



Informe y recomendaciones de la CTP sobre las nuevas tecnologías de análisis genético y nuevos marcadores de ADN de origen biogeográfico y de rasgos fenotípicos externos

Introducción

Estamos asistiendo a una nueva revolución tecnológica en el campo de la Genética Forense. Se trata de la implementación creciente en los laboratorios (tanto públicos como privados) de la metodología de la secuenciación masiva en paralelo (*MPS, Massive Parallel Sequencing*) [1-5]. Actualmente, hay un número creciente de centros de genética forense públicos y privados que están investigando y comenzando a implementar esta nueva tecnología para: (1) el análisis de marcadores de ADN forense "clásicos" (es decir, ADN de repeticiones en tándem cortas (STR) y región control del ADN mitocondrial) utilizadas en todo el mundo en la casuística forense, así como (2) para estudiar otros marcadores de ADN como son los SNP (polimorfismos de un solo nucleótido. Del inglés: *Single Nucleotide Polymorphism, SNP*) así como los INDEL (del inglés *Insertion / Deletion, INDEL*, que son pequeñas deleciones e inserciones de nucleótidos).

Los SNP / INDEL pueden usarse tanto para estudios de identificación individual forense (SNP de identidad), como en estudios de ancestralidad o ascendencia biogeográfica (SNP de ancestralidad), así como para determinar algunas características fenotípicas (color de piel, ojos y color de pelo) (SNP de rasgos fenotípicos) [6]. El empleo de esta nueva tecnología constituye una herramienta de indudable utilidad en la investigación criminal así como en la identificación de cadáveres y personas desaparecidas. Uno de los proyectos Europeos actuales más ambicioso en el campo forense es el proyecto VISAGE (<http://www.visage-h2020.eu>) cuyo objetivo principal es el desarrollo y la validación de sistemas de secuenciación masiva para el estudio de marcadores de origen biogeográfico y de rasgos fenotípicos para su uso en el campo forense.

Muchos de estos nuevos marcadores de ADN están localizados en regiones reguladoras de los genes o en los propios genes (regiones codificantes del genoma) en contraposición a otros marcadores de ADN forense (STR, SNP de identificación y región control del ADN mitocondrial) que se encuentran en regiones no codificantes del genoma. Esto plantea un nuevo reto para su aplicación en el campo forense ya que en la mayoría de las legislaciones actuales europeas este uso no se encuentra regulado siendo normalmente leyes (como en España) de bases de datos.

Otra diferencia importante entre estos nuevos marcadores de ADN (de ancestralidad y rasgos fenotípicos) y los marcadores clásicos de ADN (STR y SNP de identificación) en genética forense es que los primeros tienen solo un valor predictivo o de inferencia con valores de probabilidad (70-90%) muy alejados de los acostumbrados a obtener cuando se produce una coincidencia en el análisis genético comparativo de perfiles STR



y/o SNP de identificación. Por este motivo se usan exclusivamente en la actualidad como un instrumento de investigación

En el presente documento se identifican los marcadores de ADN y los métodos validados en genética forense para realizar inferencias de origen biogeográfico y de apariencia fenotípica y se enumeran una serie de recomendaciones con respecto a su uso futuro en la casuística forense en nuestro país.

Polimorfismos de un Único Nucleótido (SNP): generalidades

Actualmente en los laboratorios de Genética Forense la técnica más utilizada es la electroforesis capilar para la detección de STR autosómicos, STR de cromosoma Y o para la secuenciación de la región control del ADN mitocondrial. Sin embargo, cada día son más ampliamente utilizados otro tipo de marcadores genéticos, los llamados SNPs (*single nucleotide polymorphisms* o polimorfismos de un único nucleótido) que son sustituciones, inserciones o deleciones de un nucleótido y que presentan una serie de ventajas sobre los tradicionales STRs utilizados desde la década de los 90:

- Gran abundancia en el genoma
- Baja tasa de mutación (10^{-8} por nucleótido [7])
- Estabilidad de herencia para el análisis de parentesco
- Tamaño corto de amplicón que permite una potencial alta tasa de éxito en muestras con ADN degradado
- Ausencia de bandas “*stutter*” lo que facilita la interpretación de perfiles
- Facilidad de genotipado

Además, existen diferentes tipos de SNP [8], lo que les convierte en una herramienta valiosa para ser utilizada en el campo de la Genética Forense:

- SNP informativos de identidad (diferenciación de individuos)
- SNP informativos del linaje (tanto en el genoma mitocondrial como en los cromosomas sexuales)
- SNP informativos de ancestralidad o ascendencia biogeográfica (estimación del origen étnico)
- SNP informativos del fenotipo (que predicen algunos rasgos físicos como color de ojos, color de cabello, color de piel, etc.)

La Secuenciación Masiva en Paralelo en Genética Forense



En la actualidad existen distintas tecnologías con diferentes métodos de construcción de bibliotecas, química de la secuenciación, ensamblaje, etc. lo que origina la existencia de diferentes plataformas de análisis (ver Anexo 1).

Las ventajas fundamentales de la secuenciación masiva en paralelo son las siguientes:

- Permite el análisis simultáneo de miles de regiones o marcadores de ADN (potencialmente el genoma) incluyendo cualquier tipo de marcadores forenses de ADN (STR, SNP, INDEL, o por ejemplo el genoma completo del ADN mitocondrial). Esto posibilita la integración en un solo flujo de trabajo de diferentes marcadores de ADN de interés en Genética Forense, y en el futuro próximo permitirá el abaratamiento de costes y un menor tiempo de realización de experimentos.
- Permite obtener un mayor poder de discriminación que la electroforesis convencional al posibilitar análisis de un mayor número de marcadores genéticos y mayor información en el caso de los análisis de STRs (secuenciación de la unidad de repetición, regiones adyacentes, etc.).
- Es una técnica de alta sensibilidad (con un límite de detección en el rango de los picogramos de ADN genómico) pudiendo ser muy útil en el análisis de muestras con muy bajo contenido de ADN.
- Permite el desarrollo de amplicones de pequeño tamaño con lo que esta tecnología puede mostrarse muy efectiva en el análisis de muestras forenses con ADN degradado.
- Constituye una herramienta de indudable utilidad en el análisis de perfiles mezclas. Lo cual supone un avance importante en el abordaje de un tipo de perfiles, que comportan una extraordinaria complejidad en el campo de la genética forense.

Todas estas ventajas hacen que en la actualidad exista por parte de la comunidad internacional de genetistas forenses un gran interés en el desarrollo y la implementación de esta nueva tecnología de análisis de ADN, existiendo diversos proyectos internacionales para la validación, el desarrollo de estándares y de bases de datos poblacionales que permitan su aplicación en el campo forense:

- DNASEQEX [9]
- SeqforSTRs - Sequencing of forensic STRs [10]
- STRSeq [11]
- Empowering forensic genetic DNA databases for the interpretation of next generation sequencing profiles (DNA.bases) [12]



- The VISible Attributes Through GENomics (VISAGE)

VISAGE (<http://www.visage-h2020.eu/>) es uno de los proyectos europeos actuales más ambicioso en el campo forense, entre cuyos objetivos principales se encuentran:

- El desarrollo y la validación de sistemas de secuenciación masiva para el estudio de marcadores de origen biogeográfico y de rasgos fenotípicos para su uso en el campo forense.
- El diseño de un marco de interpretación que incluya un prototipo de software para la consideración estadística combinada de los marcadores anteriormente citados.
- El establecimiento de recomendaciones para su futura implementación teniendo en consideración los diferentes marcos sociales, éticos y legales de los diferentes países europeos.
- El contacto con las empresas líderes en el campo forense de cara a conseguir productos comerciales para una futura aplicación de rutina de este tipo de marcadores y tecnología en los estados miembros de la Unión Europea.

ADN y ancestralidad

Podemos definir la inferencia o predicción biogeográfica de ancestralidad (“ancestralidad”) como la estimación del origen geográfico de los ancestros biológicos de una persona basada en el análisis de ADN. Las pruebas de inferencia de ascendencia con ADN se basan en el conocimiento de la existencia de una cierta variación genómica entre las poblaciones que se originaron en diferentes ubicaciones geográficas de todo el mundo. En particular, existen ciertos marcadores genéticos que son variables, presentando dichas variables una mayor prevalencia en algunas poblaciones que en otras.

Hasta hace unos años, los marcadores de linaje nos ofrecían ciertas pistas sobre el origen geográfico de los ancestros paternos y maternos de una persona. Sin embargo, la herencia especial de estos marcadores (uniparentales y heredados sin modificaciones a lo largo de muchas generaciones, salvo mutación) imposibilita rastrear la ancestralidad completa de un individuo si pertenece a una población mezclada (con diferentes componentes ancestrales). Para predecir correctamente la ancestralidad es necesario por tanto usar marcadores que capten la historia completa del árbol genealógico de un individuo, como los SNPs localizados en el nuclear ADN autosómico (que contemplan la herencia de todos los individuos del árbol genealógico, y no sólo de los que transmitieron el ADN mitocondrial o el cromosoma Y).



Estos sistemas permiten realizar una estima probabilística del origen o proporción de ancestralidad (en individuos con diversas contribuciones ancestrales) en base a la distribución de frecuencias de los SNPs estudiados en diversas poblaciones humanas (fundamentalmente África, América, Europa, Oceanía, este de Asia, sur de Asia y suroeste de Asia).

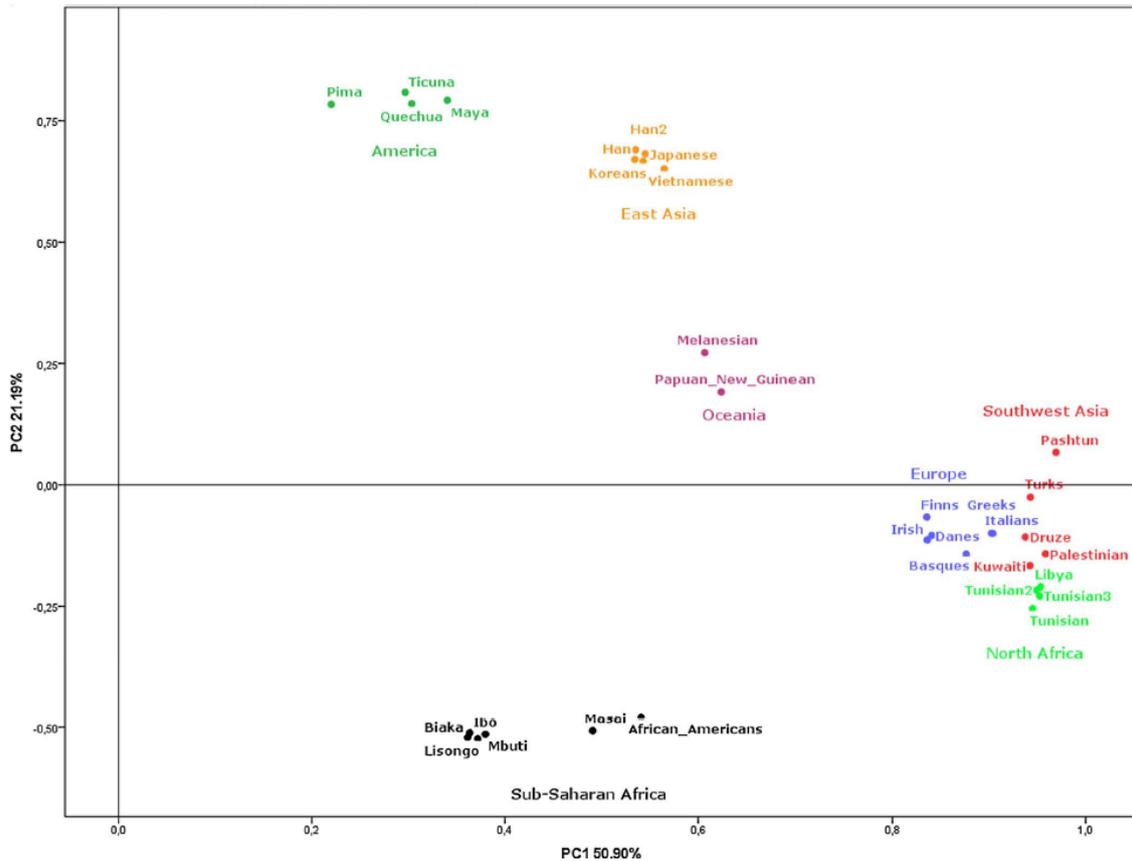


Gráfico tomado de [39] a partir de 32 poblaciones y usando los 165 SNPs del Precision ID Ancestry panel (ThermoFisher Scientific).

Obviamente y en la medida en que tengamos mayores datos poblacionales y mayor conocimiento del genoma, se podrán ir refinando estos grandes grupos poblacionales. Así, por ejemplo, ya se han descrito métodos para distinguir a los japoneses del resto de poblaciones del este asiático, usando variantes raras específicas de población [13].

Además, hemos de mencionar la progresiva comercialización de “tests genéticos de ancestralidad” por parte de compañías que ofrecen a los consumidores reconstruir sus historias familiares y determinar los orígenes geográficos de sus antepasados [14-15]:

- AncestryDNA <https://www.ancestry.com/dna/>
- 23andMe, <https://www.23andme.com/dna-ancestry/>
- FamilyTreeDNA <https://www.familytreedna.com/>



- MyHeritage (<https://www.myheritage.es/dna>)

Muchas de estas pruebas genéticas están basadas en el estudio de cientos de miles de SNPs en todo el genoma y, por lo tanto, ofrecen un poder de discriminación enorme en estudios de parentesco.

Cuestiones como los aspectos bioéticos, la privacidad de los datos genéticos, la comercialización de la información genética, etc. son cuestiones de gran importancia que no van a ser tratadas en el presente documento, aunque si hemos de poner de manifiesto la enorme potencialidad de estas bases de datos privadas como herramientas para la resolución de hechos delictivos tal y como ha sucedido en famoso caso del “*Golden State killer*” resuelto gracias al uso de la web “*open-source*” GEDmatch (<https://www.gedmatch.com/>) donde los usuarios de los kit comerciales anteriormente citados pueden volcar sus datos para contactar con posibles familiares y que ha sido el origen de una nueva disciplina científica denominada “Genealogía Forense” [16-17]. Sin embargo, no todas las muestras forenses son susceptibles de análisis por estos métodos, que requieren cierta cantidad de ADN que no se encuentre mezclado con ADN de otros individuos para poder aplicarse.

ADN y rasgos fenotípicos (*Forensic DNA Phenotyping -FDP- of External Visible Characteristics -EVC-*)

Algunos marcadores de ADN (SNPs e INDELS) ubicados en las regiones reguladoras o codificantes de genoma pueden determinar la distinta expresión de los genes o producir sustituciones de aminoácidos, alterando las propiedades funcionales de la proteína traducida y, por lo tanto, expresándose en distintos fenotipos, algunos de las cuales son las características visibles del individuo.

Existen diversos estudios que ya han evaluado la existencia de polimorfismos asociados con el color de la piel [6], color del pelo [6], el color de los ojos [6], modelos de calvicie masculina [18], el tipo/forma del cabello de la cabeza [19-20], la edad [21], la morfología facial [22-23], la altura [24], plegamiento del lóbulo de la oreja [25], encanecimiento del cabello [26], etc.

Podríamos definir el fenotipado de ADN forense como aquella herramienta investigativa mediante la cual obtenemos información fenotípica de un individuo mediante el estudio de los rasgos fenotípicos nombrados anteriormente a partir de evidencias biológicas del mismo, normalmente abandonadas en la escena del delito. Esta información puede ser de gran utilidad en la investigación de delitos con sospechosos no identificados, permitiendo así reducir el potencial número de los mismos gracias a la información fenotípica obtenida.



Como vimos anteriormente, ninguna de estas pruebas presenta una precisión absoluta, se trata más bien de un análisis predictivo o de inferencia con valores de probabilidad (70-90%) muy alejados de los acostumbrados a obtener cuando se produce una coincidencia en el análisis genético comparativo de perfiles STR. Sirva como ejemplo la siguiente tabla obtenida a partir de los 41 SNP utilizados en el sistema HIRISPLEX-S (modificada a partir de [6]) para la caracterización del color de pelo, piel y ojos (<https://hirisplex.erasmusmc.nl/>).

Precisión de la estimación (máximo de 1) a nivel poblacional	
Color de ojos azul	0'94
Color de ojos intermedio	0'74
Color de ojos castaño	0'95
Color de pelo rubio	0'81
Color de pelo castaño	0'74
Color de pelo rojo	0'93
Color de pelo negro	0'86
Tono color cabello	0'91
Color de piel muy pálida	0'83
Color de piel pálida	0'76
Color de piel intermedia	0'78
Color de piel oscura	0'98
Color de piel oscura-negra	0'99

Por este motivo se usa exclusivamente como un instrumento de investigación, para reducir el número de sospechosos potenciales en aquellos casos en los que el grupo de sospechosos es muy amplio y han fallado otros métodos de investigación (por ejemplo, tras no obtener ninguna coincidencia después de una búsqueda en la base de datos nacional forense). Es decir, la predicción de características físicas a partir de ADN no se utilizará como pruebas finales ante un tribunal, sino como una ayuda para la policía de cara a orientar sus investigaciones.

Predecir otros rasgos externamente visibles, como los rasgos faciales para obtener un retrato robot, será aún más difícil y, en la actualidad, todavía imposible, porque si bien todos los rasgos de una persona tienen una base genética, los rasgos faciales están determinados por un gran número de genes que siguen siendo desconocidos y, además, hemos de considerar la posible influencia de factores externos (nutrición, hábitos de vida, procesos habidos durante la gestación...). Por tanto, sería aventurado esperar soluciones mágicas provenientes del campo de la Genética, aunque es de esperar un gran aumento de la investigación de marcadores que determinen características físicas visibles.

Regiones codificantes y no codificantes del ADN



Las regiones codificantes de ADN se describen como las partes de ADN que codifican proteínas y que, por lo tanto, pueden proporcionar información sobre el fenotipo de un individuo (es decir, sus características observables). Por contraposición, los marcadores clásicos de ADN (marcadores STRs) utilizados en Genética Forense se localizan generalmente en regiones no-codificantes del genoma.

De forma general, se tiene la creencia de que la información sobre el ADN no codificante de un individuo no proporciona información sobre las características externas visibles de una persona, ni aporta información sobre la posible susceptibilidad individual o familiar a padecer una enfermedad de base genética y, por lo tanto, su aplicación podría considerarse menos problemática desde el punto de vista ético. Sin embargo, en los últimos tiempos el desarrollo de la medicina predictiva (Alzheimer, cáncer de mama, cáncer de colon, etc.) está cobrando una enorme importancia y está basada en SNPs, obtenidos de estudios de asociación de genomas completos (GWAS, por sus siglas en inglés), generalmente no codificantes.

En algunas legislaciones, la distinción entre ADN codificante y no codificador es clave en la regulación del uso del ADN con fines forenses, que a menudo limita los análisis legítimos del ADN a las regiones no codificantes. Así, en la legislación española (Ley Orgánica 10/2007, de 8 de octubre, reguladora de la base de datos policial sobre identificadores obtenidos a partir del ADN [27]) cita en su preámbulo *“Esta regulación contiene una salvaguarda muy especial, que resulta fundamental para eliminar toda vulneración del derecho a la intimidad, puesto que sólo podrán ser inscritos aquellos perfiles de ADN que sean reveladores, exclusivamente, de la identidad del sujeto -la misma que ofrece una huella dactilar- y del sexo, pero, en ningún caso, los de naturaleza codificante que permitan revelar cualquier otro dato o característica genética”* y *“Sólo podrán inscribirse en la base de datos policial regulada en esta Ley los identificadores obtenidos a partir del ADN, en el marco de una investigación criminal, que proporcionen, exclusivamente, información genética reveladora de la identidad de la persona y de su sexo”* en su artículo cuarto.

Sin embargo, la investigación llevada a cabo durante la última década ha demostrado claramente que en la parte no codificante del ADN puede haber regiones reguladoras que permiten activar o desactivar genes y, por lo tanto, tienen una gran importancia para la expresión de los mismos. La evidencia científica ha demostrado que una proporción considerable de la variación de la función del gen es proporcionada por elementos reguladores del ADN, que pueden estar dentro de los genes, cerca de los genes o incluso distantes a los genes, y no solo por las variantes de ADN codificantes de proteínas dentro de los genes como se había supuesto en el pasado.

Además, hay casos en los que un marcador genético de ADN no codificante específico puede estar muy cerca de un gen específico, y el análisis del marcador genético específico puede informarnos sobre la región codificadora por el hecho de que el



marcador y la región codificadora respectiva están muy cerca (ligados) y se heredan de forma conjunta. De esta forma una variante STR o SNP específica en una región no codificante puede ser predictiva de una situación patológica producida por una mutación en un gen que se encuentre ligado (desequilibrio de ligamiento).

Por todo lo anterior, en la actualidad, no hay una distinción tan clara entre el ADN codificante y el ADN no codificante en términos de la posible información suplementaria (no relacionada con el propósito de la identificación genética) que se puede obtener de su estudio. Por lo tanto, ambos tipos de marcadores de ADN necesitan una regulación legal en la que se establezcan los fines, la proporcionalidad y los límites de su uso.

Se ha de tener en cuenta, que parte del ADN codificante proporciona en realidad información que no invade la privacidad de las personas (color de ojos, color de pelo, etc.), y por tanto no se trata de información genética sensible que vulnere los derechos fundamentales del individuo al ser estudiada.

Recomendaciones para el uso de los nuevos marcadores y nuevas tecnologías de análisis del ADN

- Los marcadores de ancestralidad y fenotipo **son herramientas de investigación predictiva que han sido avalados por la comunidad científica internacional**. Su aplicación en la casuística forense ha permitido la resolución de casos forenses antiguos sin resolver. Como ejemplo recordar el caso de agresión sexual y asesinato de Eva Blanco en Algete (Madrid) que fue resuelto 18 años después gracias al uso de marcadores de ADN de ancestralidad por el Instituto de Ciencias Forenses de la Universidad de Santiago de Compostela
- Es necesario que cada laboratorio proceda a realizar los **estudios de validación interna** previos al uso de estos marcadores de ADN en la casuística forense. Sería también conveniente que dichos análisis estuvieran **acreditados bajo la norma ISO 17.025** tal y como ocurre, en la actualidad en nuestra legislación, con los marcadores genéticos de uso rutinario en investigación criminal.
- Se recomienda a los organizadores de ejercicios de suficiencia de polimorfismos de ADN en el campo forense el **desarrollo de ejercicios de suficiencia específicos que incluyan este tipo de marcadores de ADN** para que los laboratorios puedan tener un sistema de intercomparación de resultados y les permita chequear los sistemas y estrategias de análisis desarrolladas.
-



- **Necesidad de regulación legal.** Estos nuevos marcadores de ADN (SNP de ancestralidad y fenotípicos) pueden ser un valioso instrumento en la averiguación de indicios criminales, siempre que se hayan agotado otras vías de investigación y que esta herramienta investigativa sea utilizada con las siguientes garantías para que la afectación de derechos del individuo sea mínima:
 - Es solo una herramienta de investigación y no se debe utilizar como prueba concluyente de identificación
 - Se ha de utilizar solo en delitos graves
 - Se necesita una autorización expresa (judicial o fiscal)
 - Se usará para indicios que no han dado coincidencias en la base de datos y para los casos en los que se han agotado todas las líneas de investigación
- Se recomienda **la difusión de estas recomendaciones**, así como otras publicaciones de divulgación científica sobre la aplicación forense de los nuevos marcadores de ADN y las nuevas tecnologías de análisis **entre los profesionales del mundo del derecho** (jueces, fiscales, abogados, etc.)

Referencias bibliográficas

1. Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol 26: 1135-1145 (2008).
2. Mardis ER. The impact of next-generation sequencing technology on genetics. Trends Genet 24: 133-141 (2008).
3. Tucker T, Marra M, Friedman JM. Massively Parallel Sequencing: The Next Big Thing in Genetic Medicine. Am J Hum Genet 85: 142-154 (2009).
4. Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet 11: 31-46 (2010).
5. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. Genomics 107: 1-8 (2016).
6. Chaitanya L, Breslin K, Zuñiga S, Wirken L, Pośpiech E, Kukla-Bartoszek M, et al. The HirisPlex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation. Forensic Sci Int Genet 35: 123-135 (2018).
7. Nachman MW, Crowell SL. Estimate of the mutation rate per nucleotide in humans. Genetics 156: 297-304 (2000).



8. Budowle B, Churchill JD, King JL. The Next State-of-the-Art Forensic Genetics Technology: Massively Parallel Sequencing, en: Amorim A, Budowle B (Eds.), Handbook of Forensic Genetics. Biodiversity and Heredity in Civil and Criminal Investigations, World Scientific Publishing Europe Ltd., London, pp. 249-292 (2017).
9. <https://www.researchgate.net/project/DNASEQEX>
10. <https://www.researchgate.net/project/SeqforSTRs-Sequencing-of-forensic-STRs>
11. <https://www.ncbi.nlm.nih.gov/bioproject/380127>
12. <https://www.researchgate.net/project/Empowering-forensic-genetic-DNA-databases-for-the-interpretation-of-next-generation-sequencing-profiles-DNAbases>
13. Yuasa I, Akane A, Yamamoto T, Matsusue A, Endoh M, Nakagawa M, et al. Japaneseplex: A forensic SNP assay for identification of Japanese people using Japanese-specific alleles. Leg Med (Tokyo) 33: 17-22 (2018).
14. Bolnick DA, Fullwiley D, Duster T, Cooper RS, Fujimura JH, Kahn J, et al. The science and business of genetic ancestry testing. Science 318: 399-400 (2007).
15. Bandelt HJ, Yao YG, Richards MB, Salas A. The brave new era of human genetic testing. Bioessays 30: 1246-1251 (2008).
16. Phillips, C. The Golden State Killer investigation and the nascent field of forensic genealogy. Forensic Sci Int Genet 36: 186-188 (2018).
17. Syndercombe Court D. Forensic genealogy: Some serious concerns. Forensic Sci Int Genet 36: 203-204 (2018).
18. Hagenshaars SP, Hill WD, Harris SE, Ritchie SJ, Davies G, Liewald DC, et al. Genetic prediction of male pattern baldness. PLoS Genet 13: e1006594 (2017).
19. Liu F, Chen Y, Zhu G, Hysi PG, Wu S, Adhikari K, et al. Meta-analysis of genome-wide association studies identifies 8 novel loci involved in shape variation of human head hair. Hum Mol Genet 27: 559-575 (2018).
20. Pośpiech E, Chen Y, Kukla-Bartoszek M, Breslin K, Aliferi A, Andersen JD, et al. Towards broadening Forensic DNA Phenotyping beyond pigmentation: Improving the prediction of head hair shape from DNA. Forensic Sci Int Genet 37: 241-251 (2018).



21. Zbiec-Piekarska R, Spolnicka M, Kupiec T, Parys-Proszek A, Makowska Z, Paleczka A, et al. Development of a forensically useful age prediction method based on DNA methylation analysis. *Forensic Sci Int Genet* 17: 173-179 (2015).
22. Cha S, Lim JE, Park AY, Do JH, Lee SW, Shin C, et al. Identification of five novel genetic loci related to facial morphology by genome-wide association studies. *BMC Genomics* 19: 481 (2018).
23. Crouch DJM, Winney B, Koppen WP, Christmas WJ, Hutnik K, Day T, et al. Genetics of the human face: Identification of large-effect single gene variants. *Proc Natl Acad Sci USA* 115: E676-E685 (2018).
24. Lello L, Avery SG, Tellier L, Vazquez AI, de los Campos G, Hsu SDH. Accurate Genomic Prediction of Human Height. *Genetics* 210: 2477-2497 (2018).
25. Shaffer JR, Li J, Lee MK, Roosenboom J, Orlova E, Adhikari K, et al. Multiethnic GWAS Reveals Polygenic Architecture of Earlobe Attachment. *Am J Hum Genet* 101: 913-924 (2017).
26. Adhikari K, Fontanil T, Cal S, Mendoza-Revilla J, Fuentes-Guajardo M, Chacón-Duque JC, et al. A genome-wide association scan in admixed Latin Americans identifies loci influencing facial and scalp hair features. *Nat Commun* 7: 10815 (2016).
27. <https://www.boe.es/buscar/act.php?id=BOE-A-2007-17634>
28. Tawfik DS, Griffiths AD. Man-made cell-like compartments for molecular evolution. *Nat Biotechnol* 16: 652-656 (1998).
29. Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475: 348-352 (2011).
30. Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AM, et al. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309: 1728-1732 (2005).
31. Ju J, Hyun Kim D, Bi L, Meng Q, Bai X, Li Z, et al. Four-color DNA Sequencing by synthesis using cleavage fluorescent nucleotide reversible terminators. *Proc Natl Acad Sci USA* 103: 19635-19640 (2006).



32. Just RS, Moreno LI, Smerick JB, Irwin JA. Performance and concordance of the ForenSeq™ system for autosomal and Y chromosome short tandem repeat sequencing of reference-type specimens. *Forensic Sci Int Genet* 28: 1-9 (2017).
33. Xavier C, Parson W. Evaluation of the Illumina ForenSeq™ DNA Signature Prep Kit – MPS forensic application for the MiSeq FGx™ benchtop sequencer. *Forensic Sci Int Genet* 28: 188-194 (2017).
34. Moreno LI, Galusha MB, Just R. A closer look at Verogen's ForenSeq™ DNA Signature Prep kit autosomal and Y-STR data for streamlined analysis of routine reference samples. *Electrophoresis* 39: 2685-2693 (2018).
35. Köcher S, Müller P, Berger B, Bodner M, Parson W, Roewer L, et al. Inter-laboratory validation study of the ForenSeq™ DNA Signature Prep Kit. *Forensic Sci Int Genet* 36: 77-85 (2018).
36. Zeng X, King J, Hermanson S, Patel J, Storts DR, Budowle B. An evaluation of the PowerSeq™ Auto System: A multiplex short tandem repeat marker kit compatible with massively parallel sequencing. *Forensic Sci Int Genet* 19: 172-179 (2015).
37. Montano EA, Bush JM, Garver AM, Larijani MM, Wiechman SM, Baker CH, et al. Optimization of the Promega PowerSeq™ Auto/Y system for efficient integration within a forensic DNA laboratory. *Forensic Sci Int Genet* 32: 26-32 (2018).
38. García O, Soto A, Yurrebaso I. Allele frequencies and other forensic parameters of the HID-Ion AmpliSeq™ Identity Panel markers in Basques using the Ion Torrent PGM™ platform. *Forensic Sci Int Genet* 28: e8-e10 (2017).
39. García O, Ajuriagerra JA, Alday A, Alonso S, Pérez JA, Soto A, et al. Frequencies of the precision ID ancestry panel markers in Basques using the Ion Torrent PGMTM platform. *Forensic Sci Int Genet* 31: e1-e4 (2017).
40. Müller P, Alonso A, Barrio PA, Berger B, Bodner M, Martin P, et al. Systematic evaluation of the early access applied biosystems precision ID Globalfiler mixture ID and Globalfiler NGS STR panels for the ion S5 system. *Forensic Sci Int Genet* 36: 95-103 (2018).
41. Kidd KK, Speed WC, Pakstis AJ, Furtado MR, Fang R, Madbouly A, et al. Progress toward an efficient panel of SNPs for ancestry inference. *Forensic Sci Int Genet* 10: 23-32 (2014).
42. Kosoy R, Nassir R, Tian C, White PA, Butler LM, Silva G, et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. *Hum Mutat* 30: 69-78 (2009).



ANEXO 1 Plataformas de secuenciación masiva en paralelo y diferentes kits comerciales para generación de bibliotecas para su uso en el campo forense

Las dos plataformas más ampliamente utilizadas en el campo de la Genética Forense corresponden a:

- Secuenciación mediante PCR en emulsión [28] usando tecnología de semiconductores [29] (Ion Torrent PGM e Ion Torrent S5 de ThermoFisher Scientific)
- Secuenciación por síntesis (“polony”, polymerase colony) [30] con identificación de bases por fluorescencia [31] (MiSeq de Illumina)

Para estas plataformas disponemos de diferentes kits comerciales de secuenciación o paneles para generación de bibliotecas para su uso en el campo forense:

- ForenSeq™ DNA Signature Prep Kit (Illumina): 58 STRs (incluidos 27 STRs autosómicos, 7 marcadores de haplotipos del cromosoma X y 24 STRs de cromosoma Y), 94 SNPs informativos de identidad, 56 SNPs informativos de ascendencia y 22 SNPs informativos de fenotipo (ver, por ejemplo [32-35]).
- PowerSeq Auto/Mito/Y System (Promega Corporation): Diferentes combinaciones de paneles que incluyen 23 STRs de cromosoma Y, 22 STRs autosómicos y región control del ADN mitocondrial (ver por ejemplo [36-37]).
- Precision ID series (ThermoFisher Scientific). Ver por ejemplo [38-40]).
 - Panel de identificación: 90 SNPs autosómicos y 34 SNPs de cromosoma Y
 - Panel de ancestralidad: 165 SNPs autosómicos
 - Panel de ADN mitocondrial (región control y genoma completo)
 - Panel GlobalFiler™ NGS STR: 35 marcadores STR incluyendo 21 STRs de CODIS, 9 STRs adicionales y 4 marcadores de determinación de sexo
 - Panel fenotípico: 24 SNPs

De los diferentes kits comerciales anteriormente mencionados relacionados con la ancestralidad hemos de destacar el sistema Precision ID Ancestry panel (ThermoFisher Scientific) que incluye 165 SNPs autosómicos que proporcionan información biogeográfica de ascendencia. De este total de marcadores, 55 de ellos se seleccionaron basándose en el panel de Kenneth Kidd [41] y 123 marcadores se seleccionaron basándose en el panel de Michael Seldin [42] (nótese que 13 SNPs son coincidentes en ambos paneles). Por otro lado, también está el sistema ForenSeq™ DNA Signature Prep Kit (Illumina) que incluye los 55 SNPs descritos por Kenneth Kidd [41].