



UNIVERSITAT
ROVIRA I VIRGILI



LABORATORIO DE ANÁLISIS

ECHÉVARRNE

**TRABAJO DE FIN DE MÁSTER:
IMPLEMENTACIÓN DEL SOFTWARE *FAMILIAS*
EN EL ANÁLISIS DE PRUEBAS DE PATERNIDAD**

RAQUEL MARTÍNEZ ARROYO

MÁSTER UNIVERSITARIO EN GENÉTICA, FÍSICA
Y QUÍMICA FORENSE
2021-2022

TUTOR ACADÉMICO: ESTHER RODRÍGUEZ

TUTOR PROFESIONAL: RAÛL SANTAMARÍA MERINO

UNIVERSIDAD ROVIRA I VIRGILI
ECHEVARNE DPTO. GENÉTICA MOLECULAR
ABRIL – JUNIO 2022

CONVOCATORIA JUNIO 2022

Índice

1. Resumen/Abstract	1
2. Introducción	2
2.1 Fundamentos de la prueba de paternidad.....	3
2.2 Valoración estadística de la prueba de paternidad	5
2.3 Softwares bioinformáticos: PatCan, GestLab y Familias	6
2.4 Interpretación de los resultados	7
3. Objetivos	8
4. Metodología	9
4.1 Gestión de muestras de paternidad privadas y judiciales	9
4.2 Tipos de muestras biológicas	10
4.3 Extracción de ADN.....	10
4.4 Amplificación de ADN.....	11
4.4.1 Amelogenina	12
4.5 Análisis de fragmentos de ADN.....	13
4.6 Análisis bioinformático e interpretación: Familias.....	14
4.6.1 Pruebas de paternidad y comparación de resultados (15 marcadores)	15
4.6.2 Prueba de paternidad con 23 marcadores informados	15
4.6.3 Pruebas de paternidad con mutaciones	15
4.6.4 Pruebas de parentesco.....	17
5. Resultados y discusión	17
5.1 Amplificación diferencial.....	17
5.2 Pruebas de paternidad y comparación de resultados	18
5.3 Prueba de paternidad con 23 marcadores informados	19
5.4 Pruebas de paternidad con mutaciones	20
5.5 Pruebas de parentesco.....	20
6. Otras aplicaciones de los STRs	21
6.1 Enfermedad de Huntington.....	21
7. Conclusiones	22
8. Bibliografía	23
Anexo I	25
Anexo II	27
Anexo III	29

1. Resumen/Abstract

Las pruebas de paternidad permiten comparar el ADN entre individuos para probar su relación biológica paternofamiliar mediante marcadores genéticos del tipo STR o *Short Tandem Repeats*, dependiendo de la finalidad de la prueba la paternidad puede ser judicial o privada. La obtención del perfil genético de los participantes se basa en las técnicas de extracción del ADN, amplificación por PCR multiplex y electroforesis capilar. El resultado será inclusivo cuando existan compatibilidades en todos los marcadores, 3 o más discordancias será motivo de exclusión. El resultado se valora de manera estadística mediante el cálculo del LR teniendo en cuenta las frecuencias alélicas de la población de estudio, a mayor número de marcadores analizados más robusto será el resultado, el cálculo estadístico varía si se tienen en cuenta la presencia de mutaciones. Con la ayuda de softwares bioinformáticos como *Familias* el proceso es automatizado para trabajar de manera eficiente. El análisis de STRs también permiten el diagnóstico de enfermedades de origen genético como, por ejemplo, las ocasionadas por expansión de triplete.

Paternity tests allow the comparison of DNA between individuals to prove their biological parent-child relationship through the STR or *Short Tandem Repeat* genetic markers, depending on the purpose of the test paternity can be judicial or private. Obtaining the genetic profile of the participants is based on DNA extraction techniques, multiplex PCR amplification and capillary electrophoresis. The result will be inclusive when there are compatibilities in all markers, 3 or more discrepancies will be grounds for exclusion. The result is evaluated statistically by calculating the LR considering the allelic frequencies of the study population, the greater the number of markers analyzed the more robust the result will be, the statistical calculation varies if the presence of mutations is considered. With the help of bioinformatic softwares such as *Familias*, the process is automated to work efficiently. The analysis of STRs also allows the diagnosis of diseases of genetic origin, such as those caused by triplet expansion.

2. Introducción

Laboratorio Echevarne es una empresa familiar fundada en el 1958 por el Dr. Fernando Echevarne, constituye uno de los laboratorios de referencia más avanzados de Europa. Actualmente, ofrece servicio en diferentes áreas: Bioquímica, Hematología, Microbiología, Genética, Toxicología, Química, Industria...^[1] Las prácticas del presente TFM se han llevado a cabo en la Unidad de Genética, durante la estancia se han visto algunas de las pruebas y técnicas del departamento, prestando especial interés en las pruebas de paternidad.

Las pruebas de paternidad permiten comparar el ADN entre individuos para probar su relación biológica paternofilial mediante marcadores genéticos del tipo STR o *Short Tandem Repeats*, los STR son un tipo de polimorfismos de 2-6 nucleótidos repetidos en tándem. Este tipo de polimorfismos también son utilizados para el diagnóstico de enfermedades de origen genético como, por ejemplo, la enfermedad de Huntington.

La obtención de los perfiles genéticos puede realizarse de manera sencilla mediante los protocolos de extracción y amplificación del ADN y electroforesis capilar. Finalmente, la comparación de los genotipos y el resultado de la prueba de paternidad se valora de manera estadística mediante el cálculo del LR.

Aplicando los conocimientos adquiridos en el Máster, he contribuido a implementar el software *Familias* para el análisis bioinformático de las pruebas de paternidad. El programa *Familias* supondrá un avance, ya que permite analizar más marcadores de los que el software de análisis *GestLab*, que usa el laboratorio actualmente, analiza. Así como, incorporar las tasas de mutación de los marcadores en la valoración estadística y estudiar la posibilidad de realizar las pruebas de parentesco biológico (hermandades) mediante el software *Familias*. Este proyecto tiene como objetivo reanalizar más de 50 pruebas de paternidad con tal de reemplazar el software de análisis bioinformático actual, el cual se ha quedado obsoleto y es más rudimentario. La importancia de automatizar el proceso bioinformático de las pruebas de paternidad con softwares gratuitos como *Familias* supone trabajar de manera eficiente, disminuyendo tiempo de trabajo y posibles errores humanos.

2.1 Fundamentos de la prueba de paternidad

El ADN constituye el material genético de un individuo, está organizado en 22 parejas de cromosomas autosómicos y una pareja de cromosomas sexuales. A su vez, los cromosomas contienen los genes, las diferentes variantes de cada gen recibe el nombre de alelo. Debido al carácter diploide del genoma humano cada individuo tendrá dos alelos para un mismo gen, heredados por vía paterna y materna. Un individuo será homocigoto cuando los alelos de un mismo gen sean iguales, de lo contrario será heterocigoto. [2]

La secuencia del genoma humano es idéntica entre individuos en un 99.9%, únicamente un 1% de polimorfismos hacen posible distinguir el genotipo de cada persona. Los polimorfismos hacen referencia a la variabilidad que existe en un fragmento de ADN, cuanta más variabilidad, mejor es la identificación de los individuos. Existen diferentes tipos de polimorfismos, la prueba de paternidad se basa en los de tipo microsatélite o STR (*Short Tandem Repeat*). Los STR son combinaciones de 2-6 nucleótidos repetidos en tándem, el número de repeticiones da nombre al alelo y varía entre individuos (**Fig. 1**). El hecho de que la secuencia repetida sea corta implica que la probabilidad de encontrar dicha secuencia sin degradar sea mayor, es por eso que se utilizan los microsatélites en lugar de los minisatélites (repeticiones de 9-100 pb). Los marcadores STR usados en el test de paternidad se encuentran tanto en regiones de ADN génicas (codificantes) como en regiones no codificantes, aunque las más usadas en genética forense son las situadas en regiones no codificantes. [16] Las secuencias flanqueantes se conocen y pueden amplificarse mediante PCR (*Polymerase Chain Reaction*), concretamente, con el sistema PCR multiplex. [3],[4]

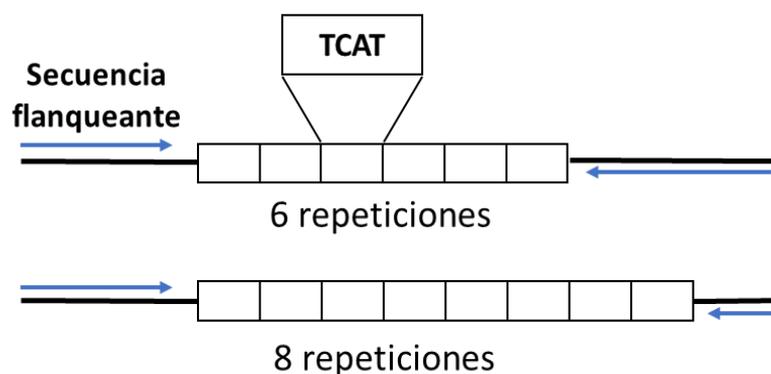


Figura 1. Para el marcador TH01 con la repetición del tetranucleótido TCAT el genotipo es 6/8.

El sistema de PCR multiplex permite amplificar varios STR a la vez, los kits comerciales incluyen múltiples parejas de *primers* marcados con distintos fluorocromos. Posteriormente, los fragmentos amplificados son separados mediante electroforesis capilar obteniendo así el perfil genético de un individuo (Fig. 2).

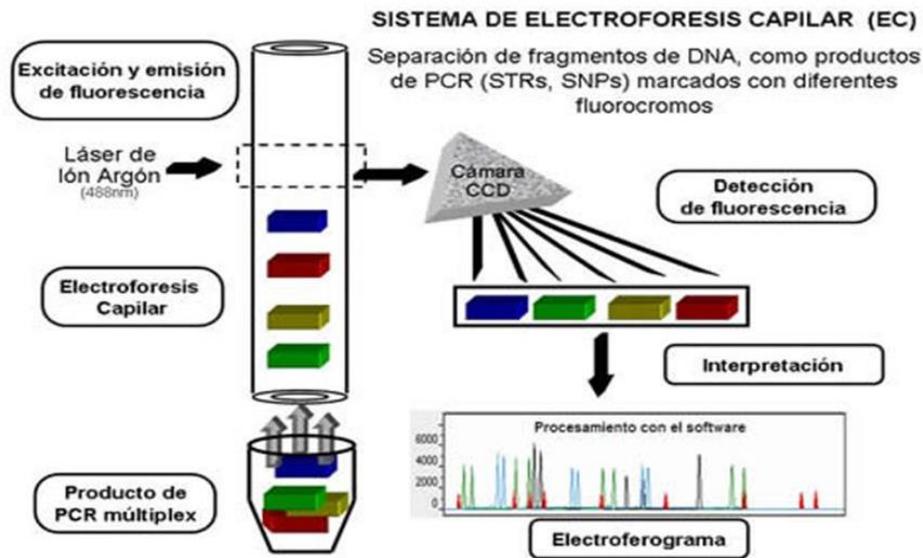


Figura 2. PCR Multiplex y electroforesis capilar

Cabe destacar que no todos los STR presentes en el genoma son útiles para la identificación genética. Solo aquellos que cumplan los siguientes requisitos serán utilizados con fin forense:

- Elevada heterocigosidad: la heterocigosidad está relacionada con el polimorfismo de un determinado marcador, un valor alto de heterocigosidad indica un alto grado de polimorfismo. Cuanto más polimórfico es un marcador más contundente es la discriminación entre diferentes perfiles genéticos.
- Tasa de mutación baja: marcadores con una tasa de mutación alta podrían presentar variaciones de una generación a otra pudiendo originar una falsa exclusión en una prueba de paternidad.
- Tamaño pequeño: Aquellos STR con menor tamaño serán fácilmente amplificables mediante PCR, los fragmentos más grandes tendrán más probabilidad de ser degradados.

2.2 Valoración estadística de la prueba de paternidad

El resultado de la prueba de paternidad se analiza de manera estadística mediante el valor LR (*Likelihood Ratio*). Para ello se establece la hipótesis nula (H0) de paternidad y la hipótesis alternativa (H1) de no paternidad, las bases del cálculo se basan en el Teorema de Bayes que define el cálculo de las probabilidades condicionadas:

H0: Probabilidad de obtener los resultados de la prueba de paternidad suponiendo que el perfil genético del supuesto padre pertenece al padre biológico.

H1: Probabilidad de obtener los resultados de la prueba de paternidad suponiendo que el perfil genético del supuesto padre pertenece a un individuo aleatorio de la población española.

$$LR = \frac{P(E|H0)}{P(E|H1)} \quad (A) \quad W = \frac{LR}{LR + 1} \quad (B)$$

Figura 3. A. Fórmula cálculo valor LR. B. Fórmula cálculo de la probabilidad de paternidad (W).

Por lo tanto, de acuerdo con la **Fig. 3A** el valor LR es una valoración estadística que permite conocer cuántas veces es más probable obtener los resultados genéticos (E, evidencia) bajo el supuesto H0, respecto la probabilidad de obtener la evidencia (perfil genético) bajo la hipótesis H1 de no paternidad, donde el supuesto padre es un individuo aleatorio de la población española.^[5]

El cálculo del LR se obtiene asumiendo que los marcadores cumplen el equilibrio de Hardy-Weinberg y que los individuos son analizados con las frecuencias alélicas representativas de la población correspondiente. Hablamos de equilibrio genético de Hardy-Weinberg para poblaciones grandes en las que no

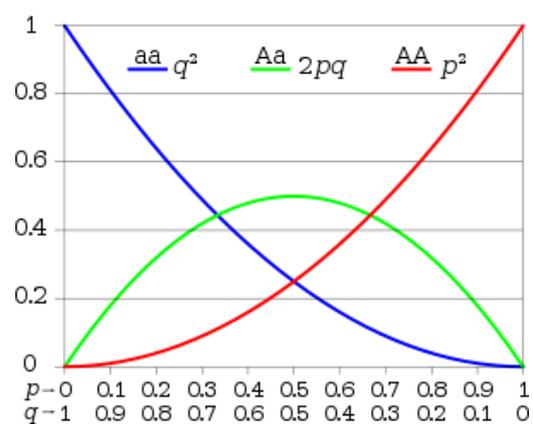


Figura 4. La Ley de Hardy-Weinberg para un locus con dos alelos (A a) con frecuencias alélicas p y q. La función predice que la frecuencia genotípica para AA es p^2 , para Aa es $2pq$ y para aa es q^2 .

intervienen factores de mutación o selección para un locus determinado y el valor de las frecuencias alélicas se mantiene de generación en generación (**Fig. 4**).

Finalmente, el índice de paternidad (IP) o LR es expresado en forma Probabilidad de paternidad (W) (**Fig. 3B**).

2.3 Softwares bioinformáticos: PatCan, GestLab y Familias

El cálculo estadístico puede realizarse con ayuda de softwares bioinformáticos como *GestLab* o *Familias*. Inicialmente, Echevarne realizaba la valoración estadística con el software sencillo *PatCan*, pero el aumento de la demanda ha hecho que ciertos procesos requieran ser optimizados con tal de agilizarlos y evitar el error humano.

Des de hace varios años, el laboratorio utiliza el software *GestLab*, dicho software fue creado en el laboratorio con la finalidad de integrar los ficheros de los perfiles genéticos obtenidos y realizar el cálculo estadístico, además de obtener un fichero con el resultado final capaz de ser enviado directamente a la interfaz de Echevarne para así crear el informe y validar el resultado.

A día de hoy, *GestLab* sigue ejerciendo su función correctamente, pero hay que tener en cuenta que fue creado con unos parámetros antiguos que han evolucionado y hacen que se quede obsoleto. Dichos parámetros son, por ejemplo, el número de marcadores que es capaz de informar, años atrás el laboratorio utilizaba el kit de amplificación de STR *NGM SElect* de *Thermo Fisher* el cual amplifica 15 STR más el marcador de sexo Amelogenina. Actualmente, con el kit de amplificación de STR *VeriFiler Express* de *Thermo Fisher* el laboratorio informa 23 marcadores autosómicos más los marcadores de sexo Amelogenina y YIndel. Por lo tanto, actualmente se están analizando más marcadores de los que el software permite informar.

Por otro lado, a menos marcadores informados, menor es el valor LR y menos robusto el resultado que se ofrece. Debido a todas estas limitaciones existe la necesidad de seguir evolucionando y encontrar un software actualizado que permita seguir trabajando de manera eficiente. En este sentido, el programa *Familias* es un software gratuito que permite calcular la valoración estadística de las pruebas de paternidad y parentesco genético incluyendo los 23 marcadores.^[6] El programa trabaja con ficheros

de formato tipo “.txt” compatibles con el programa de análisis de electroferogramas *GeneMapper* usado en el laboratorio.

2.4 Interpretación de los resultados

En un caso típico de paternidad entre un supuesto padre y un hijo/a se comparan los perfiles genéticos teniendo en cuenta que deben compartir un alelo para cada marcador, de lo contrario se considera exclusión (**Fig. 5**).

Cada laboratorio debe establecer el número de exclusiones que determina que el sujeto no es el padre biológico del hijo. El laboratorio Echevarne establece la exclusión de paternidad cuando existen incompatibilidades en 3 o más marcadores. En caso de existir 2 exclusiones suele pedirse muestra a la madre con tal de incrementar la exactitud de la prueba al obtener el alelo obligado paterno (**Fig. 6**). El alelo obligado paterno es aquel que necesariamente es heredado del padre.

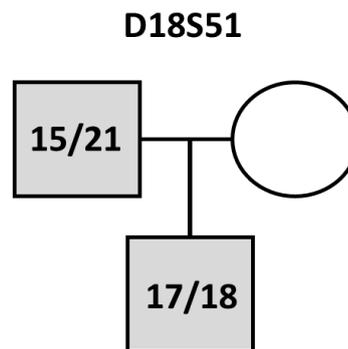


Figura 5. Exclusión para el marcador D18S51, padre e hijo no comparten ningún alelo.

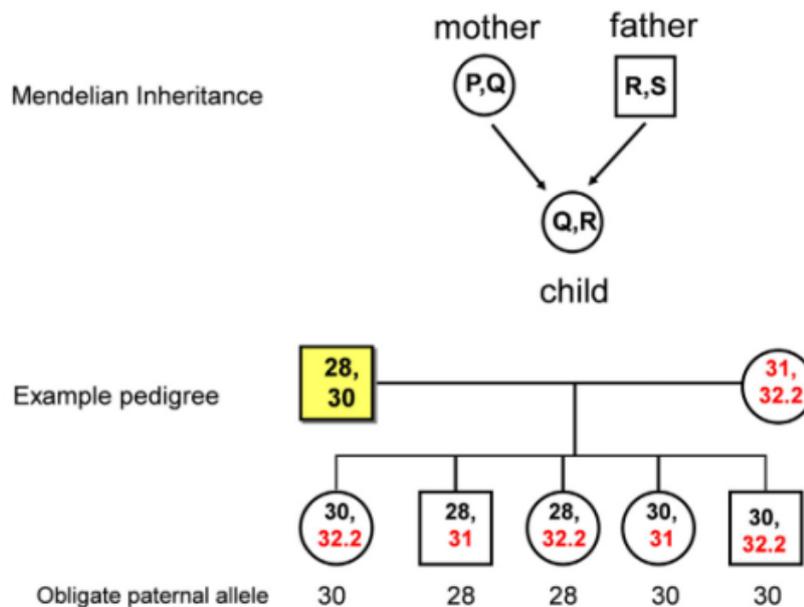


Figura 6. Representación del alelo obligado paterno siguiendo la herencia mendeliana.

En aquellos casos en que existe 1 incompatibilidad se considera la posibilidad de mutación en la línea germinal del progenitor (**Fig. 7**).

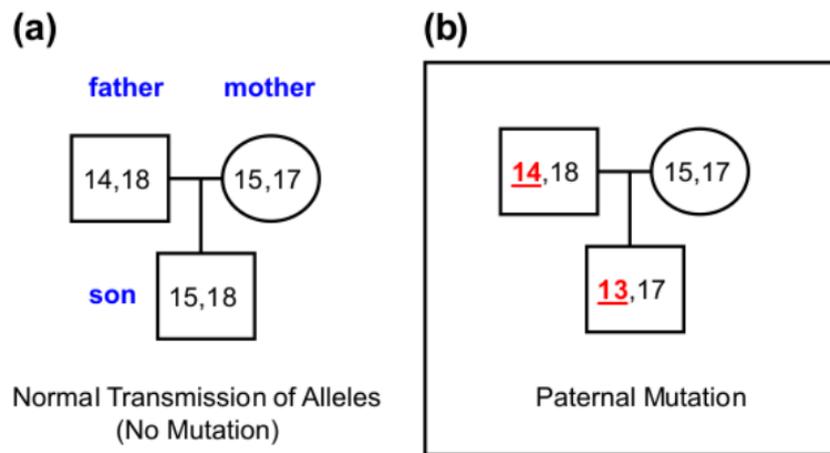


Figura 7. A. Trío Padre-Madre-Hijo con transmisión normal de los alelos. **B.** Trío Padre-Madre-Hijo con transmisión de una mutación en el alelo 14 paterno.

La paternidad se considera totalmente probada para valores de LR superiores a 10.000 y probabilidades de paternidad (W) del 99.99%.

En el **Anexo I** se muestra un ejemplo del informe proporcionado por el Laboratorio Echevarne para una prueba de paternidad.

3. Objetivos

El principal objetivo de la estancia de prácticas es contribuir a implementar el software *Familias* en el análisis de paternidades del Laboratorio Echevarne. Este desafío permitirá dejar de usar el software *GestLab* actual, el cual ha quedado obsoleto e incluye pasos más rudimentarios que dan lugar a errores humanos y suponen mayor tiempo de trabajo. Por lo tanto, la finalidad de este proyecto es optimizar el análisis bioinformático de las pruebas de paternidad.

Para llevar a cabo los objetivos ha sido necesario:

- Aprender a gestionar muestras de paternidad privadas y judiciales.
- Aprender los protocolos de extracción, amplificación y análisis de fragmentos de ADN del laboratorio.
- Procesar y almacenar los datos obtenidos de manera organizada mediante ficheros Excel.

- Analizar e interpretar los datos obtenidos con el software *Familias* respecto con *GestLab*.

4. Metodología

4.1 Gestión de muestras de paternidad privadas y judiciales

El procesamiento de las muestras comienza con su correcta recepción en el laboratorio, debemos tener en cuenta que existen dos tipos de prueba de paternidad en función del uso que se le dé al resultado: informativa (privada) o judicial. Las muestras recibidas en el laboratorio deben gestionarse acorde al tipo de prueba de paternidad.

Las paternidades informativas o privadas son aquellas en que el resultado se comunica a la persona que solicita la prueba. No hace falta identificar a todos los participantes, por este motivo no tiene ninguna validez legal y es totalmente confidencial. Las muestras biológicas han de venir acompañadas únicamente de un consentimiento informado (**Anexo II**).

Por el contrario, las pruebas legales son aquellas en que el resultado puede usarse como evidencia a nivel judicial. Requieren la identificación de todos los participantes y el seguimiento de la cadena de custodia de las muestras durante todo el proceso. Además de informar, obliga a los participantes a aceptar los resultados obtenidos por vía legal, ayudando a resolver casos vinculados al derecho civil, penal y de extranjería e inmigración. Las muestras biológicas deben llegar al laboratorio correctamente precintadas junto con tres copias del consentimiento informado judicial, la hoja de cadena de custodia (**Anexo II**) y la fotocopia de algún documento oficial en vigor que identifique a los participantes. En caso de participantes menores de edad es necesario contemplar los siguientes supuestos:

- Si el hijo/a es menor de edad debe constar una fotocopia del libro de familia.
- Si el hijo/a es menor de edad y no tiene los apellidos del supuesto/a padre/madre, debe tener la tutela y el consentimiento de algún tutor legal (madre, DGAIA...).
- Si el hijo/a es menor de edad y tiene los mismos apellidos del supuesto/a padre/madre únicamente podrá realizarse la prueba si se tiene la custodia.

4.2 Tipos de muestras biológicas

Para realizar la prueba de paternidad el laboratorio receptiona muestras de:

- **Saliva:** muestra recogida mediante frotis con hisopo bucal, el ADN se extrae de las células epiteliales que recubren el interior de la boca.
- **Sangre:** tejido formado por células sanguíneas y plasma, el ADN se extrae a partir de los leucocitos, ya que los eritrocitos son anucleados y no contienen material genético.

Las muestras deben llegar correctamente identificadas (**Fig. 8**), excepcionalmente pueden aceptarse cepillos de dientes y compresas. A continuación, se describe el protocolo de extracción de ADN para muestras de saliva y sangre.

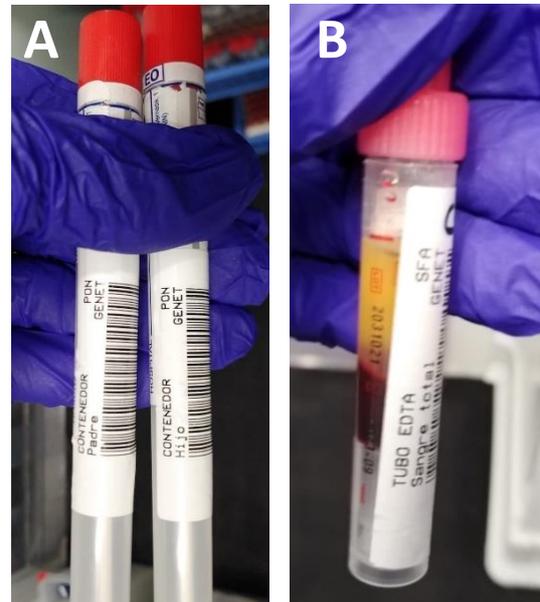


Figura 8. A. Hisopo bucal. **B.** Sangre total.

4.3 Extracción de ADN

La extracción del ADN consiste en la lisis celular (membrana citoplasmática y nuclear) y la liberación del material genético. Pueden utilizarse dos kits diferentes de extracción de ADN: *EZ1 Investigator* de Qiagen o *Prep-n-Go Buffer* de Applied Biosystems.

El kit *EZ1 Investigator* consiste en una extracción mediante bolas magnéticas, es un proceso más riguroso y se usa en caso de no tener muestras duplicadas ya que obtiene mejores resultados que Prep-n-Go. La reacción se lleva a cabo de manera automatizada en la máquina BioRobot EZ1 (**Fig. 9A**).

Por el contrario, el kit Prep-n-Go Buffer (**Fig. 9B**) extrae el ADN mediante la adición de 300 μ L de Buffer y el calentamiento a 90°C durante 20 minutos para lisar las células. Posteriormente, debe centrifugarse a 8.000 rpm y recuperar el DNA del sobrenadante. Este protocolo es más barato, aunque a veces puede fallar, por lo que se utilizará cuando las muestras existan por duplicado.

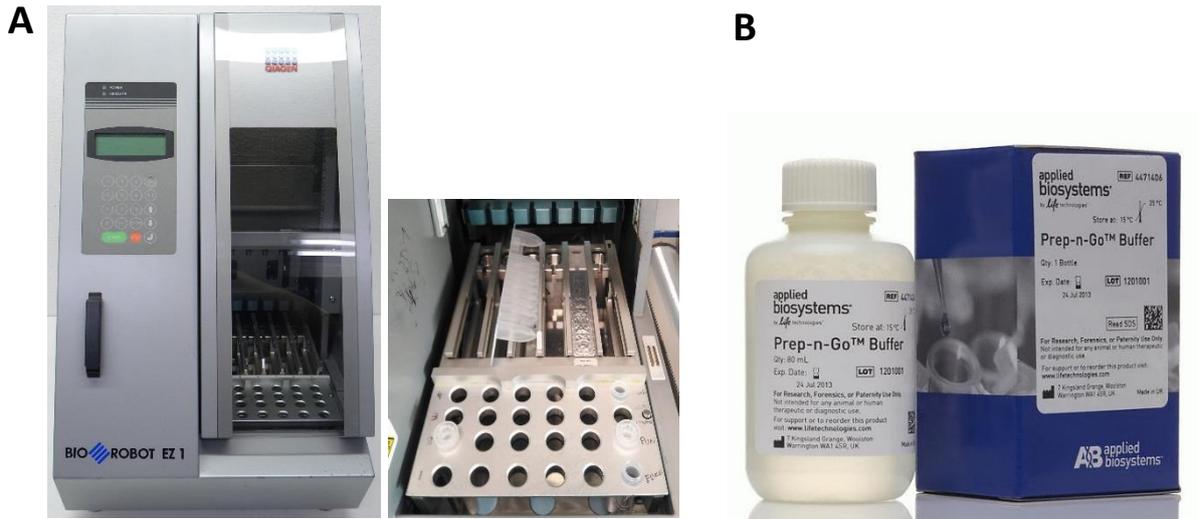


Figura 9. A. BioRobot EZ1 para extracción con EZ1 Investigator y montaje. **B.** Kit Prep-n-Go Buffer.

4.4 Amplificación de ADN

La amplificación de STR se realiza mediante PCR multiplex, el kit de amplificación *VeriFiler Plus* contiene un mix de *primers* marcados con 6 fluorocromos distintos, permitiendo así la amplificación de diferentes marcadores con distinto peso molecular simultáneamente (**Fig. 10**). La preparación de la muestra y las condiciones de la PCR se muestran en la **Fig. 11**.

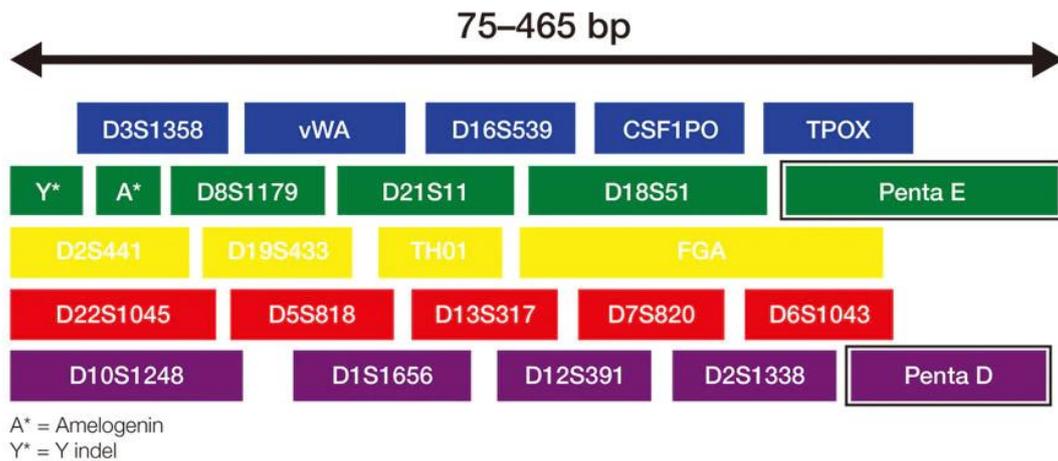


Figura 10. STR amplificados por el kit *VeriFiler Express* marcados por colores según el fluorocromo del primer.

Figura 11. Preparación de la reacción de PCR y condiciones.

Reacción PCR		Condiciones PCR		
MMIX	4 µL	95°C	1 min	x 27-29
Primers	4 µL	94°C	3 seg	
DNA	2 µL	59°C	16 seg	
		65°C	29 seg	
		60°C	5 min	
		4°C	∞	

El kit antiguo NGM Select amplifica los siguientes 16 STR y el gen de la Amelogenina: *D10S1248, vWA, D16S539, D2S1338, D8S1179, D21S11, D18S51, D22S1045, D19S433, TH01, FGA, D2S441, D3S1358, D1S1656, D12S391, SE33*. El kit actual *VeriFiler Express* permite amplificar los mismos marcadores excepto el *SE33* e incorporando: *CSF1PO, PENTA E, TPOX, PENTA D, D5S818, D13S317, D7S820, D6S1043* y *Yindel*. Es decir, en total se amplifican 23 marcadores y los marcadores sexuales Amelogenina y *Yindel* [8].

4.4.1 Amelogenina

La amelogenina es una proteína codificada en los cromosomas sexuales X e Y, que se encuentra formando parte del esmalte dental. Las diferencias que existen entre la secuencia del gen localizado en el cromosoma X (AMELX) respecto al Y (AMELY) permiten diagnosticar genéticamente el sexo de un individuo (**Fig. 12**). [11], [12]

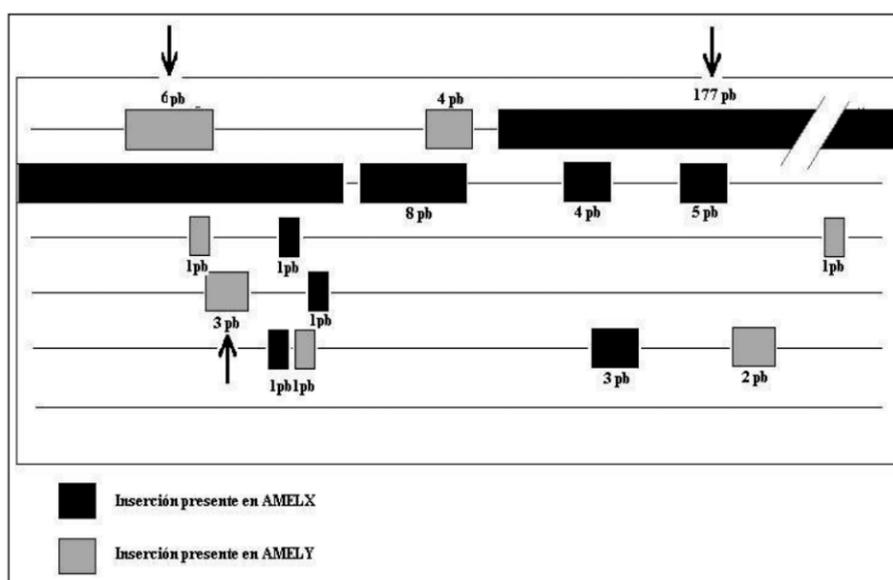


Figura 12. Divergencias del gen AMELX y AMELY, señaladas con una flecha las más usadas en el diagnóstico genético sexual.

Aquellos genes que contengan la deleción de 6pb (Chr. X) resultarán en cadenas de DNA más cortas respecto a aquellos genes que no contengan la deleción. Dado que únicamente el gen AMELX contiene la deleción, podremos diferenciar el tamaño de los fragmentos amplificados por electroforesis capilar y, por ende, el sexo genético del individuo (**Fig. 13**).

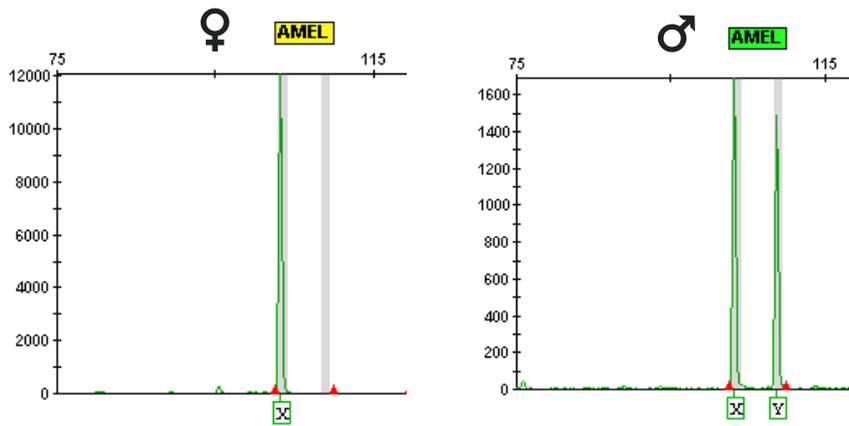


Figura 13. Amelogenina XX (sexo femenino) y XY (sexo masculino)

4.5 Análisis de fragmentos de ADN

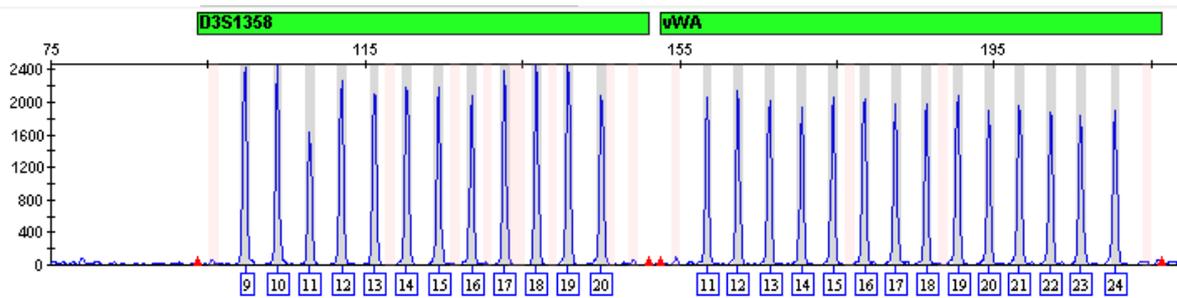
El análisis de fragmentos de ADN se lleva a cabo mediante el robot SeqStudio (**Fig. 14**), permite analizar 4 muestras a la vez y detectar 6 fluorocromos distintos [7]. Se basa en la separación de fragmentos de ADN mediante electroforesis capilar. Los fragmentos de ADN amplificados se separan por tamaño, aquellos con menor peso molecular llegan al cátodo antes, la excitación de los fluorocromos permitirá diferenciar aquellos fragmentos con el mismo peso molecular (pb).



Figura 14. Máquina Seqstudio para EC

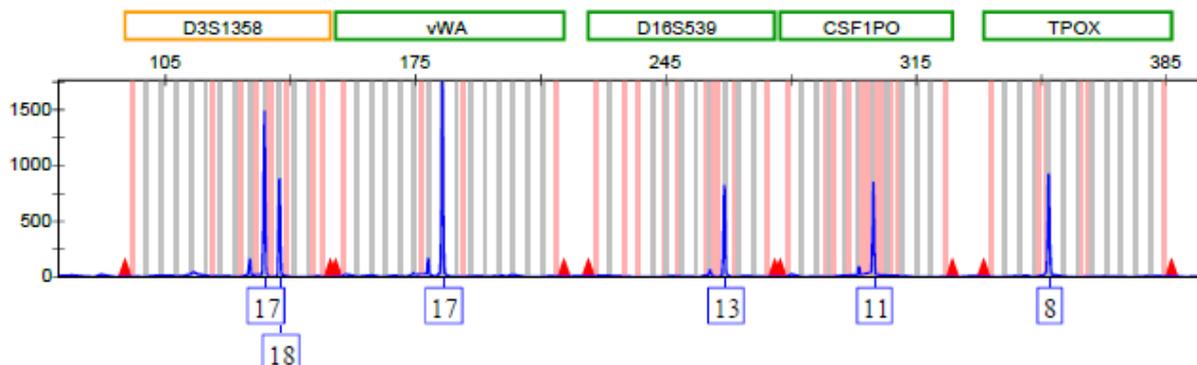
La longitud en pares de bases de cada fragmento se asigna en función de un marcador de peso molecular. Por otro lado, la asignación de los alelos se realiza mediante el *Allelic Ladder*, una mezcla artificial de los alelos más comunes presentes en la población para el conjunto de STR analizados (**Fig. 15**).

Figura 15. Allelic Ladder para los marcadores D3S1358 y vWA.



Finalmente, los datos obtenidos del perfil genético en forma de electroferograma (**Fig. 16**) con picos de fluorescencia quedan registrados en el sistema informático para su posterior análisis e interpretación.

Figura 16. Ejemplo electroferograma: el perfil genético muestra homocigosis para los marcadores vWA, D16S539, CSF1PO y TPOX. El individuo es heterocigoto para el macador D3S13568.



La técnica de electroforesis capilar supone un avance respecto a la visualización de los fragmentos amplificados mediante geles de acrilamida. Una mayor sensibilidad de los fluorocromos en EC permite detectar cantidades muy pequeñas de ADN, por el contrario en geles de acrilamida podría pasar desapercibido por el ojo humano. Por otro lado, la cuantificación del tamaño en pb de los fragmentos amplificados se determina de manera exacta gracias a la incorporación de un estándar interno, además los resultados se obtienen de manera automatizada facilitando así su análisis y su almacenaje.

4.6 Análisis bioinformático e interpretación: Familias

La implementación del software Familias se divide en cuatro vertientes:

- 1- Análisis y comparación de casos de paternidad previos con Familias, para 15 marcadores informados.

- 2- Análisis de casos de paternidad previos con Familias, para 23 marcadores informados.
- 3- Análisis de pruebas de paternidad con mutaciones.
- 4- Análisis de pruebas de parentesco biológico (hermandades).

4.6.1 Pruebas de paternidad y comparación de resultados (15 marcadores)

Esta primera vertiente tiene como objetivo probar el software Familias y comprobar que se obtienen los mismos resultados para el análisis de 15 marcadores con GestLab. Para ello se ha repetido el análisis de 50 pruebas de paternidad, realizadas e informadas previamente por el laboratorio, con el software Familias. Se pretende comparar los resultados obtenidos con los dos softwares con tal de validar los resultados del nuevo programa, comprobar que son fiables y automatizar el proceso bioinformático.

4.6.2 Prueba de paternidad con 23 marcadores informados

Una vez comprobada la funcionalidad del software Familias, se pretende mejorar la calidad del resultado de la prueba de paternidad informando más marcadores. Esta fase tiene como objetivo comparar e interpretar el valor de LR obtenido informado 15 y 23 marcadores.

4.6.3 Pruebas de paternidad con mutaciones

En ocasiones, los perfiles genéticos del supuesto padre y el hijo pueden presentar una única incompatibilidad genética, esta puede ser debida a mutaciones en la línea germinal del progenitor. Cuando esto sucede, el cálculo del valor LR se realiza sin tener en cuenta el genotipo del marcador en cuestión.

Este punto tiene como objetivo utilizar el programa Familias para calcular el valor LR teniendo en cuenta la tasa de mutación de los marcadores e interpretar los resultados. Para ello, se han utilizado las tasas de mutación de la base de datos *STRbase-NIST* y la opción de mutación simple que considera que todas las mutaciones ocurren con igual probabilidad, este modelo no tiene en cuenta la frecuencia del alelo ni el número de repeticiones (**Fig. 17**).

La tasa de mutación o *mutation rate* de un loci determinado viene dado por el número de mutaciones para dicho loci y el número de meiosis estudiadas, la probabilidad de mutación es mayor durante la división celular (replicación del DNA). Dado que las células espermáticas (espermatogonias) se dividen constantemente, prácticamente durante

toda la vida, el hombre tiene una tasa de mutación mayor respecto a la de la mujer. Por el contrario, las células germinales femeninas (ovogonias) experimentan menor número de divisiones y, por ende, su tasa de mutación es menor. Mientras que la espermatogénesis es un proceso continuo, la ovogénesis contiene periodos de latencia [13],[14]. Por este motivo, el programa Familias nos permite diferenciar la tasa de mutación en hombres y en mujeres.

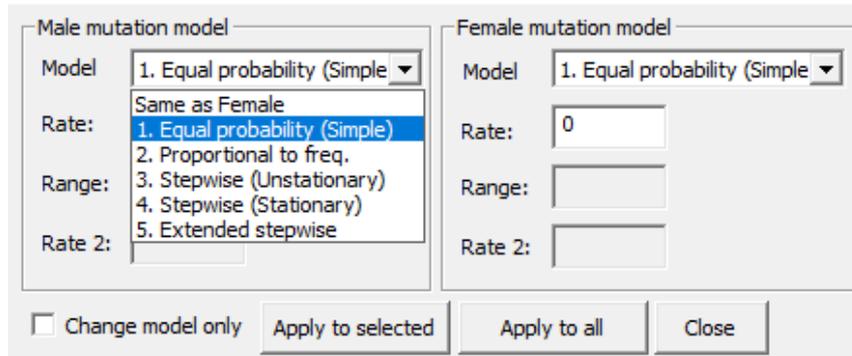


Figura 17. Existen 5 modelos de mutación en *Familias* y pueden diferenciarse en hombre y mujer.

Existen varios mecanismos de mutación, el más común se da por deslizamiento de las cadenas de DNA durante su replicación. Este fenómeno puede ocasionar deleciones o inserciones dependiendo de la hebra de DNA implicada. Cuando la cadena molde se desliza formando un *loop*, se origina como resultado una copia con menos repeticiones. Por el contrario, cuando es la cadena que se está sintetizando la que se desliza la DNA polimerasa origina una copia con más repeticiones (**Fig. 18**).^[10]

Cuando las mutaciones ocurren en la línea germinal del progenitor el hijo puede heredar el alelo mutado. Este tipo de mutaciones también pueden verse reflejados en forma de *stutter* en el electroferograma (**Fig. 19**).

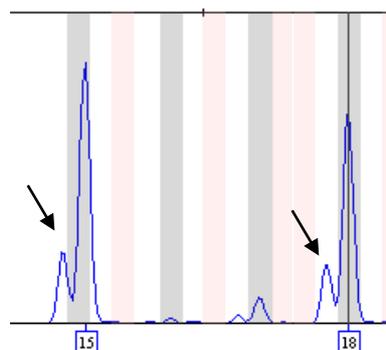
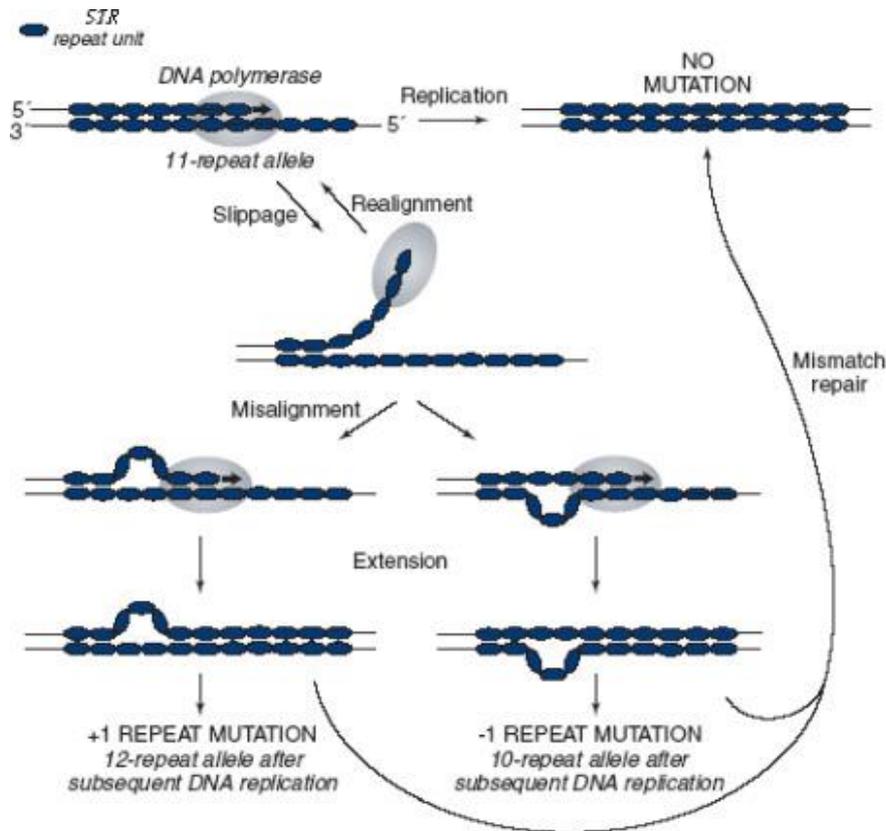


Figura 19. Picos *stutter* de los alelos 15/18.

Figura 18. Mecanismo de mutación por deslizamiento de la cadena de DNA durante su replicación.



4.6.4 Pruebas de parentesco

Las pruebas de parentesco biológico permiten establecer relaciones biológicas entre individuos (hermandad, maternidad...). Actualmente las pruebas de parentesco con STR autosómicos, STR del cromosoma Y o ADNmt son externalizadas a otra empresa. El objetivo de este punto ha sido probar la posibilidad de realizar hermandades con el programa Familias. Para ello, se han analizado tríos: Padre biológico - hijo biológico - supuesto hermano. Es decir, se han realizado hermandades entre dos personas aleatorias incluyendo al padre de uno de ellos. Se pretende analizar el LR obtenido en caso de no hermandad.

5. Resultados y discusión

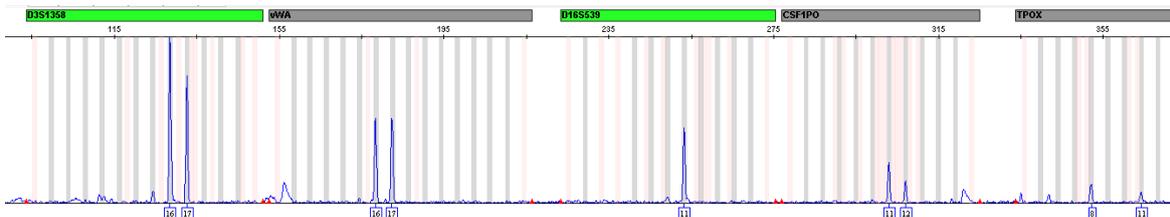
5.1 Amplificación diferencial

Tras analizar más de 50 electroferogramas es interesante remarcar que muchos de ellos presentan el mismo fenómeno, la altura de los picos disminuye de manera gradual: los picos son más altos (mayor intensidad) para los marcadores de menor tamaño. Por el

contrario, se observa una menor intensidad para los picos de los marcadores de mayor tamaño (**Fig. 20**)

Este fenómeno recibe el nombre de amplificación diferencial, en este caso ocurre por un exceso de concentración DNA en la PCR debido a la exhaustiva toma de la muestra. Dicho fenómeno también puede darse cuando la muestra está muy degradada, situación muy común en muestras de índole forense.

Figura 20. Fenómeno de amplificación diferencial, se obtiene un perfil de intensidades descendentes.



Generalmente, los marcadores de menor tamaño se amplifican mejor que los que tienen un mayor número de pares de bases. Por este motivo, observaremos un perfil de intensidades descendente. Dicho fenómeno se agudiza cuando existe mucha cantidad de ADN o cuando está degradado ya que los fragmentos más grandes serán más proclives a sufrir degradación que aquellos más pequeños.

5.2 Pruebas de paternidad y comparación de resultados

En total se han analizado 50 casos de paternidad, los valores de LR obtenidos con Familias son equiparables a los obtenidos en GestLab informando 15 marcadores. En la **Tabla 1** puede verse el resultado de 10 de las muestras, la totalidad de los resultados se muestra en el **Anexo III**. Los valores de LR obtenidos con Familias son equiparables a los valores de LR calculados por GestLab, las diferencias pueden deberse a las frecuencias alélicas utilizadas.

Tabla 1. Valor LR obtenido en GestLab y Familias

	LR GestLab	LR Familias
1	31.506.801	31.525.709
2	78.839	79.045
3	1.638.523	1.645.906
4	195.287	195.794
5	3.939	3.957

6	3.597	3.621
7	307.000	308.028
8	1.180.835	1.553.845
9	1.018	1.019
10	38.245	38.268

5.3 Prueba de paternidad con 23 marcadores informados

Una vez comprobado que *Familias* funciona correctamente, se han recalculado los valores de LR para 50 casos incluyendo los 23 STR del kit *VeryFiler Express*. La **Tabla 2** muestra el resultado de 10 de las muestras, la totalidad de los resultados pueden verse en el **Anexo III**. De acuerdo con lo esperado, el estudio de más marcadores aumenta el valor LR drásticamente en todos los casos. Se destacan las muestras 5, 6 y 9 con valores muy bajos de LR, al aumentar el número de marcadores analizados su valor es superior a 10.000, permitiendo así concluir la paternidad.

Tabla 2. Valor LR obtenido en GestLab y Familias para 15 y 23 marcadores

	15 STR		23 STR
	LR GestLab	LR Familias	LR Familias
1	31.506.801	31.525.709	674.625.278.900
2	78.839	79.045	8.082.611
3	1.638.523	1.645.906	27.744.238.590
4	195.287	195.794	413.273.477
5	3.939	3.957	75.629
6	3.597	3.621	1.807.247
7	307.000	308.028	9.555.688
8	1.180.835	1.553.845	2.935.870.018
9	1.018	1.019	1.268.975
10	38.245	38.268	1.320.520.926

5.4 Pruebas de paternidad con mutaciones

El estudio se ha realizado únicamente con 5 muestras de paternidades, de entre todos los informes del laboratorio de los últimos 3 años no se han encontrado más casos de mutaciones. En la **Tabla 3** se muestran los resultados, en la columna de “mutación” se especifica el marcador y el genotipo del padre y del hijo. Finalmente, se muestra el valor LR, para 23 marcadores informados, calculado sin tener en cuenta el marcador mutado y teniendo en cuenta su tasa de mutación.

Tabla 3. Resultado LR pruebas de paternidad con mutaciones

	Mutación	Tasa mutación (%)	IP (sin tasa mutación)	IP (con tasa mutación)
1	D12S391 P: 21/ <u>25</u> H: 18/ <u>24</u>	0,17	181.047.794	418.830
2	FGA P: <u>24</u> / <u>24</u> H: 19/ <u>22</u>	0,28	27.744.238.590	40.726.718
3	D7S820 P: 10/ <u>11</u> H: <u>12</u> / <u>12</u>	0,10	18.865.263	15.606
4	D3S1358 P: <u>17</u> / <u>18</u> H: 13/ <u>16</u>	0,12	145.997.704	6.126.740
5	vWA P: 16/ <u>17</u> H: <u>18</u> / <u>18</u>	0,17	78.881.421	128.054

De acuerdo con lo previsto, el cálculo del LR disminuye al tener en cuenta la tasa de mutación de cada marcador. Sin embargo, el valor sigue siendo superior a 10.000 y, por tanto, se sigue concluyendo una paternidad probada con probabilidades del 99.9%.

Aunque el número de muestras es bajo y poco representativo, observamos que los marcadores mutados corresponden a marcadores con mayores tasas de mutación respecto a marcadores como TH01 o TPOX con tasas de mutación del 0,01%.^[9]

5.5 Pruebas de parentesco

Se plantean 10 hermandades aleatorias con el trío: madre biológica – hijo biológico – supuesto hermano, las hipótesis se muestran en la **Fig. 21**. Se realiza el cálculo con el programa Familias para 23 marcadores.

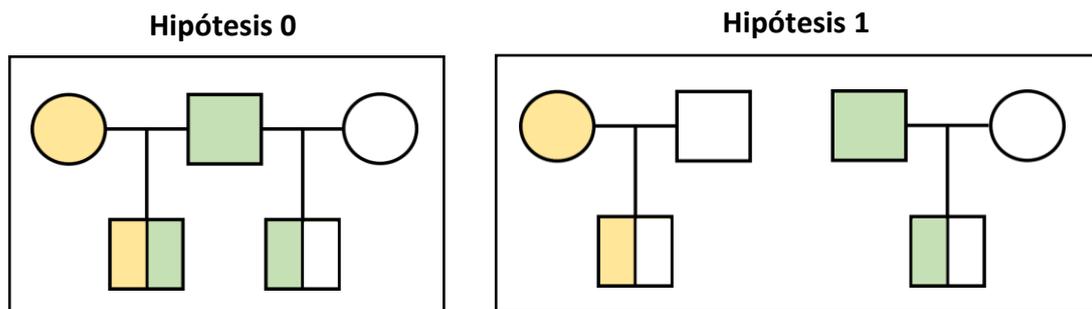


Figura 21. Pedigrís H0 y H1 para parentesco de media hermandad

Se obtienen LR muy bajos para todos los casos (**Tabla 4**) del orden de 10^{-5} a 10^{-1} , corroborando que será poco probable que un caso de no hermandad pueda dar un falso positivo ($LR > 10.000$).

6. Otras aplicaciones de los STRs

Los marcadores tipo STR (*Short Tandem Repeat*) también son útiles en el diagnóstico de enfermedades de origen genético como, por ejemplo, enfermedades por expansión de triplete: Cromosoma X Frágil, Distrofia miotónica, Ataxia de Friedrich, Ataxias espinocerebelosas, enfermedad de Huntington...

6.1 Enfermedad de Huntington

La enfermedad de Huntington es una afectación genética causada en el gen HTT ubicado en el cromosoma 4, el cual codifica para la proteína huntingtina expresada en el sistema nervioso central. Su herencia es autosómica recesiva, la mutación en uno de los alelos es suficiente para desarrollar la enfermedad.

La enfermedad se da por expansión del triplete $(CAG)_n$ que codifica para glutaminas en el extremo N-ter de la proteína. Las secuencias con más de 40 repeticiones implican una proteína demasiado grande que tiende a formar agregados neurotóxicos debido a la interacción hidrofóbica.^[15]

El análisis de STR mediante PCR y electroforesis capilar permite obtener el tamaño del fragmento, la división del tamaño del fragmento en pares de bases entre 3 (triplete) nos indica el número de repeticiones CAG del alelo (**Fig. 22**).

	IP Familias
1	1,20E-05
2	5,55E-02
3	4,25E-01
4	5,42E-03
5	3,63E-01
6	3,18E-03
7	9,12E-03
8	3,94E-02
9	4,80E-04
10	4,55E-04

Tabla 4. Valor LR hermandades

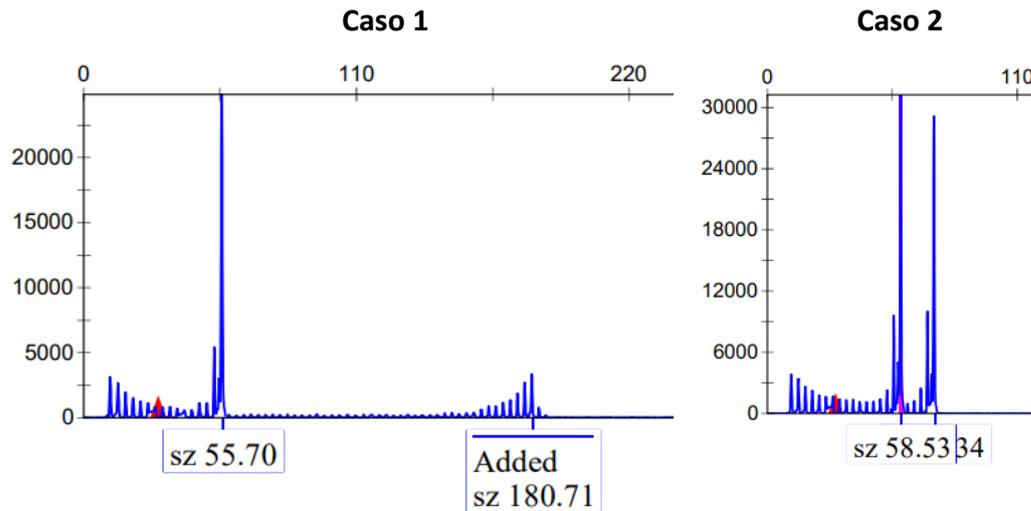


Figura 22. Caso 1. Paciente con Huntington heterocigoto con alelos de 18 y 60 repeticiones. **Caso 2.** Paciente sano con alelos de 19 y 21 repeticiones.

7. Conclusiones

La facilidad para extraer el ADN mediante el uso de kits comerciales y el estudio de los polimorfismos de tipo STR permiten realizar pruebas de paternidad, parentesco biológico y diagnosticar enfermedades de origen genético de manera sencilla. La posibilidad de integrar los datos obtenidos en programas bioinformáticos permite analizar un mayor número de datos, en este caso la implementación del software *Familias* ha supuesto la mejora de los procesos bioinformáticos del laboratorio contribuyendo a reducir el tiempo de trabajo y los posibles errores humanos. Además, ha permitido analizar un total de 23 marcadores incrementando la robustez de la valoración estadística de la prueba de paternidad, en algunos casos ha supuesto el cambio de un LR superior a 10.000.

Por otro lado, se han podido reevaluar los casos de paternidad con 1 exclusión incluyendo la tasa de mutación de los marcadores en cuestión y, de acuerdo con lo previsto se ha reducido el valor LR pero la probabilidad sigue siendo del 99.9%.

Finalmente, se ha comprobado que la posibilidad de obtener falsas hermandades en las pruebas de parentesco biológico mediante *Familias* es baja, aunque la presencia de un único progenitor dificulta el análisis. Por el momento, las pruebas de parentesco biológico deben seguir externalizándose, en un futuro podría plantearse el análisis de marcadores del cromosoma Y, cromosoma X y ADN mitocondrial.

8. Bibliografía

- [1] <https://laboratorioechevarne.com/>
- [2] Brown TA. Genomes. 2nd edition. Oxford: Wiley-Liss; 2002. Chapter 1, The Human Genome. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK21134/>
- [3] National Research Council (US) Committee on DNA Forensic Science: An Update. The Evaluation of Forensic DNA Evidence. Washington (DC): National Academies Press (US); 1996. 2, Genetic and Molecular Basis of DNA Typing. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK232609/>
- [4] <https://strbase.nist.gov/intro.htm>
- [5] Crespillo Márquez, M. C., Acón Ortego, I., Alonso Alonso, A., Amorim, A., Arenas, M., Carracedo, A., Fernández de Simón Loro, L., Fernández Oliva, C., Freire-Aradas, A. (2019). *Genética forense: del laboratorio a los tribunales*. Diaz de Santos. <http://www.ebooks7-24.com.sabidi.urv.cat/?il=10345>
- [6] <https://familias.no/es>
- [7] https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0018646_SeqStudioInstSW_UG.pdf
- [8] <https://www.thermofisher.com/es/es/home/industrial/forensics/human-identification/forensic-dna-analysis/pcr-amplification-forensic-dna-profiling/verifiler-express-pcr-amplification-kit.html>
- [9] <https://strbase.nist.gov/mutation.htm>
- [10] Fan H, Chu JY. A brief review of short tandem repeat mutation. *Genomics Proteomics Bioinformatics*. 2007 Feb;5(1):7-14. doi: 10.1016/S1672-0229(07)60009-6. PMID: 17572359; PMCID: PMC5054066.
- [11] <https://scielo.isciii.es/pdf/cmfn52/articulo2.pdf>
- [12] Sasaki, S & Shimokawa, Hitoyata. (1995). The amelogenin gene. *The International journal of developmental biology*. 39. 127-33.

[13] Sear, Rebecca. (2015). How We Do It: The Evolution and Future of Human Reproduction. *Population studies*. 69. 126-127. 10.1080/00324728.2014.969574.

[14] <https://www.science.org/content/article/males-mutate-more-females>

[15] <https://rarediseases.info.nih.gov/espanol/11993/enfermedad-de-huntington>

[16] Butler, J. M. (2006). Genetics and genomics of core short tandem repeat loci used in human identity testing. *J. Forensic Sci.* 51, 253–265. doi: 10.1111/j.1556-4029.2006.00046.x

progenitor biológico con respecto a un individuo tomado al azar de entre los de su misma población. El estudio estadístico se ha realizado teniendo en cuenta la frecuencia de los alelos en la población española (obtenidos de GHEP-ISFG).

STR	P	H	Compatibilidad
D8S1179	12/14	13/14	SI
D21S11	28/30	28/32	SI
D2S441	14.2/15	15/16.3	SI
D12S391	13/16	13/15	SI
D3S1358	15.3/16	15.3/17	SI
TH01	7/8.2	7/10	SI
D22S1045	11/16	12/16	SI
D16S539	9/10.2	9/13	SI
D2S1338	20/23	23/26	SI
D19S433	14.2/15	15/16	SI
VWA	14/16	14/18	SI
D1S1656	12/15	15/16	SI
D18S51	12/15	12/18	SI
D10S1248	9/10	9/12	SI
FGA	24/26	26/27.3	SI
AMELOGENINA	XY	XX	

CONCLUSIÓN

INCLUSIÓN

El individuo donante de la muestra identificada como "P" no puede ser excluido como padre biológico del individuo donante de la muestra identificada como "H".

La probabilidad de paternidad siguiendo los predicados verbales de Hummel corresponde a una PATERNIDAD PRÁCTICAMENTE PROBADA.

Con la técnica empleada se han analizado 15 marcadores STR y el marcador de género Amelogenina, formando el perfil genético que es único para cada individuo con una fiabilidad del 99,9999%. Se realiza el perfil genético del hijo y del supuesto padre y se comparan los marcadores obtenidos. Puesto que un hijo posee un alelo heredado de la madre y otro del padre biológico para cada marcador, se considera padre biológico cuando para todos los marcadores, uno de los dos alelos del hijo es compatible con uno de los alelos del supuesto padre. En este caso se calcula la probabilidad de paternidad, que es superior al 99,99% cuando comparten los 15 marcadores analizados.

Si hay tres o más alelos que no coinciden entre el hijo y el supuesto padre, se excluye la paternidad con una certeza del 100%. La probabilidad de paternidad será del 0%. Un solo marcador incompatible se incluirá en el cálculo de IP, teniendo en cuenta su tasa de mutación.

Anexo II

Consentimiento informado Prueba de Paternidad / Maternidad



FUNDAMENTOS

La prueba de paternidad / maternidad se basa en el análisis del material genético (ADN) de las personas. Solamente regiones de ADN repetitivo no codificante (STRs), que presentan gran variabilidad entre los individuos de la población, son analizadas. De estas secuencias no se puede deducir si la persona que dona la muestra tiene predisposición a sufrir una enfermedad o presenta una determinada peculiaridad física. Estas secuencias no proporcionan este tipo de información. Sin embargo, sí se puede comprobar una relación de paternidad / maternidad.

La información obtenida del análisis de 15 marcadores genéticos STRs, constituye un código genético identificador o perfil genético. La comparación del perfil genético entre las muestras analizadas permite establecer la relación de parentesco. Finalmente se elabora un informe de paternidad / maternidad que describe toda la información obtenida de las muestras facilitadas.

FORMULARIO DE CONSENTIMIENTO INFORMADO

- He leído este Consentimiento Informado y he comprendido la información.
- Los datos facilitados en este documento son correctos.
- Tengo constancia de que la presente prueba no tiene valor legal sobre las relaciones de filiación preexistentes, ni sobre la atribución de la patria potestad o cualquier otra de las filiaciones derivadas.
- Realizo la toma de muestra personalmente bajo mi responsabilidad y sólo yo respondo por la identidad de las muestras que entrego al laboratorio.
- Cuento con el permiso o capacidad legal para la toma de muestra de las personas implicadas en la prueba. Si un menor de edad identificado participa en la prueba, el tutor legal de este menor, si lo solicita, podrá tener acceso al resultado de la misma aunque no sea la misma persona que ha solicitado la determinación.
- Comprendo que el laboratorio podría tener dificultades en analizar una muestra y que una segunda muestra podría ser requerida.
- Mi consentimiento es completamente voluntario.
- Autorizo a Laboratorio Echevarne a conservar mis datos personales por tiempo indefinido. No obstante, siempre podré hacer uso de mi derecho de cancelación de la presente autorización.

Nombre del/la paciente:

DNI:

Parentesco: Padre Madre Hijo Hija

Dirección completa:

e-mail:

Teléfono:

Fecha:

Firma:

Nombre del representante legal:

Fecha:

Firma:

Documento de aplicación: Manual Fase Preanalítica Cap. 11

PERSONA RESPONSABLE DE LA IDENTIFICACIÓN DE LAS MUESTRAS

Nombre: _____ Apellidos: _____

DNI: _____ Centro: _____

El abajo firmante DECLARA que ha identificado correctamente a cada individuo con su DNI y que ha custodiado las muestras hasta el momento de su almacenaje.

Fecha y Hora: Firma: 

LISTADO DE IDENTIFICACIÓN DE MUESTRAS

Tipo de muestra	Nombre y apellidos	DNI o pasaporte*	Código

* Se adjunta fotocopia del DNI o pasaporte de las personas implicadas en este estudio o, en el caso de menores de edad, fotocopia del libro de familia o una fotografía reciente.

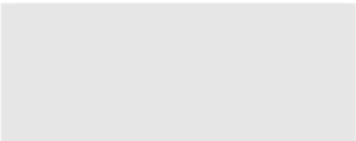
Toma de muestra

Responsable: _____ DNI: _____

Centro: _____ Hora: _____ Fecha de toma: _____

Almacenaje de las muestras

Centro: _____ Hora: _____ Fecha de entrada: _____

Realizado por: Firma: 

Condiciones de almacenaje: Temperatura ambiente Refrigeración Congelación

Envío de las muestras

Código Bolsa custodia: _____

Medio de transporte: _____ Hora: _____ Fecha de envío: _____

Remitido por: _____ Firma: 

Condiciones de envío: Temperatura ambiente Refrigeración Congelación

Anexo III

	LR GestLab 15	LR Familias 15	LR Familias 23
1	31,506,801.00	31,525,709.78	674,625,278,900.00
2	78,839.00	79,045.03	8,082,611.82
3	1,638,523.00	1,645,906.70	27,744,238,590.00
4	392,292.00	393,313.40	59,461,994.47
5	22,593,525.00	29,807,507.23	692,731,547,000.00
6	195,287.00	195,794.99	413,273,477.50
7	823,469.00	891,284.75	199,782,991.60
8	3,939.00	3,957.42	75,629.99
9	3,597.00	3,621.72	1,807,247.41
10	7,535,536.00	7,825,452.00	672,757,774,700.00
11	108,041.00	109,105.00	667,597,658.80
12	159,484.00	159,579.71	146,290,782.50
13	307,000.00	308,028.08	9,555,688.23
14	640,315.00	640,699.34	69,454,993.40
15	1,018.00	1,019.01	1,268,975.61
16	1,180,835.00	1,553,845.76	2,935,870,018.00
17	33,252.00	33,272.15	10,385,157.19
18	2,221,775.00	2,240,946.93	18,865,263.66
19	190,138.00	190,233.10	181,047,794.00
20	95,869.00	96,155.00	95,955.95
21	109,573,811.00	129,574,036.50	2,004,079,168.00
22	72,467.00	72,510.74	198,634,164.00
23	65,820.00	65,992.01	4,366,040.88
24	38,245.00	38,268.61	1,320,520,926.00
25	1,226,256.00	1,226,991.74	55,151,663.36
26	4,604,676.00	4,612,660.00	5,499,406,083.00
27	86,170.00	86,221.92	11,392,179.97
28	81,609.00	81,821.45	9,312,532.53
29	5,420.00	5,451.00	395,016,788.00
30	1,753,308.00	1,974,308.00	187,679,337.50
31	39,927.00	40,111.04	50,998,996.53
32	83,036.00	83,419.11	651,665,141.40
33	9,705,661.00	9,750,370.57	169,112,210,800.00
34	740,185.00	741,025.00	145,997,704.50
35	4,657.00	4,659.99	6,781,437.97
36	178,500.00	178,607.15	16,331,995.12
37	385,999.00	387,264.00	78,881,421.66
38	1,852,712.00	1,853,078.00	6,547,900,611.00
39	3,631.00	3,812.30	4,361,839.52
40	841,023.00	841,527.24	200,359,019.70
41	37,940.00	38,805.89	1,337,838,380.00
42	437,474.00	459,239.21	4,111,067,877.00

43	230,871.00	238,037.34	48,438,044.71
44	242,366.00	249,889.21	11,070,045.37
45	44,322.00	45,697.46	14,438,670.97
46	801,704.00	826,590.81	27,852,360.80
47	1,271,381.00	1,334,643.48	72,539,396,360.00
48	515,933.00	541,601.11	339,117,698.50
49	2,444,004.00	2,565,593.52	33,162,424,090.00
50	189,115.30	194,515.30	220,931,163.80