



**Universitat de les
Illes Balears**

Facultat de Ciències

Memòria del Treball de Fi de Grau

STRs EN LA INVESTIGACIÓ BIOLÒGICA DE LA PATERNITAT: INTEPRETACIÓ I VALORACIÓ DELS RESULTATS

Maria Francisca Bestard Vachiano

Grau de Bioquímica

Any acadèmic 2014-15

DNI de l'alumne: 41570834-J

Treball tutelat per Antònia Picornell Rigo
Departament de Genètica



S'autoritza la Universitat a incloure el meu treball en el Repositori Institucional per a la seva consulta en accés obert i difusió en línia, amb finalitats exclusivament acadèmiques i d'investigació

Paraules clau del treball:
STRs, Genètica Forense, Paternitats, Artefactes.

Abstract

Els *Short Tandem Repeats* (STRs) són regions polimòrfiques del DNA, utilitzats com a marcadors en la Genètica Forense, gràcies a la seva variabilitat entre els individus. L'objectiu del treball va ser analitzar de manera global i anònima els resultats obtinguts, tant de les assignacions positives, com de les exclusions, a 25 casos de paternitats fets a la UIB. Els marcadors que han mostrat un major poder de discriminació són el Penta E i el D2S1338 i el sistema que ha presentat menys poder de discriminació és el TPOX. A més, s'han analitzat un total de 8 artefactes biològics i tècnics. L'artefacte més freqüent és l'*stutter* i, en general, s'observa que aquells marcadors que presenten major complexitat tenen uns pics *stutter* superiors que els sistemes més simples, però sempre inferiors al 15%. El segon artefacte més freqüent és l'adenilació incompleta, però tan sols en un 8,1% l'alçada del pic superava en un 30% el pic original. Els *pull-up*, *drop-out*, *allele imbalance*, *locus-to-locus imbalance* i els artefactes tècnics s'han trobat amb menor freqüència i no han dificultat la genotipació dels STRs. Finalment, els al·lels *off-ladder*, els *drop-in* i els patrons de triple pic no s'han trobat en cap dels casos.

ÍNDEX

1. Introducció	
1.1. <i>Short tandem repeats</i> (STRs).....	pàg.3
1.2. STR en la Genètica Forense.....	pàg.6
1.3. Interpretació dels resultats i artefactes.....	pàg.7
2. Materials i Mètodes	
2.1. Procediment general	pàg.13
2.2. <i>Kits</i> emprats.....	pàg.14
2.3. Anàlisi de fragments.....	pàg.15
2.4. Anàlisi estadístiques.....	pàg.18
3. Objectius.....	pàg.20
4. Resultats	
4.1. Resultats de l'estudi de paternitat.....	pàg.21
4.2. Artefactes.....	pàg.25
5. Conclusions.....	pàg.31
6. Bibliografia.....	pàg.32

1. INTRODUCCIÓ

1.1- *Short Tandem Repeats* (STRs)

Els *Short Tandem Repeats* (STRs) també anomenats microsatèl·lits o *single sequence repeats* (SSRs) són fragments de DNA polimòrfic organitzats en unitats de 2 a 6 nucleòtids que es van repetint, quantitat que pot variar des de dues a una trentena de vegades. Els STRs es troben en abundància en el genoma humà, on es distribueixen de manera atzarosa tant en els cromosomes autosòmics com als cromosomes sexuals, així com també en el DNA mitocondrial (mtDNA) [21,40].

En l'organització del genoma humà, els STRs es classifiquen com a DNA repetitiu no codificant que es repeteix de manera moderada i en tàndem (**Figura 1**). Habitualment es troben formant part d'introns o regions intergèniques. Encara que la funcionalitat d'aquest tipus de DNA no es trobi molt definida, el que sí podem afirmar és que han adquirit una gran importància en els estudis de paternitat i en proves de identificació gràcies a la seva variabilitat entre els diferents individus de la població [27].

En estudis recents, s'han relacionat una sèrie de funcions biològiques amb els STRs; com per exemple, la importància en l'organització de la cromatina o, per altre banda, també podrien estar involucrats en l'estructura de determinats RNA o en l'activitat de gens concrets [35].

Els STRs són essencials en el camp de la Genètica Forense, però també són útils en altres camps com és en la Genètica Clínica en el diagnòstic de determinades malalties. Els STRs estan implicats en un grup de patologies on la causa primària és una expansió excessiva de les unitats de repetició, sobretot de triplets. Algunes d'aquestes malalties són: el síndrome de X fràgil, l'atròfia muscular espinobulbar, la distròfia miotònica, la malaltia de Huntington i l'atàxia espinocerebral de tipus I [31].

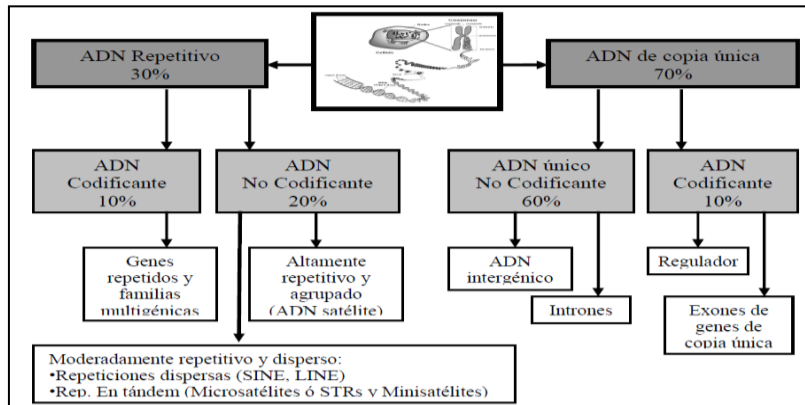


Figura 1. Organització del genoma humà [39]

1.1.1- Estructura dels STRs

Els STRs es caracteritzen per presentar un nucli de repetició (*repeat*) que comprèn entre dos i sis parells de bases, fragment que es va repetint un nombre diferent de vegades. En el camp de la Genètica Forense s'utilitzen majoritàriament els que tenen repeticions de quatre parells de bases amb al·lels que oscil·len entre els 100 i 300 parells de bases per la seva facilitat d'amplificació per la Reacció en Cadena de la Polimerasa (PCR). Finalment, aquestes unitats repetides es troben flanquejades per unes seqüències molt conservades [19, 21, 30].

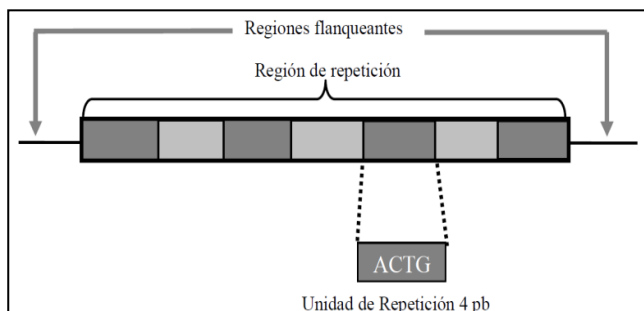


Figura 2. Estructura dels STRs [39]

1.1.2- Classificació dels STRs

Els STRs es poden classificar seguint dos criteris; el primer depèn del nombre de nucleòtids que formen part de nucli de repetició, així trobem els mono, di, tri, tetra, penta i hexanucleòtids (**Figura 3**). El segon criteri de classificació és en funció de la manera en que es van distribuint les unitats de repetició; així podem tenir STR simples, composts, complexos o hipervariables (**Figura 4**). En el cas dels STR simples, es va repetint una sèrie de vegades un mateix fragment amb absència de variació

en la seqüència, en canvi, els STR composts presenten dues o més repeticions simples. Els complexos alternen seqüències fixes a tots els individus amb seqüències de repetició variable de diferent longitud. Finalment, les repeticions hipervariables es caracteritzen per nombrosos al·lels que difereixen entre si tant en la longitud, com en la seqüència [3,38].

- **D**inucleotide (CA)(CA)(CA)(CA)
- **T**rinucleotide (GCC)(GCC)(GCC)
- **T**etranucleotide (AATG)(AATG)(AATG)
- **P**entanucleotide (AGAAA)(AGAAA)
- **H**exanucleotide (AGTACA)(AGTACA)

Figura 3. Classificació dels STRs en funció del nombre de nucleòtids presents al nucli de repetició [7]

	Locus	Secuencia	Alelos incompletos
STRs simples. Una sola UR y sin variaciones de secuencia	HUMFES/FPE	(ATTT) _n	
	HUMTOX	(AGAT) _n	
	D5S818	(GATA) _n	
	D13S317		
STRs simples. Con UR incompletas (con algunas UR de menos de 4 bases)	HUMTH01	(TCAT) _n	9.3
	HUMF13A1	(AAAG) _n	3.2
	HUMCSF1PO	(AGAT) _n	10.3
	D7S820	(GATA) _n	6.3
STRs compuestos. Con dos o más UR de distinta secuencia	HUMGABRB15	(GATA) _n (GATC) _n (TATC) _n	
STRs compuestos. Con UR incompletas	HUMVWA	(ATCT) _n (GTCT) _n (ATCT) _n	15.2
	D3S1358	TCTA (TCTG) ₁₋₃ (TCTA) _n	15.2
STRs complejos	HUMFGA D21S11	(TTTC) _a TTTT TTCT (CTTT) _n CTCC (TTCC) ₂	
STRs hipervariables	HUMACTBP2 D11S554	(TAGA) ₄ (TGA) ₀₋₁ (TAGA) ₆₋₁₆ (TAGG) ₀₋₁ (TG) ₅	

Figura 4. Classificació dels STR en funció del patró de repetició [modificada 29]

Finalment, hi són presents els al·lels incomplets, és a dir, al·lels que tenen una sèrie de repeticions completes, però, a més, presenten una repetició incompleta, ja sigui per una inserció o per una deleció. D'aquesta manera podem trobar, per exemple, l'al·lel 9.3 del marcador TH01, en el qual hi ha nou repeticions completes de 4 nucleòtids cada una i una repetició incompleta, que en aquest cas seria de 3 nucleòtids [21].

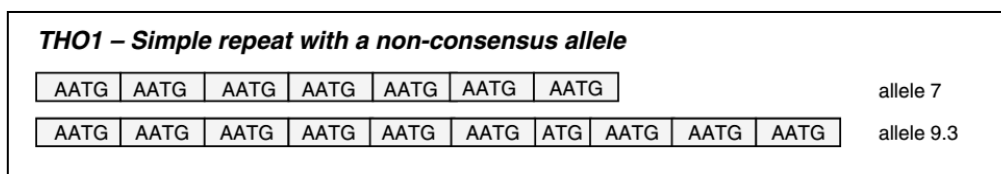


Figura 5. Representació d'un al·lel incomplet (9,3) en el marcador TH01 [21]

1.1.3- Microsatèl·lits i mutació

Aquestes regions solen presentar una elevada taxa de mutació a causa del fenomen anomenat *Polymerase Slippage* o lliscament de la polimerasa, característic de les zones de DNA repetitiu, resultat per la qual són tan polimòrfiques i tan variables entre els individus (**Figura 6**). En aquests

casos, durant la replicació, la DNA polimerasa produeix un lliscament sobre la cadena motlle que com a resultat es forma una bombolla a la nova cadena. El mecanisme de reparació elimina aquesta bombolla, provocant insercions o delecions d'una petita seqüència, fent que augmentin o disminueixin el nombre de repeticions d'un al·lel en concret [29].

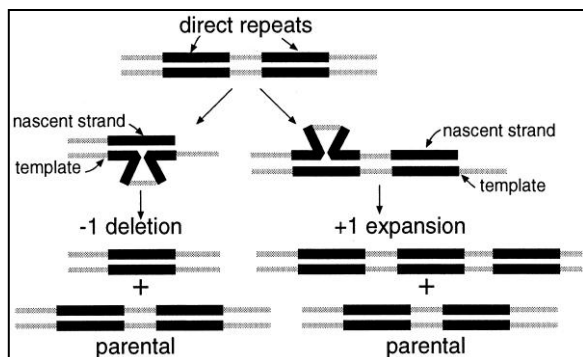


Figura 6. Conseqüències del lliscament de la polimerasa; disminució del nombre de repeticions (-1) o augment (+1) [11]

1.2- STRs en la Genètica Forense

Un dels objectius de la Genètica Forense és la resolució de paternitats disputades o altres casos de relacions de parentiu entre dos individus. En aquest cas el procés que es segueix és comparar el perfil al·lèlic d'una persona, és a dir, el conjunt d'al·lels pels diferents STRs estudiats d'un individu, amb el perfil al·lèlic d'un altre individu i establir un parentiu entre aquests dos (**Taula 1**). D'aquesta manera, primer es duu a terme la identificació dels individus per separat i, posteriorment, s'estableix una relació entre ells.

Taula 1. Perfil al·lèlic dels tres individus implicats en una paternitat

Sistema/ Mostra	Pare	Fill	Mare	Sistema/Mostra	Pare	Fill	Mare
D8S1179	12-13	13-13	13	D19S433	12-13	13-14	12-14
D21S11	30-	30-	28-30	vWA	18	15-18	15-19
	32.2	32.2		TPOX	8	8-8	8-11
D7S820	10	9-10	9-12	D18S51	13-17	14-17	13-14
CSF1PO	11-12	11-12	11	Amelogenina	XY	XY	XX
D3S1358	17-18	15-17	14-15	D5S818	11-12	11-11	11-12
THO1	6-9	9-9.3	7-9.3	FGA	22-23	21-22	20-21
D13S317	8-11	8-13	11-13	Penta E	11-13	11-13	11-12
D16S539	11-13	11-12	12-13	Penta D	11-12	11-12	12-13
D2S1338	18-24	18-19	18-19				

En Genètica Forense s'utilitzen majoritàriament els STRs envers dels minisatèl·lits, ja que aquests darrers presenten una limitació important; es requereix una gran quantitat de DNA d'elevat pes molecular que dificulta el seu estudi. De manera contrària, els STRs compleixen una sèrie de requisits que els fan idonis per aquests tipus d'estudis. Una de les característiques importants que permet que siguin utilitzats com a marcadors moleculars és que presenten una herència codominant, és a dir, són manifestos els dos al·lels: el patern i el matern [25]. A més, els STRs impliquen una amplificació òptima

i una ubicació cromosòmica concreta. Referent a la taxa de mutació, ha de ser relativament baixa com perquè es mantinguin el nombre de repeticions de pares a fills, però també ha de ser relativament alta, per tal de que siguin altament polimòrfics.

D'aquesta manera, els STRs presenten un elevat poder de discriminació, és a dir, permeten diferenciar a un individu entre un conjunt, gràcies a la seva variació. Aquest poder de discriminació augmenta quan es fa simultàniament l'estudi d'una sèrie de *loci*. Habitualment es duu a terme la realització de *PCR Multiplex* combinant marcadors amb diferent longitud i marcadors amb diferents fluorocroms. Aquestes reaccions se fan a partir de molt poca quantitat de mostra, fet que facilita l'estudi, ja que en molts de casos forenses la quantitat de mostra és mínima. Els STR emprats en un determinat estudi solen ser STRs amb un elevat nivell de variació intrapoblacional, fet que permet augmentar el poder de discriminació [39,40]. És important que els *kits* emprats incloguin els 13 STRs següents: CSF1PO, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, FGA, THO1, TPOX, vWA, ja que són els inclosos al CODIS (*Combined DNA Index System*), que és el terme utilitzat per descriure el programa de suport a les bases de dades de DNA de justícia criminal del FBI, i que s'ha adoptat com a sistema mínim consens a nivell de la comunitat internacional.

1.3- Interpretació dels resultats i artefactes dels STRs

Durant el procés de genotipació dels STRs d'un individu per la elaboració del seu perfil, es poden produir un conjunt d'artefactes que poden dificultar la interpretació dels resultats de la prova i, per tant, poden alterar la correcta tipificació dels STRs corresponents. Alguns d'aquests artefactes es donen durant la Reacció en Cadena de la Polimerasa (PCR) i altres, poden correspondre a característiques biològiques peculiars de la mostra. El perfil de cada artefacte és diferent i l'investigador ha de saber identificar-ho, ja que de manera contrària, es pot donar un resultat erroni. Els artefactes més comuns són els productes *stutter*, l'al·lel *drop-out*, els al·lells *off-ladder*, els pics *pull-up*, l'al·lel *drop in*, l'adenilació incompleta i el desequilibri de pics (entre al·lells o entre locus)[5, 7].

a-Els ***stutters*** o bandes de repeticició són els artefactes més freqüents. S'originen a causa d'un lliscament de la polimerasa durant l'amplificació del DNA, fet que genera un petit pic just abans del pic principal (**Figura 7**). Són amplicons amb una o més repeticions menys que l'al·lel original, però així i tot, la grandària del pic corresponent a l'*stutter* és habitualment molt més petita [20,5]. El nombre de pics *stutter* i la seva respectiva intensitat són inversament proporcionals a la longitud de les repeticions i directament proporcionals al nombre de repeticions presents a l'al·lel amplificat [36]. És a dir, unitats més petites de repeticició (di- o tri) generen una major quantitat d'*stutters* i de major intensitat, i dins un mateix sistema, a major nombre de repeticions, major és la intensitat del pic. A més, el nombre d'*stutters* incrementa amb la complexitat dels STRs, així com també augmenta a mesura que augmenta la quantitat A-T [2], tot i que normalment representen menys del 15% del pic principal [21].

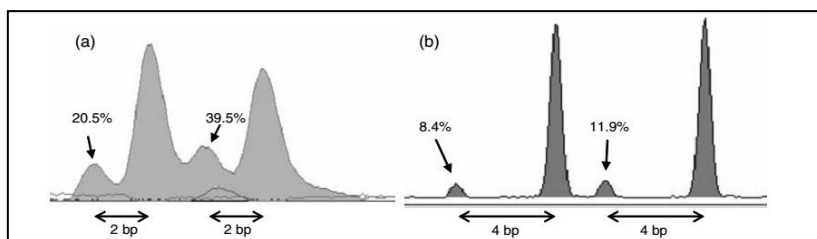


Figura 7 . Productes *stutter* durant l'amplificació del DNA. a) Repetició de dinucleòtids amb el pic del *stutter* que representa un percentatge elevat i b) Repetició de tetranucleòtids amb menor percentatge de *stutter* [21]

b-El **drop-out** al·lèlic és un artefacte on no s'amplifica l'al·lel (**Figura 8.a**), ja sigui per una mutació al lloc d'unió al *primer* (al·lels nuls) (**Figura 8.b**), per manca o degradació del DNA o per alguna condició de la PCR [6].

El fet d'utilitzar diferents *kits* implica utilitzar diferents *primers* pel mateix STR, per tant, si un dels *primers* no s'ha pogut unir a causa d'una mutació, l'altre sí que ho podrà fer i, per tant, si que hi haurà ampliació i ens permetrà detectar els *drop-out* deguts a l'existència d'un al·lel nul (**Figura 9**) [5].

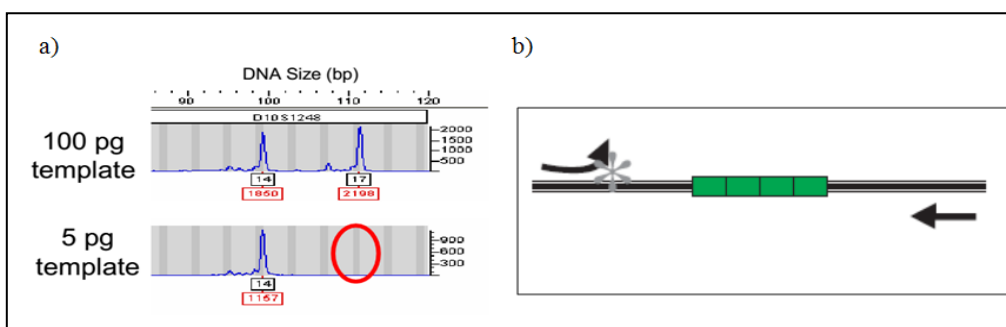


Figura 8 .a) Absència de pic a l'electroferograma a conseqüència de *drop-out* [7] b) Mutació al lloc d'unió del *primer* que impedeix la seva unió al DNA [6]

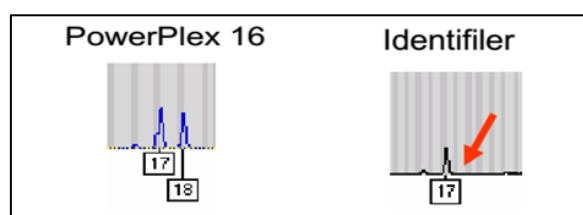


Figura 9 . La utilització de dos *kits* diferents permet detectar els *drop-out*, ja que s'utilitzen *primers* diferents [12]

c-Els **pull up** són pics també coneguts com a contaminació de color causats per un excés de mostra. Quan la intensitat de la senyal és molt elevada no es pot mesurar el vertader valor de la senyal i, per tal de compensar-ho, es genera un solapament espectral de colors. Com a resultat, es generen uns pics que solen estar presents en la mateixa posició, però amb un color diferent [15].

Durant l'anàlisi de la fluorescència, hi ha un solapament de les longitud d'ones de les diferents sondes, fet que indueix la presència de dos pics de diferent color. Per tal de solucionar aquesta incapacitat del instrument per detectar adequadament el color, el que fa en la majoria dels casos és diluir la mostra [5].

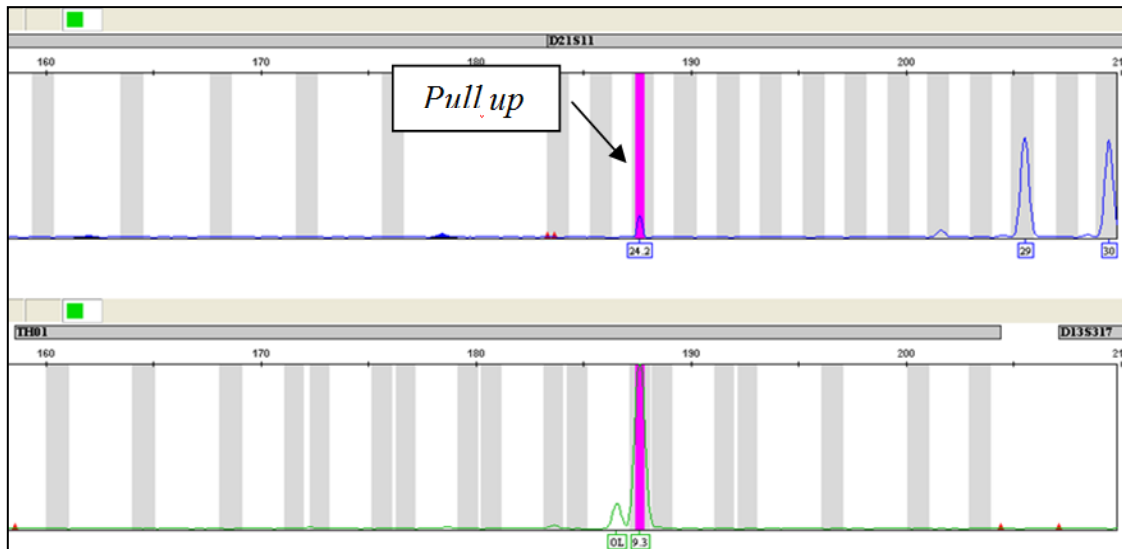


Figura 10. Representació d'un *pull up* originat per el marcador TH01

d-**Al·lels off-ladder**, és a dir, al·lels que no estan inclosos en el *ladder*. El *ladder* al·lèlic (o escaleta d'al·lels) és un marcador de DNA de referència, el qual permet identificar i assignar un valor a aquells al·lels més freqüents (**Figura 11**). Per tant, tots aquells al·lels que presenten un nombre de repeticions desconegut o lleugerament diferent són assignats com a *off-ladder*, és a dir, el programa no li assigna un valor (**Figura 12**) [5]. No tenen perquè ser sempre un artefacte del procés de genotipació, sinó al·lels nous, al·lels poc freqüents o variants dels al·lels comuns resultants d'una inserció, deleció o substitució que afecta al nombre de repeticions.

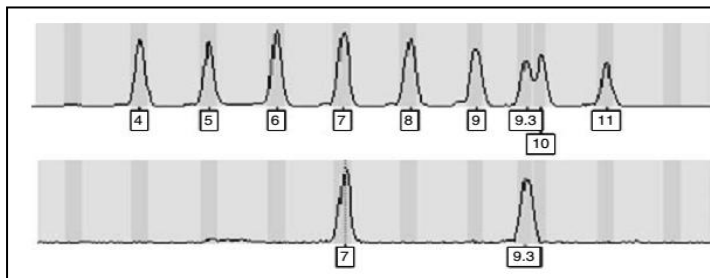


Figura 11. El *ladder* al·lèlic permet assignar els valors corresponents als pics de la mostra [21]

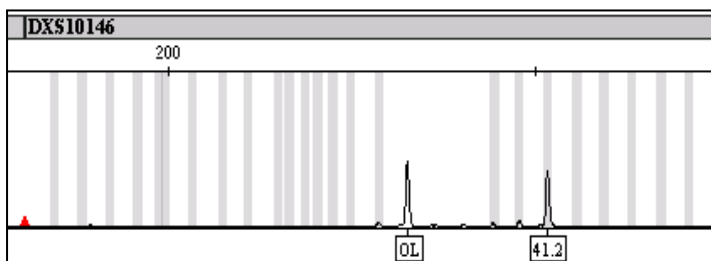


Figura 12. Representació d'un al·lel *off-ladder*

e-**Drop-in**. Es tracta de falsos al·lels que apareixen addicionalment a un perfil genètic com a conseqüència de contaminacions esporàdiques o per efectes estocàstics quan s'incrementen el número de cicles de la PCR per manca de mostra [9].

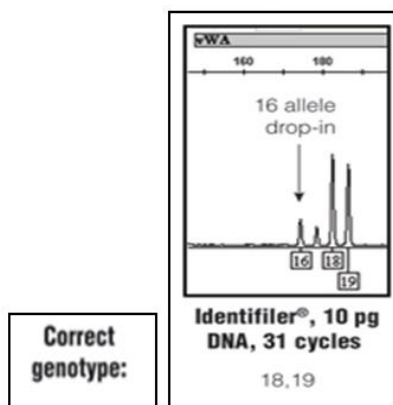


Figura 13. Representació d'un al·lel *drop-in* [9]

f-*Allele imbalance i locus-to-locus imbalance*

L'*allele imbalance* o desequilibri entre al·lells és l'amplificació preferencial d'un dels al·lells en un individu heterozigot. Normalment el pic de l'al·lel menor és lleugerament superior al del major, ja que durant els cicles de PCR com més petit és l'amplicó més còpies se'n produeixen, però en condicions normals la diferència no supera el 40%. Si el desequilibri és molt sever, com el que s'observa en la **Figura 14**, és possible la genotipació errònia com a un homozigot. Pot deure's a una mutació al lloc d'unió al *primer* que disminueixi l'eficiència de l'amplificació o ser indicatiu d'una mescla en la mostra, en aquest cas es detectarà a diferents marcadors.

El desequilibri entre locus és l'amplificació preferencial dels marcadors amb amplicons més curts. Pot donar-se per degradació del DNA motlle o per excés de DNA a la reacció de PCR [9].

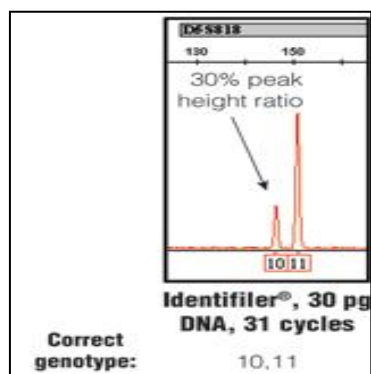


Figura 14. Desequilibri entre al·lells. L'al·lel 11 supera amb més d'un 60% l'alçada del pic corresponent a l'al·lel 10 [9]

h-Adenilació incompleta. La Taq polimerasa emprada per l'amplificació dels STRs catalitza l'addició d'un nucleòtid extra, usualment una adenina, a l'extrem 3' dels productes de PCR. Els *kits* comercials ja han estat optimitzats emprant *primers* que afavoreixen l'adenilació completa dels amplicons. No obstant, si hi ha un excés de DNA motlle o les condicions de la PCR no són òptimes (per exemple poc temps d'amplificació) es poden observar pics dividits (**Figura 15**) on l'al·lel estarà representat per dos pics que difereixen en 1 pb [7]. El *software* de genotipatge les pot analitzar con un *off-ladder* o microvariant. L'aparença d'aquests pics serà diferent segons la relació existent entre cadenes amb o sense adenilació completa (**Figura 16**) [21].

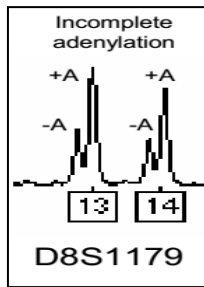


Figura 15. Representació de l'adenilació incompleta: cada al·lel presenta dos pics que difereixen de 1 pb [7]

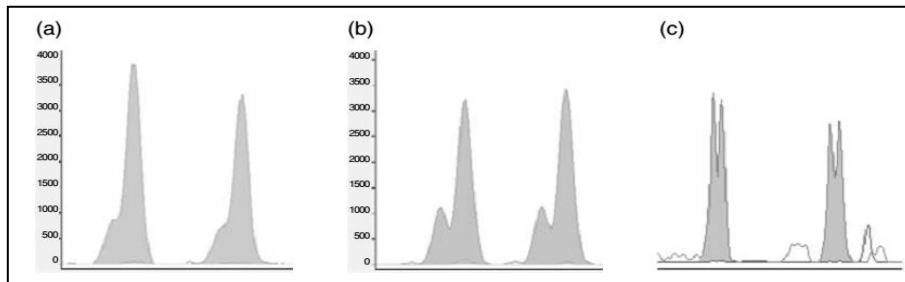


Figura 16. Diferents graus d'adenilació incompleta. En funció del percentatge d'amplicons que presenten adenilació incompleta, l'alçada dels pics variarà; a) la majoria dels amplicons per aquell pic presenta adenilació completa, generant un pic d'adenilació incompleta molt baix i gairebé solapat amb l'original i c) aproximadament la meitat dels amplicons presenten adenilació incompleta, resultant dos pics gairebé de la mateixa alçada [21]

i-Mutacions de novo. La mutació dels STR, anteriorment esmentada, pot donar problemes d'interpretació en l'anàlisi de relacions de parentiu quan es produeix a un gàmet parental.

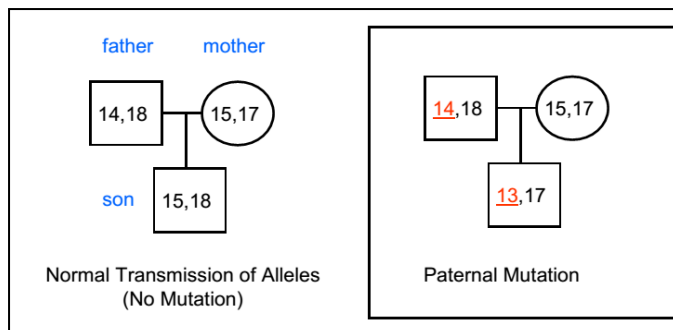


Figura 17. Mutació parental en un anàlisi de paternitat [5]

j- Patrons de triple pic. A vegades s'obtenen patrons trial·lèlics reproduïbles a la mostra en qüestió. A diferència del que ocorre a una mescla, aquest patró sols s'observa a un locus. Són conseqüència de la presència de fragments cromosòmics duplicats o mutacions que dupliquen la seqüència d'unió de l'encebador. S'han descrit més de 175 patrons trial·lèlics diferents en els 13 STRs del CODIS. El marcadors on més se n'han descrit són els següents: TPOX, D18S51, D21S11, VWA i FGA [7]. La llista es pot trobar a la NIST seguint el següent enllaç: (http://www.cstl.nist.gov/biotech/strbase/tri_tab.htm) [9].

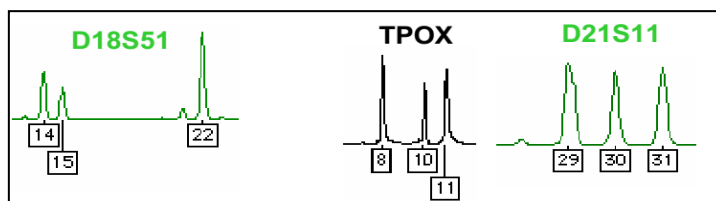


Figura 18. Exemples de marcadors que poden presentar triple pic (D18S51, TPOX i D21S11) [12]

k-Artefactes tècnics: renou, bombolles, *spikes* i contaminants

Bombolles. Alguns fluorocroms poden separar-se de l'encebador i migrar independentment a través del capil·lar durant l'electroforesi. L'agrupació d'aquests fluorocroms pot ocasionar un pic anòmal a l'electroferograma. La majoria d'aquests artefactes tendeixen a aparèixer com un pic molt ample (**Figura 19**). El perill és que poden emascarar un al·lel veritable [16].

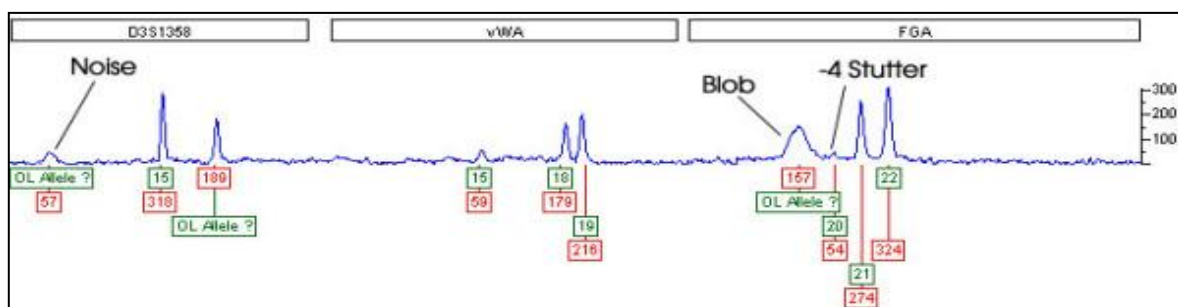


Figura 19. Representació d'artefactes tècnics (renou i bombolles) [16]

El **renou** és present en totes les mostres i s'observa com petits pics al llarg de la línia de base. A vegades, el renou pot ser un problema si és massa gran, ja que pot emascarar pics al·lèlics reals. Les bombolles d'aire, vidres d'urea i contaminació de la mostra es troben entre les possibles causes.

Els ***spikes*** són artefactes tècnics causats per pics de voltatge durant l'anàlisi. Solen ser pics de forma inusualment prima. També les bombolles d'aire o els cristalls d'urea poden donar falsos pics. Igual que amb tots els artefactes tècnics, el perill és l'impacte potencial sobre els resultats de les proves de DNA [16].

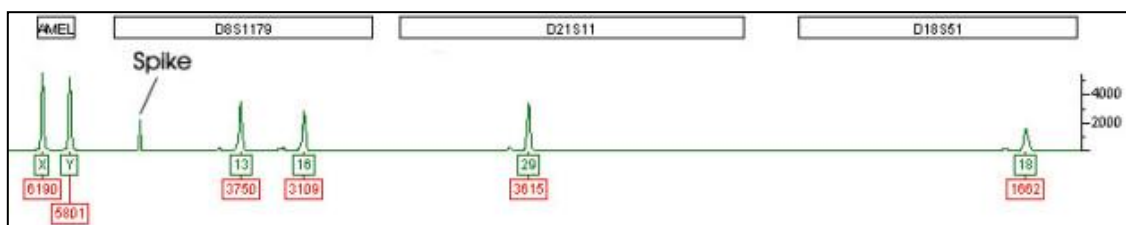


Figura 20. *Spike* present en el marcador D8S1179 [16]

Contaminants. Materials presents a la mostra que presentin fluorescència en la regió visible de l'espectre, poden interferir quan s'empra el sistema de detecció de fluorescència per a la tipificació del DNA. S'han descrits com a contaminants d'aquest tipus; antibiòtics, vitamines, aromàtics policíclics, colorants tèxtils, etc . Normalment són eliminats si es fa una extracció orgànica de DNA (per exemple

amb fenol/cloroform). A més, els pics que donen solen ser molt amplis i, per tant, es diferencien bé del productes de PCR marcats amb fluorocroms.

2.MATERIALS I MÈTODES

2.1- Procediment general de les proves de paternitat

La sol·licitud per una prova de paternitat a la UIB es pot fer per via judicial o directament posar-se en contacte amb el laboratori de Genètica de la Universitat de les Illes Balears. Posteriorment, es concreta una cita, on es recullen les mostres. La recollida de mostres és molt senzilla i no invasiva; s'extreu DNA a partir d'un frotis bucal emprant uns escovillons de cotó, on s'obtenen cèl·lules de la mucosa bucal o saliva. En tot moment s'utilitzen dues mostres de cada individu per tal de tenir un duplicat.

A continuació, s'inicia l'anàlisi del DNA obtingut. En primer lloc, es segueix el protocol d'extracció del DNA amb una quantificació final de mostra per tal d'optimitzar la PCR. Es realitzen dues PCR amb dos kits diferents; *kit AmpF/STR Identifiler* (Applied Biosystems) i el kit *Powerplex® system* (Promega). En total hi ha 17 marcadors tipificats per l'estudi: (D19S433, D2S1338, D21S11, D3S1358, FGA, D8S1179, D18S51, CSF1PO, TPOX, THO1, vWA, D16S539, D7S820, D13S317, D5S818, Penta E i Penta D). El propòsit d'emprar dos kits que dupliquen 13 dels 17 STRs és la comprovació de que l'amplificació s'ha fet correctament i, a més, permet detectar possibles artefactes.

En les dues PCR mencionades, s'utilitzen parelles de *primers* que es fonamenten amb dos oligonucleòtids de 15-30 nucleòtids (*forward i revers*) que hibriden amb la respectiva cadena complementària. Es poden marcar amb fluorescència ambdós *primers* o només un dels dos. Els *primers* permeten l'amplificació de la polimerasa, generant fragments de DNA de diferent longitud en funció del nombre de repeticions que presenta el STR. Finalment, es duu a l'analitzador genètic, on mitjançant una electroforesi capil·lar es separen els diferents fragments marcats en funció de la longitud. Un detector de fluorescència permet elaborar un electroferograma final on s'observen els pics corresponents a les diferents molècules detectades segons la longitud i el marcatge amb diferents fluorocroms (**Figura 21**) [15].

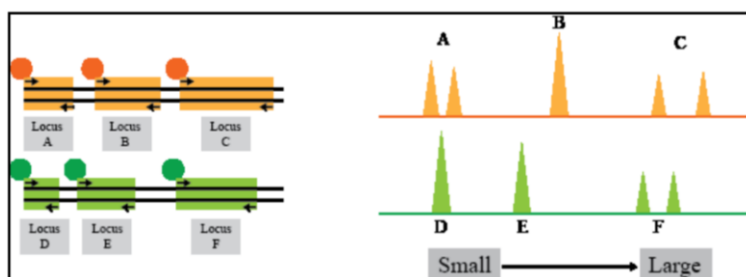


Figura 21. Separació dels fragments marcats en funció de la longitud i el marcatge que presenten [15]

Una vegada arriben els resultats de l'analitzador genètic, s'examinen amb el **Software GeneMapper®**; programa que permet l'anàlisi dels microsatèl·lits per extreure el perfil al·lèlic de cada individu. Finalment, en cas de que la paternitat sigui compatible, el darrer programa que s'utilitza és el

Famílias que ens permet calcular la probabilitat de parentiu entre diferents persones a partir de seu perfil al·lèlic.

2.2- Kits emprats

Com ja s'havia esmentat anteriorment, la distribució dels STRs és aparentment atzarosa i tant els podem trobar en els cromosomes autosòmics com en els cromosomes sexuals (X i Y) [23,34], així com també al mtDNA [22,40]. S'han descrit milers de microsatèl·lits en el genoma humà, suposant aproximadament el 3% de la totalitat del genoma. De mitjana es troba un STR diferent cada 10.000 nucleòtids [8].

Els més utilitzats actualment són els autosòmics, encara que en alguns casos on és necessari veure exclusivament l'herència paterna s'utilitzen els STRs del cromosoma Y, mitjançant el kit AmpFLSTR® Yfiler® PCR Amplification Kit. En altres casos, quan les relacions entre parentius són complexes, és interessant analitzar els microsatèl·lits del cromosoma X a través del kit Investigator Argus X-12. En casos d'interès en l'anàlisi de l'herència materna s'empra el mtDNA, tot i que en aquest cas s'empra preferentment la seqüència de regions hipervariables.

Per tal d'analitzar els STRs autosòmics s'utilitzen majoritàriament dos kits de les cases comercial *Identifiler* i *Promega*; el *AmpF/STR Identifiler Plus PCR Amplification Kit* (Applied Biosystems) i el kit *Powerplex® 16 system* (Promega).

El kit *AmpF/STR Identifiler* (Applied Biosystems) permet genotipar 15 STRs; 13 dels quals es comparteixen amb el *PowerPlex® system* (Promega) i els dos restants són específics per aquest kit [25]. De la mateixa manera, el *PowerPlex® system* (Promega) permet identificar 15 STRs; 13 serviran per fer de duplicat, en el cas de ser emprat simultàniament amb l'anterior, i els altres dos (penta D i penta E) són específics (**Figura 22**) [26]. Cal mencionar que cada un dels kits utilitza uns encebadors diferents pels 13 STRs comuns i, d'aquesta manera, si existeix alguna mutació en el lloc d'unió del primer al DNA, tindrem la seguretat que amb un kit o amb l'altre hi haurà amplificació. En total s'analitzen 17 STRs diferents i, a més, s'inclou també una regió del gen de l'amelogenina, que permet identificar el sexe de l'individu (**Figura 23**) [25].

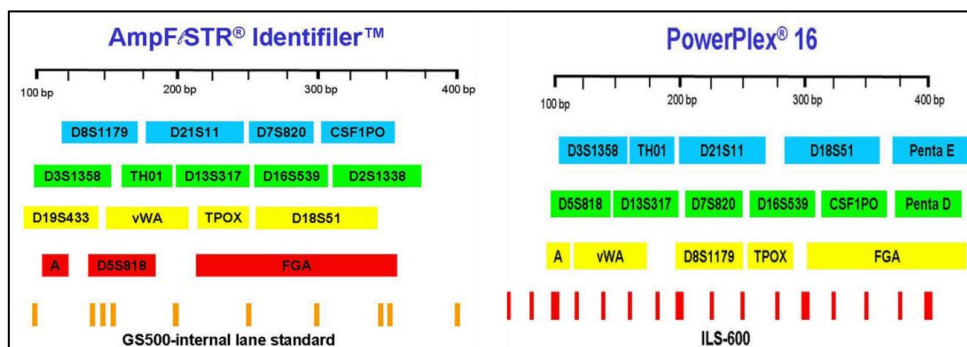


Figura 22. Diagrames dels STRs amplificats pel kits *AmpF/STR Identifiler Plus* [10]

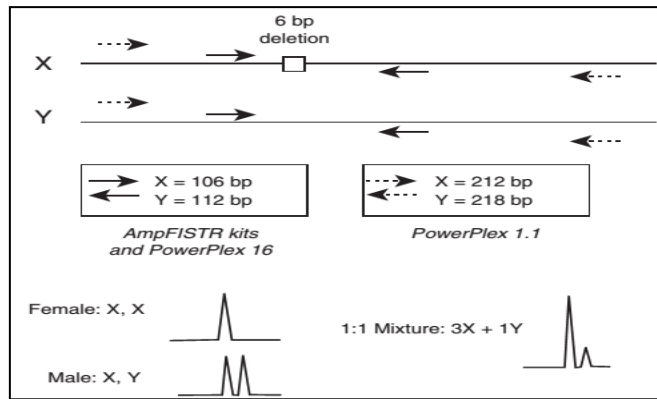


Figura 23. Determinació del sexe mitjançant l'amelogenina [6]

Actualment, s'han desenvolupat nous kits que inclouen més marcadors, per exemple el *Powerplex 21®System* o el *GlobalFiler® PCR Amplification Kit* que inclou 24 STRs. També s'han creat kits per a nous marcadors, derivats dels actuals, anomenats miniSTRs (Figura 24). En realitat són els mateixos STRs, però el protocol d'amplificació es fa amb uns primers que redueixen la zona flanquejant amplificada. L'objectiu d'aquests petits STRs és facilitar l'estudi en aquells casos on el DNA està lleugerament degradat, com per exemple en casos de DNA antic o catàstrofes aèries i naturals on la mostra ha estat sotmesa a condicions adverses per a la conservació del DNA [14]. Gràcies a una sèrie d'estudis, s'han generat protocols i kits per tal de fer que aquests miniSTR siguin útils i eficaços per a l'aplicació a la rutina de la Genètica Forense i, a hores d'ara, n'hi ha 26 de descrits [10].

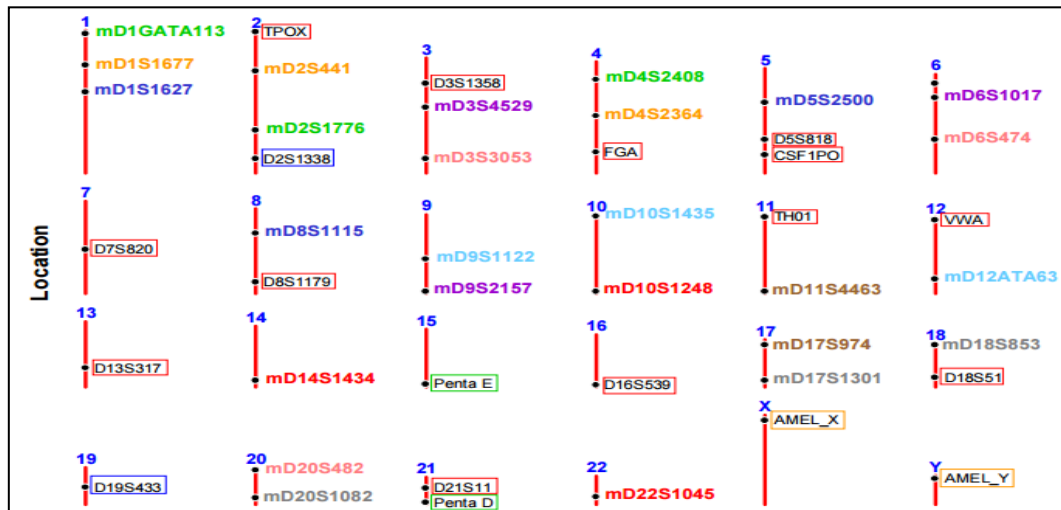


Figura 24. Posició dels 26 miniSTR en el genoma [10]

2.3- Anàlisi de fragments (programa GeneMapper)

El GeneMapper®Software és un programa d'anàlisi de fragments de DNA, que permet l'estudi de AFLP (amplificació de fragments polimòrfics), LOH (pèrdua d'heterozigocitat), microsatèl·lits i, finalment, el genotipat de SNP. En el cas de les proves de paternitat, només l'utilitzem per l'anàlisi dels STRs.

El programa es fonamenta en extrapolar un valor als pics, que correspon al nombre de repeticions per cada STRs basant-se amb els valors corresponents del *ladder* al·lèlic (conjunt d'al·lells de referència d'un determinat marcador) (**Figura 25**) i el marcador de pes molecular (**Figura 26**). D'aquesta manera, el primer que s'ha de fer és comprovar que el marcador ha corregut bé, ja que sinó el programa assigna erròniament els valors a cada un dels pics [15].

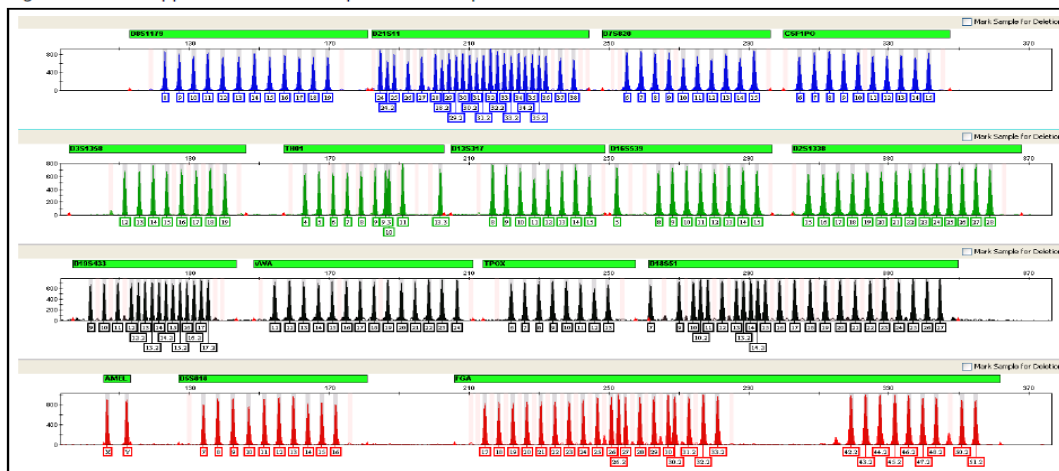


Figura 25. Ladder al·lèlic del kit *AmpF/STR Identifier* [25]

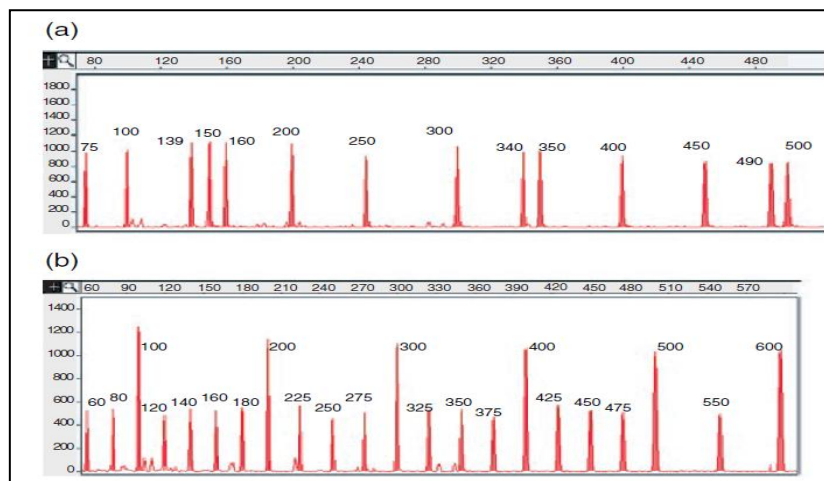


Figura 26. a) Marcadors de pes molecular emprant en els kits *AmpF/STR Identifier* i b) *Powerplex® system* [21]

Per defecte, el programa detecta altres pics que no corresponen amb els al·lells i, per tant, s'han d'analitzar un per un i eliminar aquells que s'han assignat erròniament (**Figura 27 i 28**). A més, s'han de contemplar tots els possibles artefactes i detectar-los, sobretot anar en compte amb els *stutter*, ja que són els artefactes més abundants i assegurar-se de que no interfereixen en la correcta tipificació dels STRs estudiats.

En cada un dels marcadors que s'estudien hi podem trobar un o dos al·lells; en funció de si l'individu és homocigot o heterocigot, respectivament (**Figura 29**). Una vegada anotats tots els al·lells implicats, generant el que es coneix com a perfil al·lèlic, s'ha de determinar si una paternitat és compatible o, en canvi, és una exclusió. Quan es considera exclusió, és hàbit mèdic-legal precisar, al menys, de tres sistemes incompatibles per negar la paternitat biològica (**Taula 2**). En el cas contrari, quan una

paternitat és compatible implica que el fill hereta un dels al·lels de la mare i un altre del pare i, per tant, podem continuar amb el programa *Famílias*.

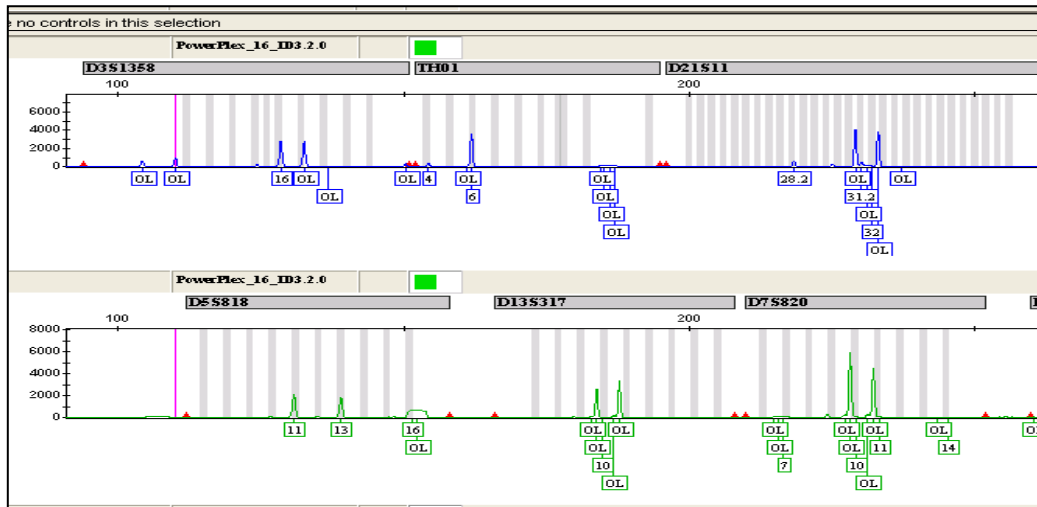


Figura 27. Electroferograma obtingut després de l'analitzador genètic

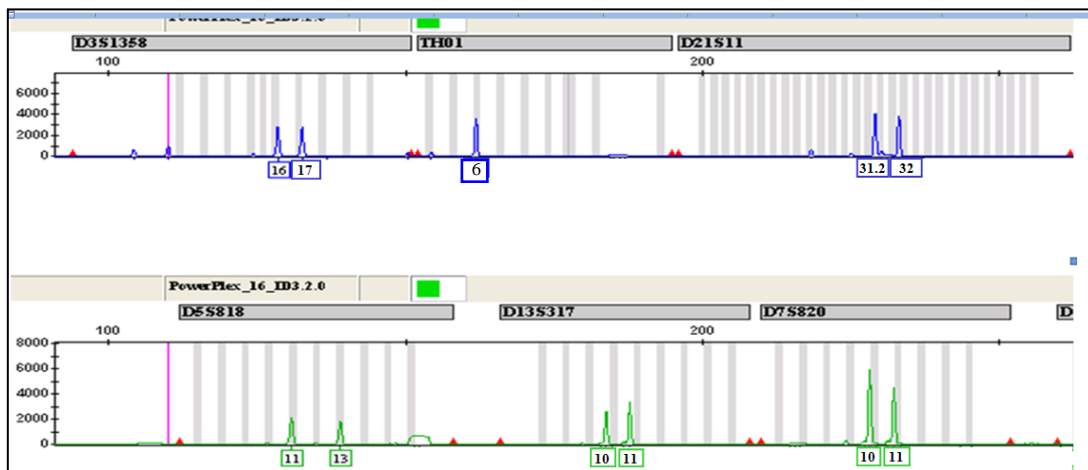


Figura 28. Electroferograma abans d'anotar el perfil al·lelic

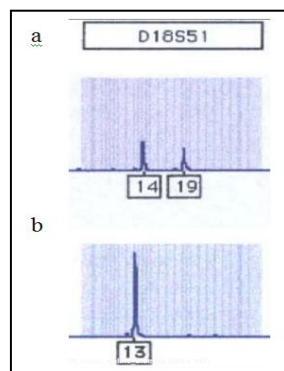


Figura 29. Individu heterozigot (a) i homozigot (b) per el marcador D18S51

Taula 2. Cas d'exclusió amb 5 marcadors incompatibles (marcats en gris)

Sistema/ Mostra	Pare	Fill	Mare	Sistema/ mostra	Pare	Fill	Mare
D8S1179	12-13	13-14	13	D19S433	12-13	11-14	12-14
D21S11	30- 32.2	30- 32.2	28-30	vWA	18	15-18	15-19
D7S820	10	9-11	9-12	TPOX	8	8-8	8-11
CSF1PO	11-12	11-12	11	D18S51	13-17	14-17	13-14
D3S1358	17-18	15-19	14-15	Amelogenina	XY	XY	XX
THO1	6-9	9-9.3	7-9.3	D5S818	11-12	11-11	11-12
D13S317	8-11	8-13	11-13	FGA	22-23	21-22	20-21
D16S539	11-13	11-12	12-13	Penta E	11-13	11-14	11-12
D2S1338	18-24	18-19	18-19	Penta D	11-12	11-12	12-13

2.4- Anàlisi estadístiques dels resultats d'una prova de paternitat (programa *Familias*)

En tots aquells casos on no hi hagi exclusió utilitzarem el *Familias*; programa que ens permet calcular les probabilitats de parentiu entre els diferents individus analitzats. El més comú és l'anàlisi de paternitats disputades, on no es dubte de la relació mare-fill i el que es calcula és la probabilitat de que el suposat pare sigui en realitat el pare biològic. En aquest cas, a partir dels perfils al·lèlics obtinguts amb el *GeneMapper* del fill, de la mare i del suposat pare i coneixent les freqüències per a cada un dels al·lèls a la població de referència, es pot calcular l'índex de paternitat (IP) i la probabilitat acumulada de paternitat (W) (veure apartat 2.4.3).

Una vegada introduïts tots els valors dels perfils al·lèlics i exposades les relacions conegudes i les que es volen conèixer, el programa genera un informe final amb tots els resultats; índex de paternitat parcial per a cada un dels marcadors (IP parcial), IP total i la W (**Figura 30**).

```

| Parent          Child
| -----
| Pare           Fill
|
| Prior probability: 0.5
| Posterior probability: 0.9999966926 Probabilitat Acumulada de Paternitat
| Ln likelihood: -93.49994587
| Posterior ratio versus Ped1: 302348.4104
| Likelihood ratio versus Ped1: 302348.4104 Índex de Paternitat
|
| System          Likelihood          LR versus Ped1
| -----
| CSF1PO          0.0285563          1.701258932
| D2S1338          0.000387282        6.93481276 Índex de Paternitat parcial
| D3S1358          0.0152703          1.754385965
| D5S818          0.0400017          2.923976608
| D7S820          0.00536047         2.136752137
| D8S1179          0.0283726          3.278688525
| D13S317          0.00931399         0.8389261745
| D16S539          0.0176357          0.8500510031
| D18S51          0.00253369         2.016129032
| D19S433          0.00300734         2.144082333
| THO1            0.00767141         1.287332647
| TPOX            0.0279242          0.9914733294
| vWA             0.00494019         0.9259259259
| D21S11          0.00011385         10
| FGA             5.12e-005          3.125
| Penta e         0.000837577        6.393861893
| Penta d         0.00372426         0.9796238245

```

Figura 30. Exemple de resultat final del programa *Familias* en l'anàlisi d'una paternitat

2.4.1- Bases de dades

Les bases de dades fan un recull de les freqüències i seqüències de tots els al·lells descrits fins el moment i tots aquells que van aparegut de nou per cada un dels STRs que s'utilitzen com a marcadors. Una de les bases de dades de referència *on line* és el "The National Institute of Standards and Technology (NIST)" que ha recopilat i conservat tota la informació obtinguda sobre els STRs des del 1997 [10]. Aquesta base de dades és un recurs per els investigadors en Genètica Forense ja que hi trobem tots els detalls dels marcadors que s'utilitzen en l'estudi dels DNA, així com totes les novetats sobre els STRs en el camp de la Genètica Forense [33].

Apart d'aquesta base de dades *online* és habitual que es tinguin bases de dades nacionals o locals on es recullin les freqüències al·lèliques a poblacions concretes, ja que és molt important aplicar en cada cas les freqüències adients segons el marc poblacional on s'ha produït el cas a analitzar.

2.4.1.1- Freqüències al·lèliques

Les freqüències al·lèliques, que es troben en les bases de dades, fan referència al nombre de vegades que apareix un al·lel en un determinada població dividit per el total d'al·lells d'aquella població [6]. Per tant, les freqüències al·lèliques són el resultat de multitud d'estudis per a cada un dels marcadors realitzats a molts individus d'una població/ètnia.

És important emprar la base de dades adient, sobretot si el cas analitzat es dona en una regió geogràfica amb subestructuració poblacional, per exemple una regió on conviuen diferents grups ètnics o on hi ha poblacions aïllades.

2.4.1.2- Freqüència al·lèlica mínima

En qualque cas es troba la transmissió d'un al·lel amb freqüència molt baixa o, fins i tot, no descrit anteriorment a la base de dades de la població que s'empra. A l'hora de realitzar el càlcul de probabilitat, la freqüència d'aquest al·lel tot sol donaria una probabilitat de paternitat altíssima i tal vegada inexacte. Per minimitzar l'efecte d'aquest al·lel, s'empra per fer els càlculs de probabilitat, el que es coneix com la freqüència al·lèlica mínima; valor conservatiu basat en la idea de que cada al·lel s'hauria d'haver observat un mínim de 5 vegades a la base de dades per ser inclòs en càlculs estadístics fiables.

Segons *The National Research Council* (1996), si un al·lel es manifesta una o dues vegades en una determinada població, la freqüència al·lèlica calculada seria inexacte i, per aquest motiu, no apareixen en les bases de dades. Només els al·lells que es manifesten com a mínim cinc vegades en una determinada població poden estar inclosos en l'estadística. A partir d'aquest criteri, es va crear una fórmula per tal de calcular la freqüència al·lèlica mínima; $5/2N$, on N fa referència al nombre d'individus estudiats en una determinada població i el 2 representa les dues còpies de cada cromosoma. És per aquest motiu, que cada base de dades té un valor mínim concret per cada marcador, en funció del nombre d'individus estudiats [6].

Per exemple, si les mostres analitzades pel marcador D16S539 a la nostra base és de 600, la freqüència al·lèlica mínima és 0,0042. Si a una paternitat trobem un al·lel molt rar, que presenta una

frequència, per exemple, de 0,0021 (menor de 0,0042), doncs el que hem de fer a l'hora de calcular la probabilitat de paternitat és emprar 0,0042 enlloc del 0,0021.

2.4.2- Índex de Paternitat i Probabilitat Acumulada de Paternitat

El conjunt de STRs estudiats presenten una probabilitat d'exclusió del 99,999999%, és a dir, exclourien en un 99,9999% dels casos a un individu que no fos el pare biològic. En cas de que no es doni una exclusió, s'ha de calcular la probabilitat d'assignació positiva, mitjançant dos paràmetres; l'índex de paternitat (IP) i la probabilitat acumulada de paternitat (W). Ambdós es calculen a partir dels mètodes Bayesianes clàssics, partint de que la probabilitat *a priori* de que el presumpte pare ho sigui és del 50%.

L'índex de Paternitat ($IP = X/Y$) és una relació que s'obté comparant dues hipòtesis. Indica, tenint en compte els resultats obtinguts en les proves, quantes vegades és més probable que el fill sigui un descendent del presumpte pare i de la mare (X), comparat amb la possibilitat de que el fill sigui descendent d'un individu a l'atzar de la població (Y). Aquesta relació es calcula per a cada un dels marcadors (IP parcial) i posteriorment s'obté l'IP total.

La Probabilitat de Paternitat es calcularia segons la fórmula $W = X/(X+Y)$ o $W = IP/IP+1$. Per facilitar la interpretació dels valors de l'índex de paternitat al context judicial s'han establert una sèrie d'equivalències o predicats verbals, com per exemple el que s'exposen a la **Figura 31** [21].

Likelihood ratios	Verbal equivalent
1-10	Limited support for prosecution hypothesis
10-100	Moderate support for prosecution hypothesis
100-1000	Moderately strong support for prosecution hypothesis
1000-10 000	Strong support for the prosecution hypothesis
>10 000	Very strong support for prosecution hypothesis

Figura 31. Interpretació del índex de paternitat (IP) [21]

3.OBJECTIUS

L'objectiu general d'aquest treball era analitzar de forma global i anònima els resultats de 25 paternitats realitzades al laboratori de Genètica de la UIB en el darrer any.

Els objectius concrets foren:

- Aprenentatge del conceptes i de la metodologia necessària per la realització i anàlisi de les proves biològiques de paternitat.
- Analitzar els electroferogrames de cada individu implicat obtinguts mitjançant el *kits AmpF/STR* Identifiler Plus I Powerplex® 16 system a fi d'obtenir el seu perfil genètic.
- Una vegada obtinguts els perfils genètics de les persones implicades en el cas de paternitat, comprovar si existeixen incompatibilitats i, en cas de detectar-ne, veure a quins i a quants de marcadors es donen.

- En cas de no poder excloure el suposat pare com a pare biològic del fill, calcular l'Índex de Paternitat i la Probabilitat Acumulada de Paternitat.
- En base als resultats anteriors determinar quins dels marcadors tenen més poder de discriminació.
- A cada electroferograma identificar els productes *stutter* de cada al·lel i de cada marcador i mesurar el percentatge de l'alçada del pic respecte al pic de l'al·lel.
- Analitzar la presència d'altres artefactes (*drop-out*, *drop-in*, microvariants, pics *pull-up*, adenilacions incompletes, desequilibris severos entre al·lells o entre locus, artefactes tècnics, patrons tri-al·lèlics i mutacions parentals) i com poden afectar aquests en la interpretació dels resultats.

4.RESULTATS

S'han analitzat 25 casos de paternitat realitzats en el darrer any en el servei de paternitats de la UIB, on s'inclouen 20 paternitats amb 3 membres (pare, mare i fill) i 5 paternitats amb 2 membres (pare i fill). Per altre banda, el nombre d'individus implicats en l'estudi són 66, ja que hi trobem individus que estan implicats en dues paternitats. Per cada individu s'han analitzats els STR inclosos en el *kit AmpF/STR Identifiler* i en el *kit Powerplex® 16 system*; per tant, tenint en compte els marcadors inclosos en cada *kit*, s'han analitzat i interpretat un total de 1980 cromatogrames.

4.1- Resultats de l'estudi de paternitat

Dels casos estudiats, un 64% corresponen a assignacions positives i un 36% corresponen a exclusions de paternitat. Aquests resultats són semblants als obtinguts en el Servei de Paternitats de la UIB al llarg de la seva trajectòria (69% i 31%, respectivament)(dades no publicades).

4.1.1- Exclusions

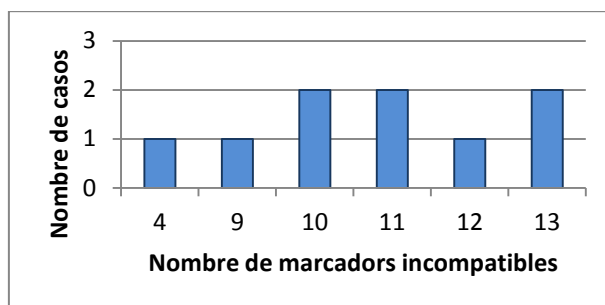
Per tal de que una paternitat sigui considerada exclusió, almenys tres sistemes han de ser incompatibles. La incompatibilitat es manifesta quan el fill té un al·lel que no està present ni a la mare ni al suposat pare (incompatibilitat de primer grau) o quan el fill no té un al·lel que el pare hagi de transmetre necessàriament a tota la seva descendència (incompatibilitat de segon grau) (**Taula 3**). És important indicar de quins tipus es tracte, ja que en funció de si és de primer o de segon grau, és més o menys possible que es degui a una mutació o a qualche artefacte.

Taula 3 . a) Exemple d'exclusió de primer grau: l'al·lel 9.3 no ho presenta ni la mare ni el presumpte pare. b) Exemple d'exclusió de segon grau: el presumpte pare no coincideix amb cap al·lel del fill

Marcador	Pare	Fill	Mare
a. THO1	6-7	6-9.3	6-8
b. D13S317	8-9	11-12	11-12

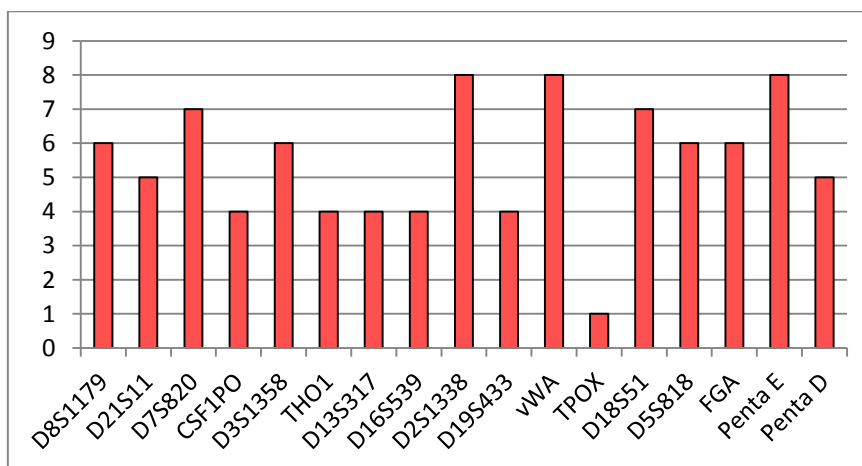
En les paternitats estudiades, el nombre de incompatibilitats de primer grau és de 66, en canvi, les exclusions de segon grau només hi són presents 24 vegades.

Com s'ha mencionat anteriorment, 9 casos varen correspondre a exclusions, però el nombre (**Gràfic 1**) i el tipus de marcadors (**Gràfic 2**) que donen incompatibilitat són diferents en cada cas. Dels 17 marcadors tipificats, s'han analitzat quants marcadors exclouen en cada una de les paternitats i, a més, amb quina freqüència han estat implicats en incompatibilitats.



Gràfic 1. Nombre de marcadors incompatibles en els casos d'exclusió

Els marcadors que més han estat implicats en exclusions són el D2S1338, el vWA i el Penta E, presents en 8 de les 9 exclusions de paternitat. Aquests resultats concorden amb l'esperat tenint en compte que el Penta E és el sistema que més variació presenta entre els diferents individus de la població segons la base de dades emprades, és a dir, és el que més al·lels presenta. A altres poblacions mediterrànies també s'ha descrit el Penta E i el D2S1338 com els STRs més discriminants [18]. En canvi, el marcador que menys vegades (sols en un cas) indicava incompatibilitat en les paternitats on hi havia exclusió és el TPOX, sistema descrit com a poc variable a la població espanyola [32]. A més, cal destacar, que cap dels STRs ha presentat incompatibilitat en totes les paternitats o en cap d'elles.



Gràfic 2. Nombre de vegades que han presentat incompatibilitat cada un dels marcadors

4.1.2- Assignacions positives

Els resultats dels casos en la pràctica forense han d'acompanyar-se de valors estadístics que permetin valorar i interpretar adequadament els resultats obtinguts. Per una banda, és necessari saber quins dels marcadors presenta major poder de discriminació *a priori*, és a dir, tenim més probabilitat de donar una incompatibilitat en cas de que el suposat pare no sigui realment el pare biològic del fill.

Com hem mencionat anteriorment, alguns marcadors presentaran més variants al·lèliques que d'altres, fet que implica que tindran major poder de discriminació.

Per una altra banda, quan s'analitza un cas concret amb assignació positiva s'ha de calcular l'IP parcial de cada marcador en els individus estudiats, que serà major com menys freqüent sigui l'al·lel compartit entre el fill i el suposat pare. L'IP total (producte dels parcials) s'incrementa cada vegada que afegim un marcador, per tant, podem deduir, que a major nombre de marcadors estudiats, més precís serà l'anàlisi de DNA realitzat, el que es traduirà amb una probabilitat acumulada superior.

En els 16 casos analitzats que foren compatibles es va calcular mitjançant el programa *Famílias* els IP parcials i totals, així com la probabilitat acumulada del cas. En la següent taula es mostren el valors totals per a cada un dels casos, així com, la mitjana de l'IP parcial per a cada un dels marcadors.

Taula 4. Probabilitat total dels casos de paternitats estudiats i informativitat dels marcadors

PROBABILITAT TOTAL DELS CASOS			INFORMATIVITAT DELS MARCADORS	
Cas	Probabilitat (W)	Índex de paternitat (IP)	Marcador	Mitjana IP parcial± desviació estàndard
1	0,9999999983	589 117 581	D8S1179	3,68±2,12
2	0,9999999997	3 306 556 363	D21S11	4,72±4,68
3	0,9999999989	931 680 086	D7S820	3,01±0,98
4	0,9999999966	294 557 503	CSF1PO	2,17±1,57
5	0,999999996	221 248 255	D3S1358	2,65±1,34
6	0,9999999993	1 442 293 281	THO1	2,52±1,32
7	0,9999966926	302 348	D13S317	3,59±2,92
8	0,9999999699	33 274 384	D16S539	1,99±0,94
9*	0,999988944	90 448	D2S1338	4,70±2,41
10	0,999994	167 313	D19S433	5,39±4,47
11	0,999999997	294 557 503	vWA	2,91±1,93
12	0,999999939	16 491 585	TPOX	2,45±2,07
13**	0,4526	0,8268	D18S51	3,61±1,15
14	0,999999984	63 871 826	D5S818	1,89±1,06
15	0,999999626	2 672 235	FGA	3,19±1,17
16	0,999999568	2 312 942	Penta E	5,24±3,69
			Penta D	4,74±9,38

Es pot observar que la probabilitat de cada paternitat i l'IP total és diferent, depenent dels al·lels dels individus implicats als sistemes polimòrfics estudiats, però en tots els casos (excepte el número 9 i 13, que s'explicarà a l'apartat de mutacions) el resultat és molt resolutiu, ja que els índex de paternitat totals ens indiquen que la probabilitat de que el suposat pare sigui el pare real del fill és entre 167.313 i 3.306.556.363 vegades major que la probabilitat de que ho sigui un home escollit a l'atzar dins la població. El rang de probabilitats estaria entre el 99,999% i el 99,9999999%, superant per tant, en tots

els casos el valor del 99,73% que segons els predicats verbals de Hummel [24] és suficient per considerar aprovada una paternitat des d'un punt de vista pràctic. La jurisprudència espanyola ha admès com a vàlid aquest valor (Sentència Tribunal Suprem, dia 24 de novembre de 1992), tot i que és pràctica habitual recomanar valors majors del 99,9% com a mínim.

Es pot observar que tres dels casos de paternitat on no es disposava del perfil genètic de la mare (indicats en negreta a la taula) són els que han donat una probabilitat menor, per aquest fet, es recomana sempre que la mare participi en les anàlisi.

Finalment, podem observar una clara relació entre els marcadors que, com hem vist en l'apartat anterior, presentaven incompatibilitat en la majoria dels casos amb els que, en aquesta taula, presenten una mitjana de l'IP parcial major. Aquests marcadors són; Penta E i D19S433, representant als que tenen un major nombre d'al·lels i, per tant, tenen més probabilitats de presentar genotips diferents en diferents individus. Finalment, permeten excloure als qui no són els pares biològics, per una banda, i per l'altra, al tenir més al·lels major és la possibilitat de que en presentin amb menor freqüència al·lèlica, el que permetrà tenir IP parcials superiors. També cal destacar l'alta desviació que presenten algun sistemes, como el Penta E, això es deu a importants diferències de freqüències al·lèliques del l'al·lel compartit entre el pare i el fill a diferents casos.

Freqüència al·lèlica mínima

Abans de començar a calcular les probabilitats, el que s'ha de fer és mirar si qualque al·lel presenta una freqüència inferior a la freqüència al·lèlica mínima calculada per cada marcador de la base de dades emprada i, si és així, s'han de modificar les freqüències en el programa, introduint la respectiva freqüència mínima.

En una de les paternitats el marcador FGA presentava l'al·lel 23.2 amb una freqüència de 0,0018, per tant, en el seu lloc es va introduir la freqüència mínima que en aquest cas era de 0,0058.

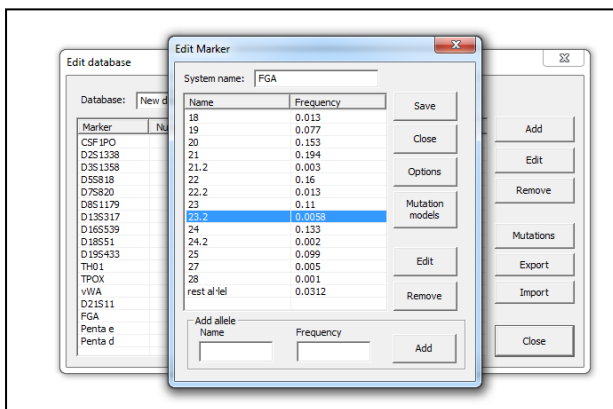


Figura 32. Canvi de les freqüències en el programa *Famílias*

Mutacions

Tot i que en els STRs l'aparició de mutacions és bastant freqüent i de fet, és el que permet que siguin tan polimòrfics, la taxa de mutació per generació és d'un 0,2%, fet que implica que tant sols esporàdicament l'observarem en una transmissió pare-fill o mare-fill. Així i tot, en els 25 casos estudiats, s'han trobat tres mutacions, dues de les quals corresponen a la mateixa paternitat, fet altament excepcional. En les dues paternitats implicades el marcador mutat és el D2S1338 i l'altre

marcador, només present en una paternitat, és el D5S818. Cada mutació s'ha d'estudiar per separat, ja que si l'al·lel mutat del fill prové del pare o de la mare o, en alguns casos, no es té clar, s'han de seguir unes fórmules diferents.

A títol d'exemple, es mostra a continuació la fórmula aplicada en una de les mutacions del marcador D2S1338, on l'al·lel mutat provenia de pare.

$IP = \frac{\mu(a+b)}{2ab}$; on μ és la taxa de mutació, a és la freqüència de l'al·lel A i b és la freqüència de l'al·lel B.

Aquests valors els trobem en les base de dades i el que s'ha de fer és calcular l'IP parcial per aquell marcador amb presència de la mutació.

$$IP = \frac{0.001526(0.0817+0.0745)}{2 \cdot 0.0817 \cdot 0.0745} = 0.0195806.$$

En el tres casos es calculà l'IP parcial tenint en compte la freqüència de mutació del marcador implicat i s'incorporà al càlcul de l'IP total. Cal indicar que la presència de mutacions redueix la probabilitat total del cas de paternitat. Així, en el cas on s'observà una mutació, l'IP fou més baix que l'habitual (99,998%) i en el cas on s'observaren dues, fou extremadament baix (45,26%). En aquest darrer cas es genotiparen 38 marcadors autosòmics addicionals, obtenint-se finalment un valor total del 99,992%.

4.2- Artefactes

a-Stutter

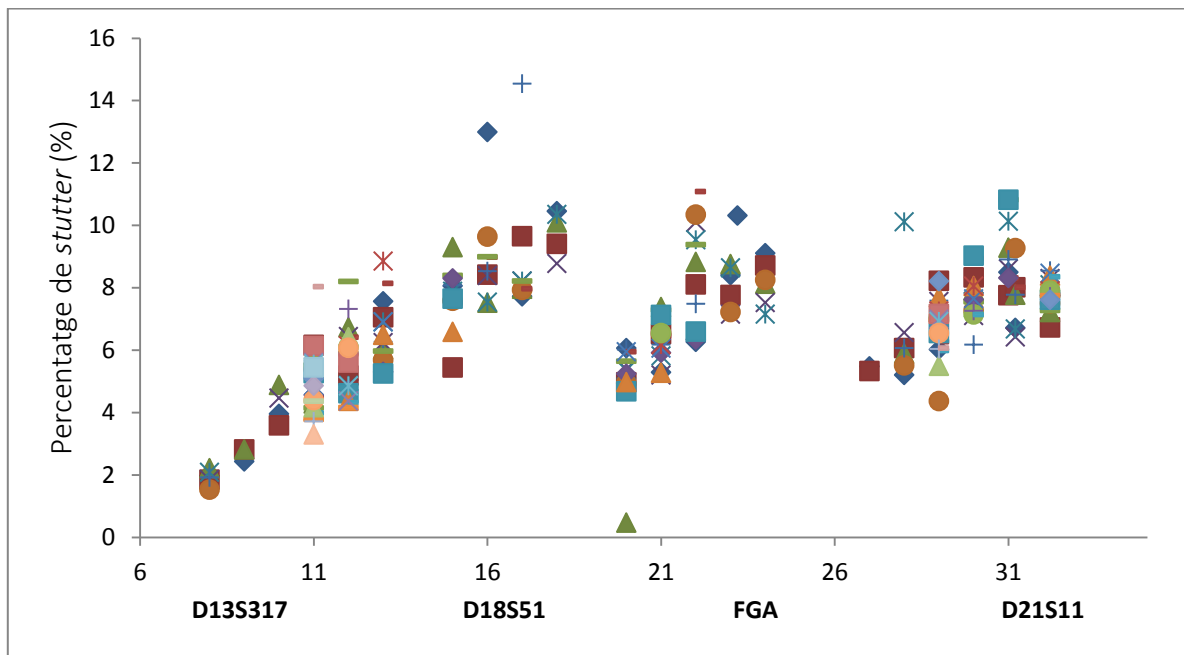
Les condicions de la PCR són iguals per a tots els marcadors, per tant, és molt difícil fer que el resultat de l'amplificació sigui el màxim i el més efectiu per a cada un d'ells. Aquest és un dels motius que explica que el percentatge d'*stutters* varia entre els STRs. També la presència d'aquest artefacte, com s'ha explicat en la introducció va lligada a l'estructura i longitud del *repeat*.

Per tal de quantificar el percentatge dels 1157 productes *stutter* trobats, es va comparar l'alçada de cada pic *stutter* amb cada pic original corresponent, analitzant un total de 2314 pics. El percentatge del producte *stutter* es determinà dividint l'alçada del pic *stutter* entre l'alçada del pic original.

Com podem observar en el **Gràfic 3**, dins cada marcador a major longitud de l'al·lel, major és el percentatge d'*stutter*, independentment del marcador estudiat, concordant amb el descrit a la bibliografia [32]. Cal mencionar però, que aquesta correlació no es veu tant clara en tots els marcadors, tal vegada conseqüència de les poques mostres analitzades. Així i tot, en la majoria dels STRs, el pic de l'*stutter* no representa més del 15% respecte al pic original, que és l'esperable quan són STRs com els emprats en aquest *kits*, majoritàriament tetanucleotídics.

En la **Taula 5** hi trobem el percentatge que representa cada *stutter* en relació al pic original i, a més, la freqüència en que s'han trobat en totes les mostres estudiades. Es pot observar que en la gran majoria de marcadors es veu una bona correlació entre l'estructura del *repeat* i l'aparició d'aquest tipus d'artefacte. Així veim que el sistemes Penta E i Penta D, que són els únics pentanucleotídics i que a més presenten una estructura simple, són els que menys vegades presenten *stutter* i si ho fan,

la intensitat del pic és baixa (3,05 i 5,39%, respectivament), concordant amb el descrit per altres autors [1]. També el TPOX presenta molt baixa intensitat, possiblement perquè té una composició simple i no presenta cap al·lel incomplet, concordant amb el descrit a la bibliografia [6].



Gràfic 3. Percentatge d' stutter corresponent als diferents al·lels de 4 marcadors

En general s'observa que els sistemes amb més complexitat presenten una intensitat superior dels pics *stutter* (>7,5%) que els sistemes simples, amb l'excepció del sistema D18S51 i en menor mesura del THO1. En el cas del sistema D18S51, aquest fet ha estat descrit anteriorment [6], l'estructura no n'explica els valors observats, per tant podria ser que les condicions de la PCR multiplex no fossin les més idònies per aquest sistema. En el cas del THO1, el valor mitjà que hem trobat no concorda amb la bibliografia [6]; és un sistema en general en poca presència d' *stutter*, especialment perquè l'al·lel més freqüent a la població és un al·lel incomplet (9.3), i aquests tipus d'al·lels donen menys *stutter* que els complets. No obstant, s'ha de mencionar que la desviació trobada és molt alta, per tant s'haurien d'estudiar més casos per valorar adequadament la presència d'aquest artefacte al sistema THO1.

Resumint, el percentatge i la intensitat observada d'aquests tipus de pic ha estat en general baix (sempre inferior al 9%), cosa que permet diferenciar-los de l'original. No obstant, s'han de tenir en compte a l'hora d'assignar el valor dels pics, ja que en alguns casos poden causar confusió i ser assignats com a al·lel.

Taula 5. Percentatge i freqüència d'*stutter* en cada un dels marcadors

Marcador	Seqüència del <i>repeat</i>	Mitjana % <i>Stutter</i> ± Desviació estandard	Freqüència absoluta
D8S1179	TCTA Complex	7,68±1,09	65
D21S11	TCTA Complex	7,80±2,26	100
D7S820	GATA	5,68±2,06	60
CSF1PO	AGAT	6,90±1,99	62
D3S1358	TCTA Complex	8,05±4,65	81
TH01	AATG	7,33±6,65	69
D13S317	TATC	5,79±6,25	86
D16S539	GATA	6,20±2,68	86
D2S1338	Compost (TGCC) (TTCC)	8,64±1,87	103
D19S433	Hipervariable	8,73±2,45	79
vWA	TCTA Complex	7,78±2,26	69
TPOX	AATG	2,87±2,29	50
D18S51	AGAA	8,44±2,16	75
D5S818	AGAT	6,02±1,69	61
FGA	TTTC complex	6,81±1,81	77
Penta E	AAAGA	3,05±1,16	16
Penta D	AAAGA	5,39±3,55	18

b-Pull-up

El *pull-up* és un dels artefactes més freqüent en l'estudi de les paternitats, però gràcies a que el programa té la capacitat d'indicar, mitjançant una línia fúcsia, a quina zona concreta s'ha generat, es facilita molt la seva detecció.

En les 25 paternitats s'han trobat 43 casos de *pull-up*, alguns que es veia clarament que eren un artefacte, ja que el pic produït quedava fora del *ladder* al·lèlic i, en canvi, altres que es podien confondre amb un pic al·lèlic. Aquest darrer cas és el que trobem en la **Figura 33**, però si observem el perfil del pic corresponent a l'artefacte i, gràcies a que ja hi ha dos pics ben definits, es pot descartar aquest tercer com a pic bo.

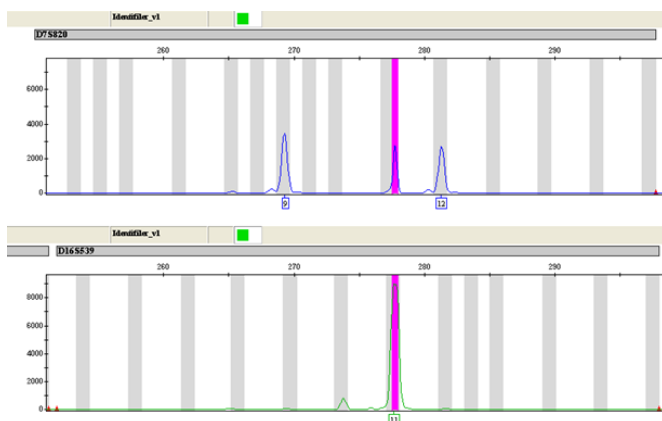
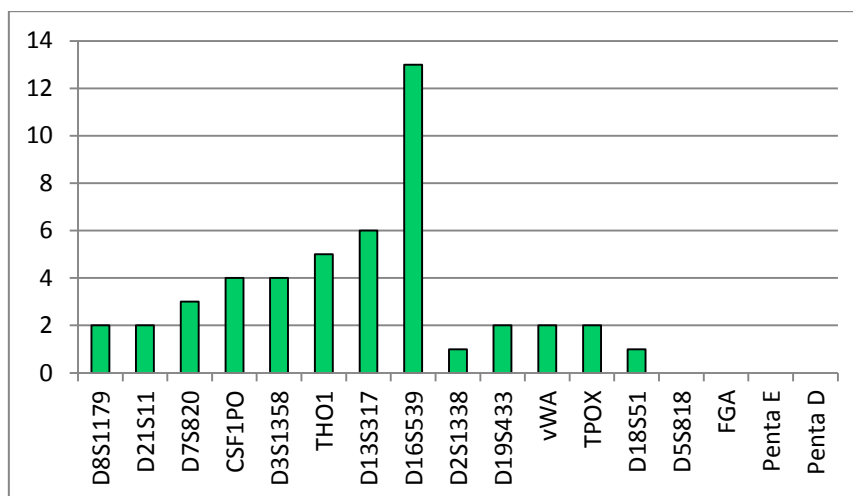


Figura 33. Exemple de *pull-up*

A continuació, es mostra el recompte de vegades que un mateix marcador ocasionava l'artefacte; resultant el D16S539 el que més casos de *pull-up* va generar (13 vegades), seguit del D13S317 (6 vegades), presentant els dos fluorescència verda. Altres marcadors com el D5S818, el FGA, el Penta D i el Penta E no ocasionaren en cap ocasió un pic *pull-up*.



Gràfic 4. Recompte de *pull-up* ocasionats pel mateix marcador

c-Drop-out

El *drop-out* pot aparèixer com a conseqüència d'un DNA motlle degradat o amb poca quantitat, o, per altre banda, a conseqüència d'una mutació al lloc d'unió al *primer*, el qual es coneix com al·lel nul.

En aquest estudi s'ha trobat un *drop-out*, que sembla ser del tipus al·lel nul, al marcador Penta E, però gràcies a que es fa un duplicat amb un *kit* diferent, el qual utilitza *primers* distints, es va poder detectar l'artefacte. Si no s'empren amplificacions amb *primers* diferents és difícil detectar els al·lells nuls, sobretot si l'al·lel en qüestió no determina incompatibilitat en el cas de paternitat en qüestió, ja que si fos així, si que es tendria present com a mutació o al·lel nul.

Taula 6. Cas de *drop-out* observat en el present estudi

Sistema d'amplificació	Pare	Fill	Mare
<i>AmpF/STR</i> Identifiler	13	5-7	7 <input type="checkbox"/> → <i>Drop-out</i>
Powerplex® 16 system	13	5-7	7-18

d-Adenilació incompleta

La presència d'adenilació incompleta és bastant freqüent, tot i que hi són presents amb diferents graus. A la **Figura 34** veiem que tant en la mostra 1 com en la mostra 2, el marcador D3S1358 presenta adenilació incompleta, encara que amb un grau d'intensitat molt diferent.

Una vegada analitzades tots els cromatogrames, en el **Gràfic 5** es pot observar quins són els marcadors que presenten una major incidència d'aquests tipus d'artefacte.

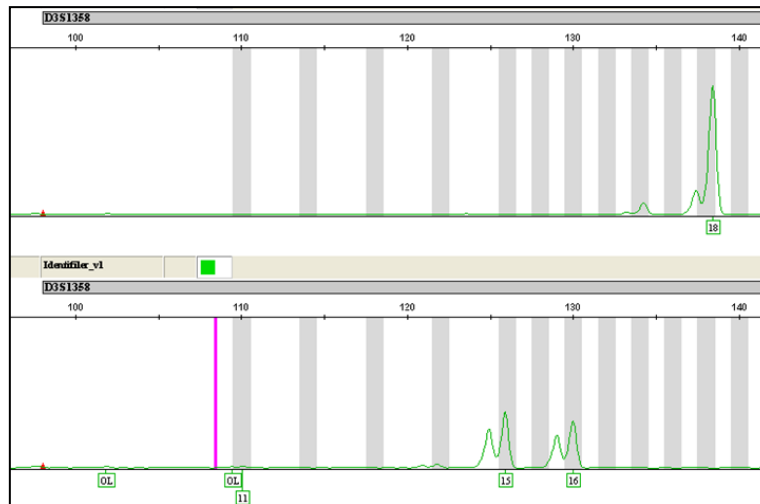
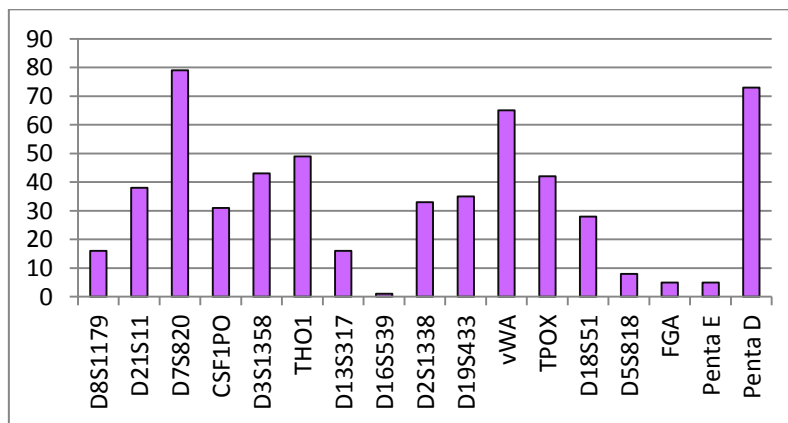


Figura 34. Exemple d'adenilació incompleta; el marcador D3S1358 presenta una major proporció d'amplicons amb adenilació incompleta



Gràfic 5. Incidència d'adenilació incompleta

Els marcadors que han presentat un major nombre de pics d'adenilació incompleta són el Penta D i el D7S820, seguit del vWA i el THO1. Aquest fet es pot explicar per les condicions de la PCR o per un excés de mostra, encara que també, podria ser a conseqüència dels *primers* utilitzats que afavoreixen menys l'adenilació completa en aquests 4 marcadors.

Els casos on el pic (-A) té una alçada important respecte al pic al·lèlic (>30%) han suposat sols un 8,1% del total, destacant els marcadors D8S1179 i el D3S1358 com els que, en un major nombre de casos, han presentat adenilacions incompletes elevades (38% i 28%, respectivament). Així i tot, s'ha de mencionar que en cap cas s'han donat pics amb una alçada que fes córrer el perill d'una incorrecte interpretació al·lèlica.

f-Al·lels *off-ladder* (Microvariants)

En els 1122 marcadors analitzats en aquestes 25 paternitats no s'ha trobat cap cas de microvariant no descrita, però cal mencionar que trobar-ne alguna és molt difícil, ja que els al·lels nous que van sortint en els nombrosos estudis relacionat amb els STRs humans usats en el camp de la Genètica Forense, es van tipificant i registrant en la base de dades.

g-Drop-in

El fet de que les mostres es trobin en bon estat i en gran quantitat, fan que la presència de *drop-in* sigui nul·la, almanco en aquestes paternitats estudiades.

h-Patrons de triple pic

No s'ha observat en cap cas un patró de triple pic, fet que és explicable per la raresa d'aquest fenomen.

g-Allele imbalance i locus-to-locus imbalance

En general podem dir que no s'ha observat una amplificació preferencial severa (diferència superior al 30%) entre els al·lels d'un individu heterozigot. És lògic que sigui així ja que, per una banda, no s'espera que hi hagi mescla degut a la procedència indubtable de les mostres analitzades i, per l'altra, la possibilitat d'una mutació al lloc d'unió al *primer* que disminueixi la eficiència de l'amplificació, és molt remota.

El desequilibri entre locus s'ha observat en alguns casos, molt probablement per excés de DNA a la reacció de PCR, tot i que en cap cas ha estat molt acusat. En el cas que es mostra a la **Figura 35**, s'observa un desequilibri entre locus i també una amplificació preferencial important entre els dos al·lels dels sistema TPOX. En aquest cas, es va repetir la reacció de PCR emprant menys quantitat de DNA motlle, eliminant-se així l'artefacte.

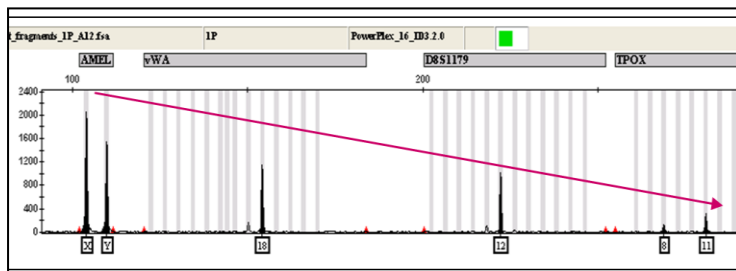


Figura 35. Desequilibri d'amplificació entre locus

h-Artefactes tècnics: renou, bombolles, *spikes* i contaminants

Atès l'origen de les mostres emprades no era esperable trobar la presència de contaminants. Tampoc s'han observat pics de tipus *spike*.

El que sí s'ha observat en bastants de cromatogrames és la presència de renou o bombolles, però com que el seu perfil està molt definit, són identificats fàcilment i no han dificultat en cap cas la lectura.

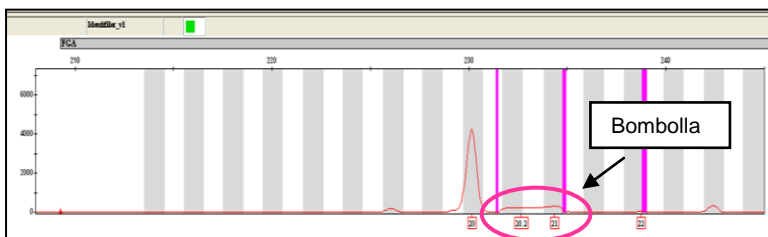


Figura 36. Exemple de bombolla

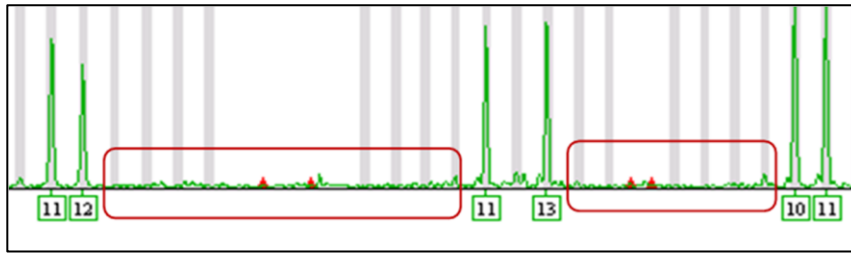


Figura 37. Exemple de renou

5.CONCLUSIONS

Actualment, els STRs són essencials en el camp de la Genètica Forense, fent possible l'anàlisi biològic de les paternitats. L'elevada variació intrapoblacional que presenten fa que els 17 STRs estudiats tinguin un alt poder de discriminació. Així i tot, s'han trobat marcadors que tenen un poder de discriminació superior que la resta, com és el cas del Penta E i el D2S1338, representant als STRs més discriminats. De manera contrària, el TPOX és el marcador que ha presentat un poder de discriminació inferior, concordant amb el que està descrit a la població espanyola [38].

En tots els casos compatibles, exceptuant els dos casos amb presència de mutació, el valor corresponent a l'índex de paternitat és entre 167.313 i 3.306.556.363 i el rang de probabilitat és entre 99,999% i 99,9999999%, sent aquests uns resultats molt resolutius i en el cas de les probabilitats, superen el valor (99,73%) establert per els predicats verbals de Hummel [24]. Així i tot, la certesa mai és del 100%, per tal manifestar que la certesa absoluta no existeix en aquest tipus d'anàlisi.

Malgrat la intenció d'optimitzar al màxim les condicions de la PCR per part de les cases comercials, en la majoria dels casos hi són presents els artefactes. Els *stutter* són els més freqüents, presents sobretot en els marcadors amb una organització més complexa i amb una seqüència de *repeat* més curta, com és el cas dels D2S1338 i el D19S433. Seguint el mateix criteri, el Penta E i el Penta D, els quals presenten un estructura simple i, a més, són pentanucleotídics, presenten una freqüència i intensitat molt menor de pics *stutter*. A més, dins un mateix marcador, a mesura que incrementen el nombre de repeticions, augmenta la intensitat del pic *stutter*. Així i tot, el percentatge corresponent al pic *stutter* en cap cas supera el 15%. El segon artefacte més freqüent és l'adenilació incompleta, present sobretot en els sistemes Penta E i D7S820, però tan sols en un 8,1% del total l'alçada del pics superava en un 30% el pic original. Els *pull-up*, *dropt-out*, *allele imbalance*, *locus-to-locus imbalance* i els artefactes tècnics s'han trobat, però amb menor freqüència i no han dificultat la genotipació i, finalment, la resta d'artefactes no s'han trobat. D'aquesta manera és necessari reconèixer cada un d'ells i assignar-los com a tal, així com reconèixer els seus perfils i les seves freqüències, ja que poden derivar a resultats erronis. La gran quantitat d'artefactes presents fa pensar en intentar trobar una estratègia millor per optimitzar encara més les condicions de la PCR, sobretot en una estratègia on les condicions siguin individualitzades per a cada un dels marcadors, encara que això complicaria molt l'estudi i seria molt més costós.

Finalment, per tal d'obtenir uns resultats molts més definits s'haurien d'incloure més paternitats, ja que en alguns d'ells hi havia poques mostres i determinats artefactes no s'han trobat. Malgrat això, les tendències han estat evidents i, en general, no divergeixen amb la bibliografia consultada.

6. BIBLIOGRAFIA

1. Bacher, JW, Hennes, LF, Gu, T, Tereba, A, Micka, KA, Sprecher, CJ, Schumm, JW (1998). **Pentanucleotide Repeats : Highly Polymorphic Genetic Markers Displaying Minimal Stutter Artifact.**
2. Brookes C, Bright J, Harbison S, Buckleton J (2012). **Characterising stutter in forensic STR multiplexes.** Forensic Science International: Genetics. Volume 6, Issue 1, Pages 58–63.
3. Butler JM, McCord BR, Jung JM, Allen RO (1994) **Rapid separation of the short tandem repeat HUMTH01 by capillary electrophoresis.** *BioTechniques* 17: 1062-1070
4. Butler JM (2001). **Forensic DNA typing. Biology, Technology, and genetics of STR markers.** Segona edició. Editorial: Academic Press.
5. Butler JM (2001). **Commonly used short tandem repeat markers, Biology of STRs: stutter products, non-template addition, microvariants, null alleles and mutation rates, Additional DNA markers: amelogenin, Y-chromosome STRs, mtDNA, SNPs, Alu repeats.** En: Butler JM (ed), Forensic DNA typing, Academic Press, San Diego, USA.
6. Butler JM (2005). **Forensic DNA typing. Biology, Technology and Genetics of STR markers.** Segona Edició. Editorial: ELSEVIER
7. Butler JM (2007). **Advantages for STR Markers STR and Molecular Biology Artifacts Types of STR Repeat Units**, 51, 1–6.
8. Butler JM, Hill CR, Kline MC, Duewer DL, Sprecher CJ, McLaren RS, Rabbach DR, Krenke BE, Storts DR (2009) **The single most polymorphic STR locus: SE33 performance in U.S. populations.** *Forensic Science International: Genetics Supplement Series (Progress in Forensic Genetics 13)* 2: 23-24. [Disponible en <http://www.fsigeneticsup.com>]
9. Butler JM, Hill CR (2010) **Scientific Issues with Analysis of Low Amounts of DNA.**
10. Butler JM, Reeder J. **NIST Standard Reference Database.** Consulta: [20/3/2015]. Disponible en: <http://www.cstl.nist.gov/biotech/strbase/>.
11. Bzymek M, Lovett S (2001). **Instability of repetitive DNA sequences: The role of replication in multiple mechanisms.** The National Academy of Sciences, vol. 98, no. 15, p. 8319–8325.
12. Clayton TM, Guest JL, Urguhart AL, Gill PD (2004). **A genetic basis for anomalous band patterns encountered during DNA STR profiling.** *J Forensic Sci.* Nov;49(6):1207-14.
13. Clayton TM, Hill SM, Denton LA, Watson SK, Urguhart AJ (2004). **Primer binding site mutations affecting the typing of STR loci contained within the AMPFISTR SGM Plus kit.** *Forensic Sci Int* ;139:255-9.
14. Coble M, Butler J (2005). **Characterization of New MiniSTR Loci to Aid Analysis of Degraded DNA***. *J Forensic Sci*, Jan. 2005, Vol. 50, No. 1
15. **DNA Fragment Analysis by Capillary Electrophoresis.** Consulta: [25/04/2015]. Disponible en: <https://tools.lifetechnologies.com/content/sfs/manuals/4474504.pdf>
16. **Forensic Bioinformatics.** [Consulta 26/04/2015]. Disponible en: <http://www.bioforensics.com/dna-testing-issues>
17. Franc F, Castell A, Verd F (2008). **El diagnóstico genético del sexo mediante el test de la amelogenina : Métodos y posibles fuentes de error .** Sex typing through the Amelogenin test: Methods and, 14(52), 119–125.
18. Garofano L, Pizzamiglio M, Donato F (1999). **Italian population data on two new short tandem repeat loci : D2S1338 and Penta E**, 105, 131–136.
19. Geada, H (n.d.). **Algunos aspectos puntuales sobre la investigación biológica de la paternidad**, 1–10.
20. Gill, P (1996). **The role of Short Tandem Repeat (STR) DNA in forensic casework in the UK-past, present and future perspectives**, 1–23.
21. Goodwin W, Linacre A, Hadi S (2007). **An introduction to Forensic Genetics.** Editorial: John Siley & Sons.
22. Goodwin W, Linacre A, Hadi S (2011). **An introduction to forensic genetics.** Segona edició. Editorial: Wiley
23. Guo WJ, Ling J, Li P (2009). **Consensus features of microsatellite distribution: Microsatellite contents are universally correlated with recombination rates and are preferentially depressed by centromeres in multicellular eukaryotic genomes.** *Genomics*. 93:323–331.
24. Hummel K, Gerchow J (1981). **Biomathematical Evidence of Paternity.** Springer Verlag, Berlin; 1981.
25. kitAmpFISTR® Identifier® PCR Amplification Kit User's Manual (2012). [Consulta: 26/02/2015]. Disponible en: https://tools.lifetechnologies.com/content/sfs/manuals/cms_076395.pdf
26. kitPowerplex® system (Promega) kit User's Manual (2013). [Consulta: 26/02/2015] Disponible en: <https://www.promega.es/~media/files/resources/protocols/technical%20manuals/101/powerplex%2016%20system%20protocol.pdf>
27. Luque A, Herráez A (2001). **Organización del genoma Eucariota, Clonación acelular: reacción en cadena de la polimerasa.** En: Luque JA, Herráez A. *Biología Molecular e Ingeniería genética.* Madrid: Ediciones Harcourt.
28. Martínez B (1999). **Introducción al estudio de los polimorfismos del ADN en Medicina Forese.** En Martínez Jareta MB (ed), *La Prueba del DNA en Medicina Forense.* Barcelona: Masson.
29. Oliva R, Oriola J, Ballesta F, Clària J, Mercadal L (2013). **Genètica mèdica.** Segona Edició. Editorial: UBE.
30. Owerbach D, Aagaard L (1984). **Analysis of a 1936bp polymorphic region flanking the human insuline gene.** *Gene*; 32: 475-479.
31. Panzer S, Kuhl DP, Caskey CT (1995). **Unstable triplet repeat sequences: a source of cancer mutations?.** *Stem cells.* Pages 146-57
32. Pérez A, Calafell F, Clarimón J, Bosch E, Mateu E, Gusmao L, Amorim A, Benchemsi N, Bertranpetit J (2000). **Allele frequencies of 13 short tandem repeats in population samples from the Iberian Peninsula and Northern Africa**, 208–214.
33. Ruitberg C, Reeder D, Butler J (2001). **STRBase: a short tandem repeat DNA database for the human identity testing community.** *Nucleic Acid Research.* Vol.29, N°1. Pages 320-322.
34. Sharma PC, Grover A, Kahl G (2007). **Mining microsatellites in eukaryotic genomes.** *Trends in Biotechnol.* 25:490–498.
35. Shi J, Huang S, Fu D, Donghui F, Jinyin Y, Xinfa W, Wei H, Shengyi L, Guihua L, Hanzhong W (2013). **Evolutionary Dynamics of Microsatellite Distribution in Plants: Insight from the Comparison of Sequenced Brassica, Arabidopsis and Other Angiosperm Species**
36. Shinde D, Lai Y, Sun F, Arnheim N (2003). **Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (GA/GT)_n and (A/T)_n microsatellites.** *Nucleic Acids Res.* 31:974–980.
37. Taula de freqüències per a l'exercici de intercomparació (2013). Disponible en: <http://www.cstl.nist.gov/strbase/>
38. Urguhart A, Kimpton CP, Downes TJ, Gill P (1994). **Variation in short tandem repeat sequences – a survey of twelve microsatellite loci for use as forensic identification markers.** *Int J Legal Med*; 107: 13-20.
39. Vasquez P, Garde M, Martínez B (2002). **Amplificación y tipado de marcadores genético-moleculares de tipo microsatélites (STRs) autosómicos y del cromosoma Y.**
40. Willems TF, Gymrek M, Highnam G, Mittelman D, Erlich, Y (2014). **The landscape of human STR variation.** *Genome Research.* doi:10.1101/gr.177774.114.