Aquestes mostres poblacionals van ser recol·lectades en diverses campanyes coordinades pel doctor Pedro Moral Castrillo i el seu grup investigador de la Unitat d'Antropologia de la Facultat de Biologia de la Universitat de Barcelona i pel professor Jean-Michel Dugoujon del Centre d'Anthropologie de la Université Paul Sabatier de Toulouse, seguint els protocols aprovats per les comissions de bioètica de les respectives institucions.

Descripció de les poblacions estudiades

En aquest moment, es procedirà a la descripció detallada de cada una de les mostres poblacionals utilitzades en l'estudi plantejat per al desenvolupament d'aquesta tesi doctoral.

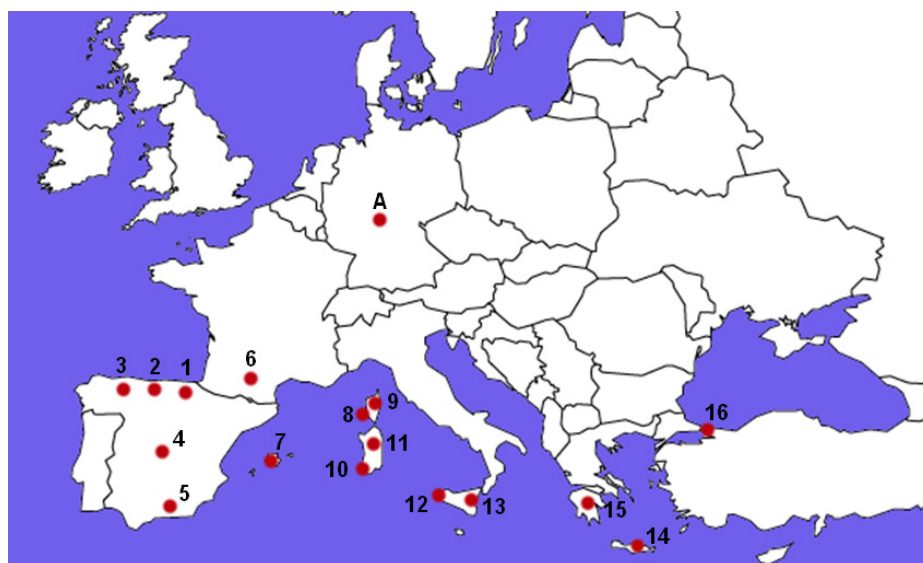


Figura A1. Localització geogràfica de les poblacions europees analitzades en l'estudi.

- 1: Bascos
- 2: Nord Peninsular: Asturians
- 3: Vall del Pas (Cantàbria)
- 4: Centre Peninsular (Serra de Gredos)

- 5: Sud Peninsular: Andalusia (Alpujarras)
- 6: Sud de França (Toulouse)
- 7: Illes Balears (Mallorca)
- 8: Còrsega - Costa: Ajaccio
- 9: Còrsega - Interior: Corte
- 10: Sardenya - Costa: Oristano
- 11: Sardenya - Interior: Nuoro
- 12: Sicília Occidental
- 13: Sicília Oriental
- 14: Creta
- 15: Grècia
- 16: Turquia
- A: Alemanya

a) Mediterrani Occidental Europeu:

1. **País Basc.** La mostra de bascos la componen un total de 112 individus, 57 homes i 55 dones. Tots els individus provenen de zones rurals de la província de Guipúscoa, al País Basc. La mostra va ser recollida durant una campanya realitzada pel Dr. Pedro Moral Castrillo de la Unitat d'Antropologia de la Facultat de Biologia de la Universitat de Barcelona i pel Dr. José Manuel Cárdenas del Banco de Sangre de Guipúscoa. L'ascendència basca de tots els individus mostrejats es va controlar mitjançant el requisit de possessió dels quatre cognoms dels pares d'origen basc, a més de mitjançant interrogació directa del recol·lector a cada una de les persones participants. Aquesta informació va permetre contrastar les dades obtingudes prèviament pels representants locals del Banco de Sangre de Guipúscoa.

2. **Nord Peninsular: Asturians.** Aquesta població pot considerar-se com una mostra significativa de la població actual del Nord de la península Ibèrica, sense particularitats especials com les dels bascos o les dels pasiegos. En aquest cas, la mostra està composta d'un total de 114 individus (93 homes i 21 dones) recollits en col·laboració amb la Dra. Carmen Rodríguez Menéndez de l'Hospital Central d'Astúries a Oviedo, seguint els criteris antropològics descrits anteriorment.

- 3. Vall del Pas.** La mostra de la Vall del Pas comprén 96 individus (67 homes i 29 dones) originaris, com a mínim en les últimes tres generacions, de la regió de Las Machorras (Burgos) al sud de la Vall de Pas. La mostra va ser recollida pel Dr. Mauricio de Grado Molinero del Centro de Salud de Espinosa de los Monteros (Burgos). La Vall del Pas és una regió muntanyosa del Nord de la península Ibèrica entre les províncies de Burgos i Cantàbria, formada per un conjunt de valls separades latitudinalment per la cadena muntanyosa, diferenciant geogràficament als pasiegos càntabres (2114 habitants al 2001) dels pasiegos burgalesos (2062 habitants al 2001). Els Pasiegos habiten en aquesta regió desde fa molt de temps. Es tracta d'una població ramadera que practica la transhumància estacional dins de cada vall. Aquest aspecte junt amb altres característiques socio-culturals particulars, ha afavorit la seva diferenciació dins l'àmbit peninsular. Encara que la presència humana a la zona és molt antiga, el poblament permanent de la regió dataria dels segle XI, quan ja hi han documents que acrediten la concessió de pastures de la regió. Desde el s. XVIII, la població ha oscil·lat al voltant dels 5000 habitants, i desde llavors, la població s'ha mantingut relativament aïllada amb un alt grau de consanguinitat i endogàmia. L'origen d'aquesta població és incert, però existeixen diversos postulats, alguns semblen indicar un origen directe en una població ancestral, altres evidències donen suport a un origen en poblacions germàniques refugiades a la vall davant la pressió de la invasió islàmica a la península, finalment altres dades apunten a que són descendents de moriscs i jueus refugiats a la zona per evitar l'expulsió durant el regnat dels reis catòlics.
- 4. Centre Peninsular.** Aquesta mostra es compon de 116 individus, 50 homes i 66 dones. La recol·lecció es va portar a terme en 2 campanyes seguint els criteris antropològics esmentats abans, en primer lloc la recopil·lació de parets d'una mostra de malalts cardiovasculars de l'Hospital de la Vall d'Hebron originaris de zones rurals de les províncies d'Àvila i Segòvia i, en segon lloc, la mostra es va completar amb mostres poblacionals de la província d'Àvila (Serra de Gredos), recollides en el desenvolupament del projecte PB88-0119 (DGCYT) dirigit per la Dra. M^a Soledad Mesa del Departamento de Zoología y Antropología Física de la Universidad Complutense de Madrid i amb la participació del Dr. Pedro Moral.

5. Sud Peninsular: Andalusia (Alpujarras). La regió muntanyosa de Las Alpujarras es localitza al Sud de Sierra Nevada. Aquesta mostra poblacional inclou 100 individus (53 homes i 47 dones), pertanyents als nuclis habitats de Ohanes, Berja, Adra i Laujar de Andarax, tots ells situats a la província d'Almeria. Aquesta mostra va ser recollida en campanya pel Dr. Pedro Moral Castrillo conjuntament amb el Dr. Francisco Luna de la Universidad Complutense de Madrid i en col·laboració amb la Dra. Gracia Villegas del Departamento de Hematología i la Hermandad de Donantes de Sangre del Hospital de Torrecárdenas d'Almeria, utilitzant unitats mòbils que es desplaçaven a cada una de les zones rurals on s'havia de recol·lectar. La regió de l'Alpujarra es pot considerar com un refugi en les repetides invasions que ha patit Andalusia al llarg de la història. Aquest fet ha condicionat la composició genètica final dels grups humans d'aquesta regió. Així mateix, es considera que l'Alpujarra hauria sigut una regió amb aports genètics considerables de la regió Nord-Africana derivats del període d'ocupació musulmana de la península Ibèrica.

6. Sud de França (Toulouse): Aquesta mostra inclou un número total de 56 individus (39 homes i 17 dones). La recol·lecció va ser portada a terme pel Dr. Francis Roubinet del Laboratoire d'Immunohématologie del Etablissement Français du Sang Pyrénées-Méditerranée de Toulouse, sota la coordinació del Prof. Jean-Michel Dugoujon. Aquesta mostra correspon a població francesa del Sud de França i es va recol·lectar com a element comparatiu en front de les mostres ibèriques analitzades en la present tesi doctoral.

b) Illes Mediterrànies:

7. Illes Balears (Mallorca). Els individus que componen la mostra poblacional de les Illes Balears van ser recol·lectats a l'illa de Mallorca. Amb aquesta mostra es pretén tenir una representació adequada de les poblacions humanes insulars més occidentals de la Mediterrània, i clarificar les seves possibles peculiaritats respecte a d'altres poblacions d'illes mediterrànies. El número d'individus que componen la mostra suma un total de 64, dels quals n'hi han 35 homes i 29 dones. Aquesta mostra s'ha obtingut en col·laboració amb el Prof. Laurent Varesi de la Faculté des Sciences et Techniques de la Université de Corse a Corte (Còrsega, França).

8-9. Còrsega: Corte (Interior) i Ajaccio (Costa). La mostra poblacional de Còrsega és divideix en dues submostres diferenciades. La primera correspon a 49 individus (16 homes i 33 dones) de la regió interior de Corte, una zona molt muntanyosa que ha afavorit l'aïllament cultural de la població al llarg de la història i que potser es remunta a períodes de la pre-història. La segona és una població de 48 individus (38 homes i 10 dones) de la província d'Ajaccio a la costa occidental corsa que, en principi, hauria mantingut un grau d'aïllament menor que la població corsa de l'interior, a causa dels contactes més o menys intensos amb diferents colonitzadors i navegants durant els temps històrics. La heterogeneïtat poblacional dins la illa, es fa evident a partir dels estudis lingüístics que mostren la subdivisió en diferents dialectes de la mateixa. D'altra banda, els estudis des del vessant genètic són encara escassos i poc conclusius, tot i que semblen indicar també aquesta diferenciació en regions. La recollida de les dues mostres va ser portada a terme pel Prof. Laurent Varesi del Laboratoire de Génétique et de Biologie Humaine de la Faculté des Sciences de la Université de Corse (França) seguint els criteris habituals en l'antropologia biològica.

10-11. Sardenya: Nuoro (Interior) i Oristano (Costa): La població de Sardenya – Costa correspon a una mostra recollida a la regió d'Oristano. Els individus inclosos són un total de 48 individus (22 homes i 26 dones). La província d'Oristano es troba a la costa occidental de l'illa i presenta característiques ecològiques de tipus lacustre. Això ha propiciat la permanència de la malària a la regió com a malaltia endèmica fins a mitjans del segle passat. La mostra anomenada Sardenya – Interior és localitzada a la regió de Nuoro i inclou en aquest anàlisi un total de 48 individus (38 homes i 10 dones). Podem centrar aquesta regió interior de Sardenya en el massís del Gennargentu, concretament en una zona muntanyosa de difícil accés i d'aïllament poblacional al llarg dels temps històrics. Aquestes mostres sardes van ser recollides en una campanya coordinada pel Prof. Giuseppe Vona del Dipartimento di Biologia Sperimentale, Sezione di Scienze Antropologiche de la Università degli Studi di Cagliari (Itàlia).

12-13. Sicília: Occidental i Oriental. Aquesta mostra insular també es divideix en dues submostres en funció de la seva localització geogràfica. Per una banda, una

mostra de la Sicília Occidental que compren 50 individus (14 homes i 36 dones), i de l'altra, una representació de la Sicília Oriental de 37 individus (22 homes i 15 dones). Les dues mostres van ser recolectades seguint criteris antropològics a partir de població reclusa de l'illa. La campanya va ser portada a terme pel Prof. Giuseppe Vona i els seus col·laboradors del Dipartimento di Biologia Sperimentale, Sezione di Scienze Antropologiche de la Università degli Studi di Cagliari (Itàlia).

- 14. Creta.** Aquesta mostra insular del Mediterrani oriental inclou un total de 96 individus (74 homes i 22 dones). L'illa de Creta va estar poblada des del vuitè mil·lenni a.C. Entre els seus pobladors antics destaquen la cultura Minoica (entre el 3000 i 2000 a.C), els aqueus que imposen la cultura micènica (1500 a.C) i els doris del nord de Grècia (1150 a.C.). L'any 67 va ser conquerida pels romans i entre el 395 i el 1204 va romandre en mans bizantines.

c) Mediterrani Europeu Oriental:

- 15. Grècia.** La mostra poblacional grega comprén un total de 105 individus, 62 homes i 43 dones, recollides seguint els criteris habituals en hospitals de la ciutat d'Atenes. Aquesta és una mostra significativa i característica de població general actual de Grècia. Els responsables del treball de recol·lecció pertanyen al Department of Pathology de la Medical School de la University of Athens (Grècia) i van ser dirigits per la Dra. Angelica Saetta.
- 16. Turquia:** Formen part d'aquesta població turca un total de 100 individus (36 homes i 64 dones), estudiants de la Fatih University d'Istanbul, tots ells individus turcs, amb cognoms d'origen turc i amb un mínim de tres generacions contrastades d'aquella regió. La recol·lecció va ser coordinada per la Dra. Nisrine Bissar-Tadmouri del Department of Biology de la Fatih University d'Istanbul (Turquia). Aquesta mostra pot considerar-se com a representativa de la població turca, i com un altre representant de les poblacions humanes del Mediterrani Oriental, junt amb la tradicionalment contrastada població grega.

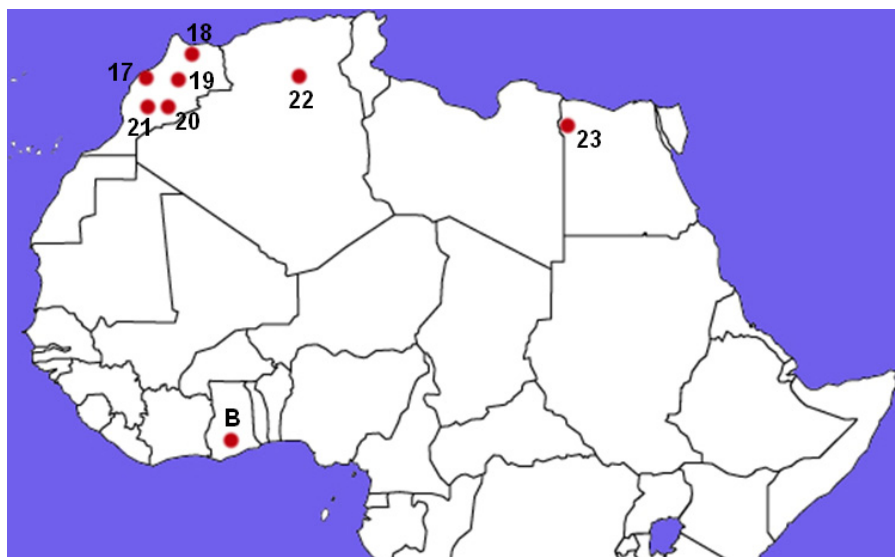


Figura A2. Localització geogràfica de les poblacions africanes analitzades en l'estudi.

- 17: Àrabs de Marroc (Franja Central)
- 18: Berbers del Marroc Nord-Oriental (Sidi Bouhria)
- 19: Berbers de l'Atlas Mitjà: Centre Marroc (Khénifra)
- 20: Berbers de l'Alt Atlas: Sud de Marroc (Vall d'Asni)
- 21: Berbers de l'Alt Atlas: Sud Marroc (Vall d'Amizmiz)
- 22: Berbers d'Algèria: Mozabites
- 23: Berbers d'Egipte: Oasis de Siwa
- B: Costa d'Ivori (Ahizi)

d) Nord d'Àfrica:

16. Àrabs de Marroc (regió costanera central). La mostra àraboparlant del Marroc comprèn 101 individus en que la pertinença al grup etnolingüístic àrab s'ha comprovat remuntant-se com a mínim a tres generacions. La població recol·lectada inclou 58 homes i 43 dones, pertanyents a diferents localitats de la franja central marroquina: Chaouia, Doukkala, Abda i Tadla. La campanya de recollida de sang va necessitar la col·laboració de diferents hospitals de Marroc a través de la coordinació del Dr. Mustafà Kandil del Département de Biologie de la Faculté des Sciences de la Université Chouaib Doukkali de El Jadida (Marroc).

17. Berbers del Marroc Nord-Oriental (Sidi Bouhria). Aquesta població és localitzada a la província de Berkane i la major part del seu contingent prové de la tribu Beni Snassen una mica més al nord. En tot cas, es tracta de població parlant berber, de les estribacions nord-orientals de l'Atlas. S'ha seleccionat aquesta regió per la seva diversitat de tribus berbers i perquè és l'origen del crani de l'home de Tafoghalt. La població inclou un total de 128 individus (65 homes i 63 dones) i va ser recol·lectada en coordinació amb el Professor Mohamed Melhaoui de la Faculté des Sciences d'Oujda de la Université Mohamed I (Marroc). Com en tots els casos, es va assegurar el compliment dels criteris antropològics ja esmentats per garantir la representativitat de la mostra.

18. Berbers de l'Atlas Mitjà: Centre Marroc (Khénifra). Dins d'aquesta mostra s'inclouen individus d'origen ètnic berber i de parla també berber (Tamazight) que es remunten com a mínim a les tres últimes generacions. S'inclouen un total de 140 individus, equitativament repartits en els dos sexes. Els individus que componen la mostra van ser recol·lectats a la ciutat de Khénifra, considerada com a capital de totes les tribus berbers de l'Atlas Mitjà (Moyen-Atlas). Tots els components de la mostra pertanyen a aquesta regió que inclou 4 confederacions berbers (grups de tribus similars): Zayane, Ichkern, Ait Sgougou i Beni M'Guild. La campanya de recol·lecció va ser duta a terme pel Dr. Nourdin Harich del Département de Biologie de la Faculté des Sciences de la Université Chouaib Doukkali d'El Jadida (Marroc).

19 i 20. Berbers de l'Alt Atlas: Sud Marroc (Vall d'Asni i Vall d'Amizmiz). Els berbers de l'Alt Atlas marroquí comprenen un ric i autòcton grup de pobles amb una cultura que es remunten a milers d'anys enrera. Principalment pertanyen a tribus del grup Masmouda. Les condicions climàtiques junt amb el difícil accés a les valls de l'Alt Atlas han afavorit l'isolament d'aquests grups i han contribuït a la conservació de la seva cultura mil·lenària. Les localitats emprades en el nostre estudi pertanyen a dues valls independents de l'Alt Atlas: la vall d'Asni i la vall d'Amizmiz. És tracta de vil·les berbers situades a més de 1800 metres d'alçada en dues valls paral·leles al sud de Marrakech. La primera mostra, d'Asni, comprèn 135 individus, 59 homes i 76 dones. La segona mostra, d'Amizmiz, inclou 85 representants, 29 homes i 56 dones. La recol·lecció va ser portada a terme pel Prof.

Mohamed Cherkaoui del Laboratoire d'Ecologie Humaine de la Faculté des Sciences Semlalia de la Université Cadi Ayyad de Marrakech (Marroc).

21. Berbers d'Algèria: Mozabites. Els Berbers Mzab provenen de cinc vil·les principals localitzades en una àrea al voltant d'un oasi del Sahara, a 500 km al Sud d'Alger. Aquesta regió d'Algèria també es coneix amb el nom de Mzab i està habitada per un grup de berbers que parlen una branca particular de Tamazight, el dialecte Tumzabt que també presenta un tipus d'escriptura particular. Actualment es calcula que són uns 220 000, són musulmans i encara conserven un sistema social especialment rígid amb la família com a element vertebrador de primer ordre. En la nostra mostra estan representats un total de 107 individus antropològicament ben caracteritzats (68 homes i 39 dones). La recol·lecció va ser portada a terme pel Prof. Issad i el Dr. Hariti del Laboratoire d'Immunologie de la Faculté de Medecine d'Alger et Ghardaia.

22. Berbers d'Egipte: Oasis de Siwa. Aquesta població és especialment interessant doncs es localitza en l'extrem més oriental del territori berber i perquè parlen Siwi, un dialecte particular del Berber o Tamazight. L'oasis de Siwa es troba en el desert de Líbia a 300 km de la costa mediterrània i a 25 km de la frontera amb Líbia. La regió està habitada des del Paleolític superior i s'especula que els primers pobladors provindrien d'una migració de protoberbers del nord-oest d'Àfrica cap a Líbia i Egipte (5000-6000 AC). Les restes arqueològiques trobades guarden una certa semblança amb les de la cultura capsiana, però els especialistes tendeixen a classificar-les dins del subgrup Líbico-Capsià, diferenciat de la indústria ateriense típica de les poblacions proto-berbers prehistòriques del Magreb. Es van recol·lectar un total de 86 mostres de sang (55 homes i 31 dones), amb la col·laboració del Prof Farha El Chennawi de la Faculté de Médecine de la Université de Mansoura (Egipte).

e) Poblacions Extra-Mediterrànies:

A. Alemanya: La població alemana analitzada està composta de 48 individus (26 homes i 22 dones) d'entre 26 i 60 anys d'edat, recollits pel Dr. Heiko Witt de la Medizinische Fakultät der Humboldt de la Universität zu Berlin. En aquest cas

només es va comprovar l'origen dels pares (no es disposava de les dades dels avis) i és pot considerar com a mostra representativa de població general d'Alemanya.

B. Àfrica Sub-Sahariana: Costa d'Ivori (Ahizi). La població de Costa d'Ivori utilitzada com a referent extern sub-saharià, pertany a la ètnia Ahizi del grup dels Akans. Són un total de 122 individus (50 homes i 72 dones) recol·lectats al poblat de Nigui-Saff localitzat a les rodalies d'Abidjan. La mostra va ser recol·lectada pel Prof. Chaventre del Laboratoire d'Anthropologie et Démographie Génétiques de la Université de Bordeaux 2 (França).

ANNEX 2

Base de dades per a l'estudi poblacional comparatiu

Conjuntament amb les dades resultants de l'anàlisi directe al laboratori de les poblacions esmentades anteriorment. Un cop finalitzat el treball experimental, es va procedir a la selecció bibliogràfica de dades sobre la variació de freqüències d'un total de 8 elements *Alu* polimòrfics a diferents poblacions de tot el món. Aquestes dades provenen de diferents estudis ja publicats per diferents grups investigadors, on s'analitzen poblacions antropològicament ben caracteritzades de diversos continents.

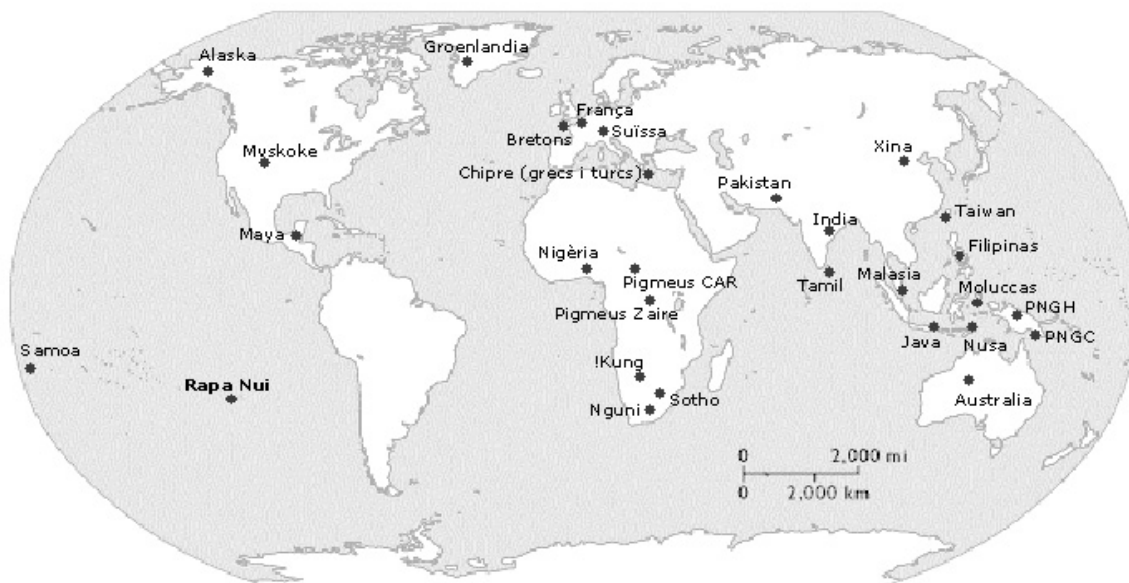


Figura A3. Poblacions seleccionades de la bibliografia per a l'anàlisi comparatiu a nivell mundial (basat en la variació per a 8 elements *Alu* polimòrfics).

La nostra intenció, és disposar d'una taula de freqüències de les insercions *Alu* més utilitzades a la bibliografia i que també van ser analitzades en les nostres poblacions, per així permetre'ns ampliar el nostre anàlisi poblacional i de reconstrucció filogenètica a tota l'espècie humana, per tal d'inferir patrons de migració i d'evolució de les poblacions i grups humans actuals. D'aquesta manera, després d'un anàlisi exhaustiu de la bibliografia actual sobre el tema es va construir una taula de freqüències de vuit polimorfismes *Alu* en un total de 30 poblacions humanes que s'uneixen a les dades ja disponibles per a les nostres 25 poblacions recollides en aquest treball de tesi doctoral.

Taula A1. Poblacions mundials i les seves freqüències per a 8 elements *Alu* polimòrfics (CAR: República Crentre-Africana; PNGH: Terres altes de Papua Nova Guinea; PNGC: Costa de Papua Nova Guinea)

	N	<i>ACE</i>	<i>TPA25</i>	<i>APO</i>	<i>FXIII B</i>	<i>A25</i>	<i>PV92</i>	<i>D1</i>	<i>B65</i>
Alaska	41	0,580	0,300	0,920	0,920	0,150	0,620	0,420	0,450
Austràlia	69	0,910	0,130	0,870	0,650	0,350	0,150	0,040	0,390
Bretons	54	0,480	0,560	0,900	0,400	0,160	0,270	0,390	0,560
Xina	49	0,670	0,350	0,820	0,710	0,100	0,860	0,170	0,350
Greco-Xipriotes	48	0,390	0,530	0,950	0,620	0,120	0,250	0,270	0,650
Turco-Xipriotes	33	0,330	0,580	0,980	0,390	0,090	0,330	0,350	0,640
Filipines	47	0,530	0,630	0,980	0,720	0,140	0,800	0,360	0,570
França	53	0,480	0,560	0,990	0,420	0,160	0,230	0,460	0,570
Groenlàndia	41	0,550	0,330	0,940	0,790	0,170	0,610	0,450	0,190
Índia	28	0,520	0,340	0,850	0,660	0,050	0,520	0,100	0,350
Java	32	0,860	0,390	0,780	0,920	0,060	0,840	0,420	0,580
!Kung	40	0,290	0,170	0,880	0,170	0,610	0,200	0,160	0,500
Malàisia	47	0,640	0,500	0,760	0,730	0,020	0,720	0,270	0,420
Maia	51	0,680	0,650	0,940	0,900	0,210	0,790	0,450	0,270
Moluccas	48	0,670	0,560	0,760	0,780	0,000	0,690	0,190	0,260
Mvskoke	50	0,700	0,490	0,960	0,790	0,210	0,570	0,460	0,480
Nguni	43	0,400	0,210	0,600	0,120	0,410	0,240	0,270	0,600
Nigèria	11	0,270	0,410	0,500	0,080	0,220	0,090	0,000	0,830
Pakistan	42	0,440	0,510	0,720	0,230	0,070	0,300	0,170	0,370
PNGC	48	0,660	0,160	0,660	0,300	0,020	0,360	0,170	0,270
PNGH	68	0,740	0,160	0,680	0,300	0,040	0,240	0,010	0,180
Pigmeus CAR	17	0,120	0,210	0,740	0,000	0,350	0,260	0,470	0,780
Pigmeus Zaire	17	0,320	0,240	0,850	0,030	0,530	0,350	0,590	0,820
Sotho	48	0,380	0,330	0,680	0,180	0,390	0,290	0,310	0,480
Suïssa	43	0,370	0,450	0,940	0,480	0,120	0,200	0,340	0,580

Taiwan	46	0,500	0,640	0,930	0,970	0,220	0,900	0,380	0,540
Tamil	47	0,690	0,560	0,810	0,610	0,170	0,560	0,340	0,550
Nusa Tengarras	90	0,640	0,380	0,780	0,810	0,050	0,500	0,190	0,400
Samoa	48	0,890	0,480	0,710	0,910	0,010	0,530	0,190	0,450
Rapa Nui	54	0,862	0,489	0,407	0,930	0,087	0,463	0,526	0,488

ANNEX 3

Marcadors genètics i anàlisi molecular al laboratori

Insercions Alu

En total, han estat genotipades en el conjunt de poblacions analitzades, 19 insercions *Alu* polimòrfiques en l'espècie humana. Es tracta d'un conjunt divers de polimorfismes *Alu* seleccionats en funció de la seva utilitat en la reconstrucció de la genètica de poblacions humanes i distribuïdes homogèniament pel genoma.

L'anàlisi de la variació per presència/absència d'insercions *Alu* és relativament senzill, només requereix l'amplificació per PCR de cada un dels marcadors, seguida d'una electroforesi en gel d'agarosa d'entre l'1 i el 2 %. D'aquesta manera, s'obté una banda gran per la presència d'*Alu*, o una petita per l'absència d'inserció (els individus heterozigots presenten les dues bandes). La detecció es realitza per tinció del gel en una solució de bromur d'etidi (10 mg/ml) durant un temps aproximat de 30 minuts, seguida de la visualització de les bandes sota la llum ultraviolada.

A la taula A2, es mostra el llistat de les insercions *Alu* estudiades, els oligonucleòtids utilitzats per a l'amplificació per PCR, la temperatura d'*annealing* per a cada marcador, la mida de bandes obtingudes en la PCR, la localització cromosòmica de cada seqüència *Alu*, i la referència bibliogràfica originària en que va ser descrita.

L'amplificació per PCR es va realitzar en termocicladors Perkin-Elmer 9600 (Applied Biosystems), Perkin-Elmer 9700 i Touchgene Gradient (Techne). Els volums de PCR per mostra van ser de 20 µl i la següent composició estandard: 250 ng DNA, 1 U Taq Polimerasa, 200 µM dNTPs, 10 pmol de cada oligonucleòtid, 1.5 mM MgCl₂. Els cicles de PCR van consistir en 5 minuts de desnaturalització a 94 °C; seguits de 30-35 cicles de 60 segons de desnaturalització a 94°C, 60 segons d'anellament a la temperatura d'*annealing* corresponent i 90 segons d'elongació a 72 °C. En algun cas, la composició de

la PCR es va completar amb un 10% de DMSO, per facilitar l'accés de la polimerasa i l'extensió de les cadenes de DNA en situacions de replegament i estructura secundària.

Algunes de les seqüències *Alu* analitzades van ser multiplexades per parelles, de manera que en una mateixa PCR, s'obtenia l'amplificació de dos marcadors (això va ser possible en els casos en que la mida de les bandes obtingudes era perfectament diferenciable i quan les temperatures d'*annealing* eren coincidents). Concretament les parelles multiplexades van ser: Sb19.12 i HS2.43; FXIIIB i A25; PV92 i D1.

Taula A2. Característiques d' amplificació dels polimorfismes *Alu* analitzat.

Element <i>Alu</i>	Oligonucleòtids	T ^a an(°C)	Crom	Referència	Al·lels
<i>TPA 25</i>	5'-GTAAGAGTTCCGTAACAGGACAGCT-3' 5'-CCCCACCTAGGAGAACCTCTCTTT-3'	58	8	Arcot 1995a Batzer 1996	570 260
<i>ACE</i>	5'-CTGGAGACCACTCCCATCCTTTCT-3' 5'-GATGTGGCCATCACATTCGTTCAGAT-3'	58	17	Arcot 1995a Batzer 1996	490 190
<i>APO A1</i>	5'-AAGTGCTGTAGGCCATTTAGATTAG-3' 5'-GTAAGAGTTCCGTAACAGGACAGCT-3'	50	11	Arcot 1995a Batzer 1996	~450 ~150
<i>FXIIIB</i>	5'-TCAACTCCATGAGATTTTCAGAAAGT-3' 5'-CTGGAAAAAATGTATTACAGGTGAGT-3'	63	1	Arcot 1995a Batzer 1996	~800 ~500
<i>PV92</i>	5'-AACTGGGAAAATTGAAGAAAAGT-3' 5'-TGAGTTCTCAACTCCTGTGTGTTAG -3'	54	16	Arcot 1995a Batzer 1996	443 129
<i>D1</i>	5'-TGCTGATGCCAGGGTTAGTAAA-3' 5'-TTTCTGCTATGCTCTTCCCTCTC-3'	70	3	Arcot 1995a Batzer 1996	~550 ~250
<i>CD4</i>	5'-AGGCCTTGTAGGGTTGGTCTGATA-3' 5'-TGCAGCTGCTGAGTGAAAGAAGCTG-3'	50	12	Edwards 1992	1501 1249
<i>B65</i>	5'-ATATCCTAAAAGGGACACCA-3' 5'-AAAATTTATGGCATGCGTAT-3'	60	11	Arcot 1995a Batzer 1996	410 115
<i>YAP</i>	5'-CAGGGGAAGATAAAGAAATA-3' 5'-ACTGCTAAAAGGGGATGT-3'	51	Y	Hammer 1994	455 150
<i>Sb19.3</i>	5'-TCTAGGCCAGATTTATGGTAACTG-3' 5'-AAGCACAATTGGTTATTTCTGAC-3'	60	19	Arcot 1998	456 154
<i>Sb19.12</i>	5'-TTAACATCCCTGCAACCCATC -3' 5'-GATTATAGTCACCCGTGTGTGC -3'	65	19	Arcot 1998	~600 ~300
<i>A25</i>	5'-CCACAAATAGGCTCATGTAGAAC -3' 5'-TATAATATGGCCITGGATTATACC -3'	63	8	Arcot 1995b	~550 ~250

<i>HS2.43</i>	5'-ACTCCCCACCAGGTAATGGT-3'	64	1	Arcot 1996	482
	5'-AGGGCCITTCATCCAGTTTGT-3'				184
<i>HS4.69</i>	5' - GTC TGA ATG TTC GTG TCC CC- 3'	64.5	6	Arcot 1996	601
	5' - GTC CAA GTT CAA GGC ACC AG- 3'				289
<i>HS4.32</i>	5' - GTT TAT TGG GCT AAC CTG GG- 3'	60	12	Arcot 1996	572
	5' - TGA CCT GCT AAC TTG TAC TTT AAC C- 3'				262
<i>DM</i>	5'- CTG TAT ACT CAG CTA CTA GGG T -3'	60	19	Mahadevan 1993	1008
	5'- AAA TAG GCT GGA CCG CGG -3'				494
	5'- CTC AGG GGT TAT CTA AAG TGG C -3'				
<i>Yb8NBC120</i>	5'-CAGTGGATCTCCATTTTACCTCTC-3'	62	22	Carroll 2001	532
	5'-GGAAAAGGTTTCAGGAAGAAAAGTG-3'				212
<i>Yb8NBC125</i>	5'-AGCCAGAAACCCTGAACAAG-3'	62	22	Carroll 2001	415
	5'-AAAGGCCCCAGAAGTATACCA-3'				97
<i>Ya5NBC221</i>	5'-CAGTTTTTCATATACATGTGGGTTTC-3'	60	22	Carroll 2001	640
	5'-TAGTGTTAAGAGGCCCATTTTCTAC-3'				313

A les figures A4 i A5 es mostren dos exemples del resultat de l'electroforesi d'uns gels d'agarosa fruit d'una PCR normal i una PCR de marcadors *Alu* multiplexats.

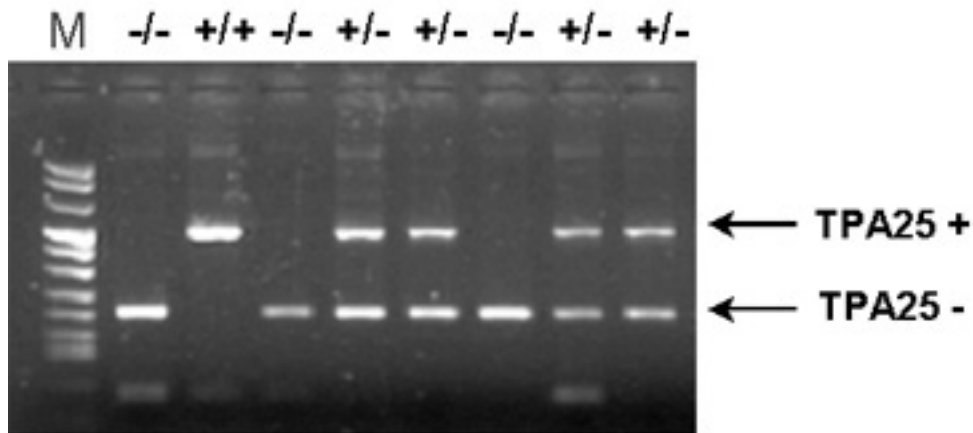


Figura A4. Gel d'agarosa obtingut per electroforesi dels productes de PCR del marcador *Alu* TPA25 (+: inserció *Alu*; -: absència d'inserció).

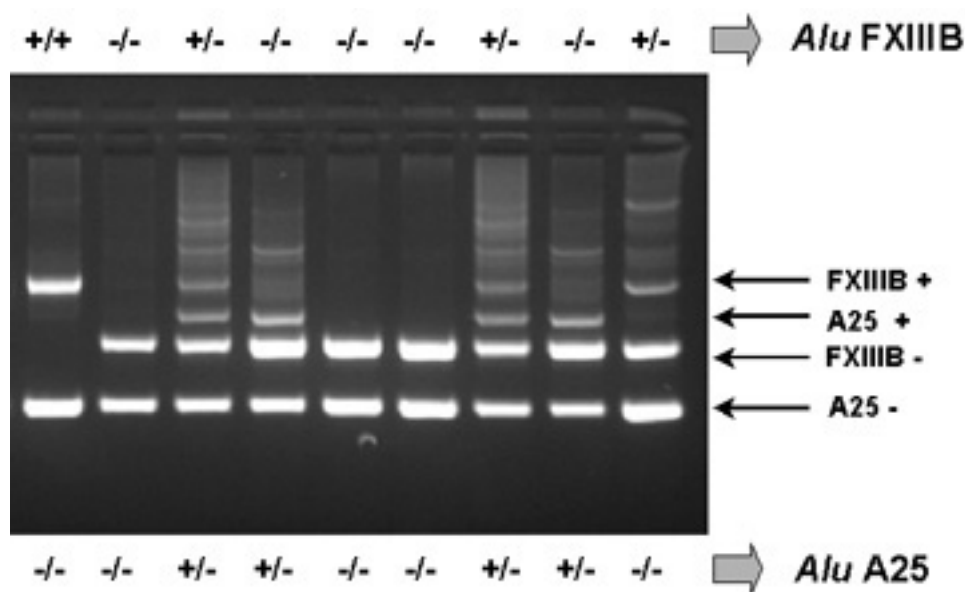


Figura A5. Gel d'agarosa obtingut per electroforesi dels productes de PCR-Multiplex dels marcadors *Alu* FXIII B i *Alu* A25 (+: inserció *Alu*; -: absència d'inserció).

Microsatèl·lits associats a elements Alu

Dins d'aquest grup van ser analitzats un total de 3 polimorfismes de tipus STR amb la intenció de reconstruir posteriorment, i en combinació amb elements *Alu* del mateix locus, haplotips altament informatius per a la reconstrucció de les històries poblacionals. Els marcadors seleccionats són: CD4 (CCCCT)_n, FXIII B (TTTA)_n i DM (CTG)_n. A la figura A6 es pot observar una reconstrucció esquemàtica d'aquests marcadors en combinació amb un element *Alu* en tres *loci* específics del genoma.

Al *loci* DM, localitzat al cromosoma 19, la inserció i el trinucleòtid es troben separats una distància de 5 kb. Pel que fa al *loci* FXIII B, en el cromosoma 1, la distància que separa l'*Alu* del tetranucleòtid, correspon aproximadament a 4100 parells de bases. Finalment, en el *loci* CD4, del braç petit del cromosoma 12, el pentanucleòtid està situat a 9,8 kb de la inserció polimòrfica. Aquestes distàncies globals, inferiors sempre a les 10 kb, fan pensar que el grau de desequilibri entre els marcadors de l'haplotip ha de ser considerable o pràcticament complet i, en conseqüència, permet un anàlisi haplotípic i evolutiu com el dut a terme en aquesta tesi doctoral.

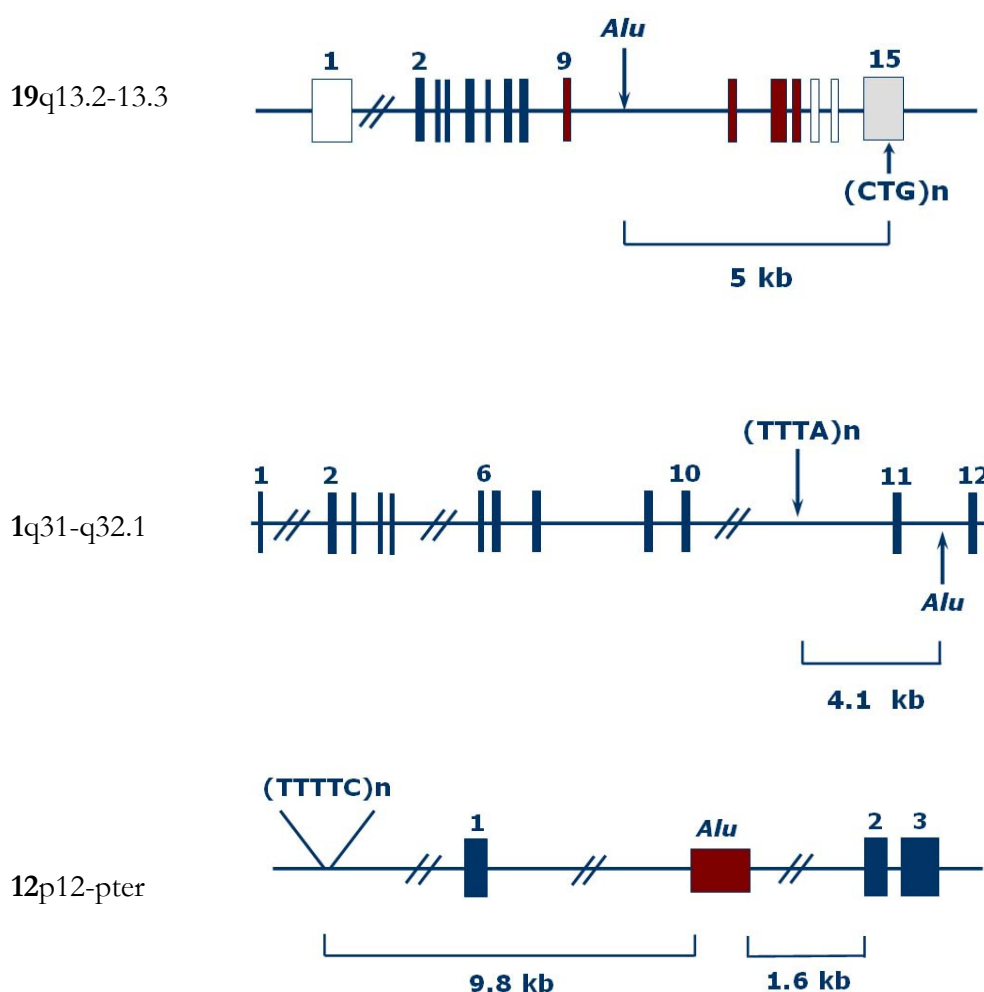


Figura A6. Esquema dels 3 loci analitzats en l'anàlisi d'haplotips *Alu*/STR (DM, FXIIB i CD4).

Aquests marcadors reben el nom de microsatèl·lits o STR's (*Short Tandem Repeats*). Els al·lels d'aquests polimorfismes es diferencien pel nombre de repeticions d'una petita seqüència, d'entre 3 i 5 nucleòtids, en el nostre cas, que dona el nom al polimorfisme. Degut a l'elevada variabilitat d'aquests marcadors (alguns presenten més de 20 al·lels) i la precisió requerida per diferenciar entre al·lels, es va optar pel seu anàlisi mitjançant amplificació per PCR amb un oligonucleòtid marcat amb fluorescència (VIC, PET, NED o 6-FAM) i posterior detecció dels al·lels a través d'un analitzador de microfragments.

El disseny de les parelles d'oligonucleòtids es va desenvolupar tenint en compte que els productes d'amplificació dels diferents polimorfismes no presentessin rangs de mida superposats i, quan els rangs eren propers, que fossin marcats amb fluorocroms que

emetessin llum de longitud d'ona diferent (verda per VIC, vermella per PET, groga per NED i blava per 6-FAM). Aquest disseny experimental va permetre que cada mostra preparada per l'analitzador de fragments contingués una barreja de productes d'amplificació de tots els microsatèl·lits analitzats, per tal d'optimitzar els anàlisis i les despeses globals.

En tots els casos, l'oligonucleòtid marcat corresponia al *forward*, que contenia al seu extrem 5' un fluorocrom. El fluorocrom emprat per cada polimorfisme està indicat amb la seqüència dels oligonucleòtids a la taula A3. Els STR's FXIIIIB (TTTA)_n i DM (CTG)_n van ser amplificats mitjançant una PCR multiplex amb una barreja de les dues parelles d'oligonucleòtids.

Taula A3. Característiques d'amplificació dels polimorfismes STR's lligats a elements *Alu* analitzats.

STR	Oligonucleòtids	T ^a an.	Crom.	Referència	Al·lels
<i>penta CD4</i>	VIC 5'-GCCTGAGTGACAGAGTGAGAACC -3' 5'-TTGGAGTCGCAAGCTGAACTAGC -3'	55	12	STRbase	81-141 bp (3-15 rep)
<i>tetra FXIIIIB</i>	PET 5'-TGAGGTGGTGTACTACCATA -3' 5'-GATCATGCCATTGCACCTCA -3'	58.8	1	STRbase	168-192 bp (6-12 rep)
<i>tri DM</i>	NED 5'-GCTCGAAGGGTCCCTGTAGCCGGGAA -3' 5'-CCTGGCCGAAAGAAAGAAATGGTCT -3'	58.8	19	Mahadevan 1992	77-143 bp (5-27 rep)

Un cop realitzades les amplificacions, es comprovaven amb una electroforesi en gel d'agarosa al 1.5%. Si les amplificacions havien estat satisfactòries, es procedia a la preparació dels productes d'amplificació per ser analitzats simultàniament en l'analitzador de microfragments, elaborant una solució que contenia:

- 1 microlitre dels diferents productes d'amplificació (diluits aproximadament 1/8 en aigua destil·lada);
- 10,5 microlitres de formamida desionitzada;
- 0,5 microlitres de marcador de pes molecular LIZ500 (Applied Biosystems), marcat amb fluorocrom LIZ.

A continuació, les mostres es sotmetien a 100°C durant 5 minuts per tal de garantir la desnaturalització ja iniciada per l'efecte que la formamida realitza sobre el DNA. Immediatament després, les mostres eren dipositades en gel.

L'analitzador de fragments emprat en les determinacions va ser un seqüenciador de DNA per electroforesi capil·lar ABI PRISM 3730 (Applied Biosystems) dels Serveis Científico-Tècnics de la Universitat de Barcelona. En l'analitzador, les mostres eren sotmeses a una electroforesi en capil·lars que contenen el polímer POP-6 (Applied Biosystems) per separar els al·lels de cada polimorfisme en funció de la longitud del número de repeticions. Aquesta separació va acompanyada per un sistema de lectura que utilitza un làser que excita els fluorocroms amb que estan marcats els oligonucleòtids i per un filtre que llegeix la llum emesa per cada fluorocrom. D'aquesta lectura se'n derivaven electroforogrames que van analitzar-se amb el programa GeneMapper 3.0 (Applied Biosystems). L'addició del marcador de pes molecular marcat amb LIZ (que emet llum taronja) permetia l'assignació de tamanys (en pb) dels al·lels de cada marcador. En la figura A7 es mostra un electroforograma amb un exemple de lectura dels microsatèl·lits analitzats.

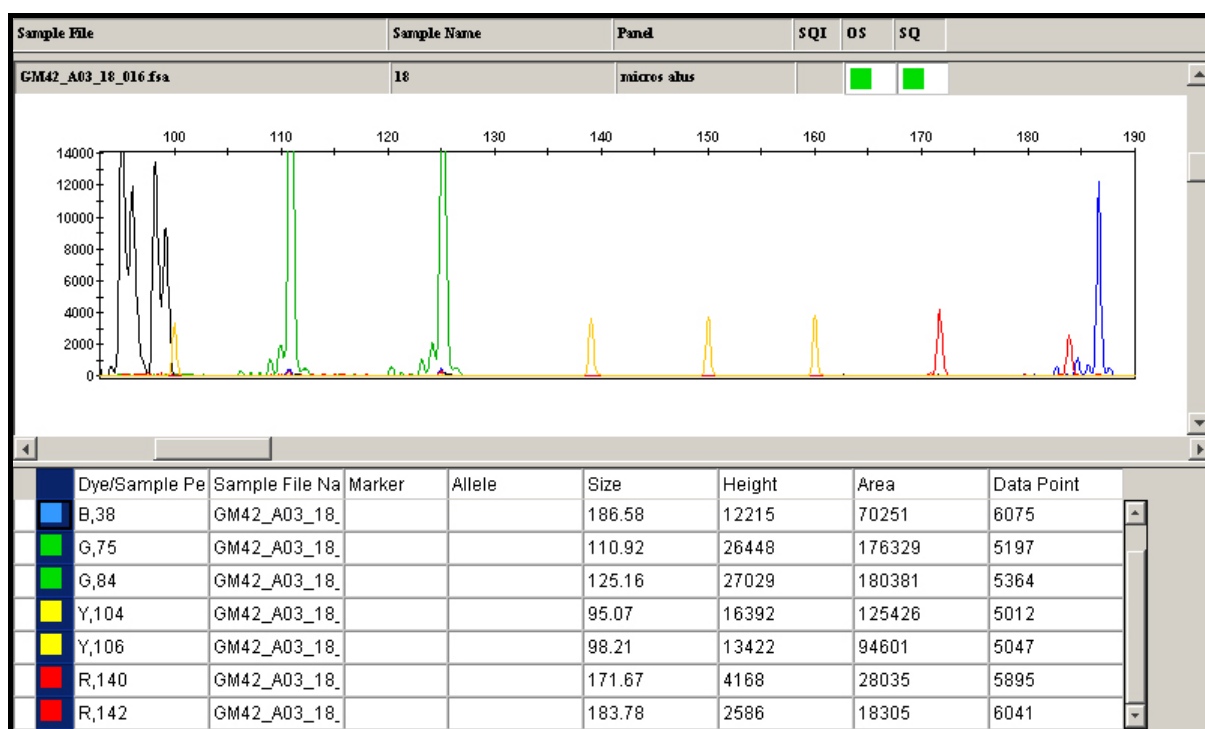


Figura A7. Electroforograma obtingut per als productes de PCR dels STR's tri-DM (color negre, groc a la taula), penta-CD4 (color verd), tetra-FXIIB (color vermell) i tetra-DYS19 (color blau). El marcador LIZ500 apareix en color taronja.

Per tal de confirmar les assignacions de les mides reals dels al·lels, es van seqüenciar per cada microsatèl·lit diversos individus homozigots. Un exemple d'aquest tipus de comprovació prèvia per seqüenciació es mostra a la figura A8.

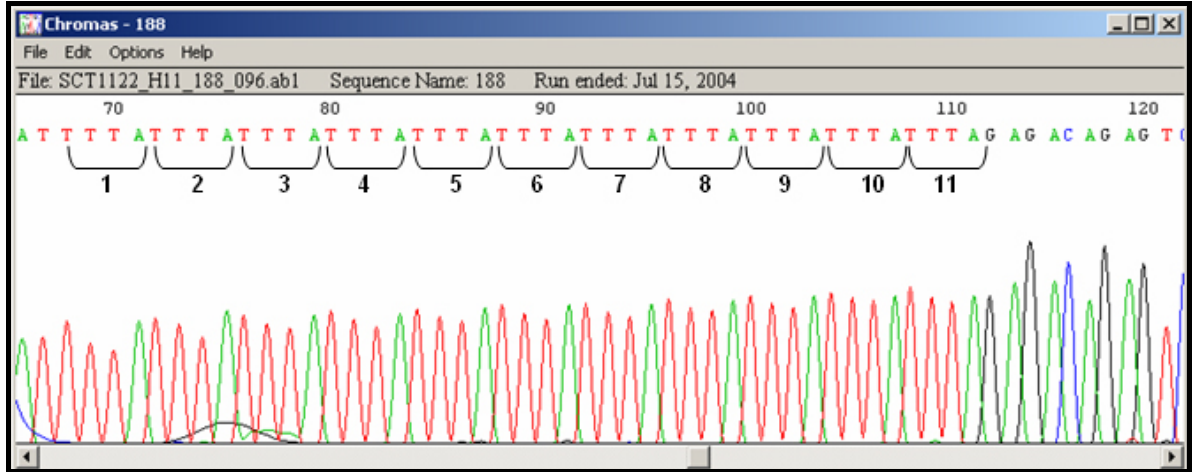


Figura A8. Comprovació per seqüenciació directa d' homozigots de la correspondència entre número de repeticions del STR FXIIB (11 repeticions TTTA) i la mida de l'al·lel (188 pb).