



Revista de Bioética y Derecho

www.bioeticayderecho.ub.edu – ISSN 1886 -5887

ARTÍCULO

Acceso a la secuencia del genoma (I). Interés en el acceso a la secuencia cruda y políticas de las instituciones secuenciadoras

Accés a la seqüència del genoma (I). Interès en l'accés a la seqüència crua i polítiques de les institucions seqüenciadores

Access to the genome sequence (I). Interest in the access to the raw sequence and policies of sequencing institutions

Teresa Pàmpols¹, Pilar Nicolás Jiménez², María Fenollar-Cortés³,
Ignacio Blanco⁴, Antonio Pérez Aytés⁵, Juan Francisco Quesada Espinosa⁶,
José Miguel García Sagredo⁷

1. Teresa Pàmpols. Consultora senior. Sección de errores congénitos del metabolismo-IBC. Servicio de Bioquímica y Genética Molecular. Hospital Clínico de Barcelona (jubilada), España. Email: pampols.ros@cofb.net. ORCID: <https://orcid.org/0009-5091-1353>. 2. Pilar Nicolás Jiménez. Departamento de Derecho Público, Facultad de Derecho Universidad del País Vasco/Euskal Herriko Unibertsitatea (España). Email: mariapilar.nicolas@ehu.eus. ORCID: <https://orcid.org/0000-0003-4166-5941>. 3. María Fenollar-Cortés. Facultativa especialista. Unidad de Genética Clínica. Servicio de Análisis Clínicos. Instituto de Medicina del Laboratorio. Hospital Clínico San Carlos, Madrid (España). Email: mariadelmar.fenollar@salud.madrid.org. ORCID: <https://orcid.org/0000-0002-7068-6306>. 4. Ignacio Blanco. Servicio de Genética. Hospital Universitario Germans Trias i Pujol, Badalona, Barcelona (España). Email: Iblanco.germanstrias@gencat.cat. ORCID: <https://orcid.org/0000-0002-7414-7481>. 5. Antonio Pérez Aytés. Grupo de Investigación en Perinatología. Instituto de Investigación Sanitaria. Hospital La Fe. Valencia (España). Email: aperezaytes@gmail.com. ORCID: <https://orcid.org/0000-0003-3401-6943>. 6. Juan Francisco Quesada Espinosa. Facultativo Especialista de Área. Servicio de Genética. Hospital Universitario 12 de Octubre, Madrid (España). Email: juanf.quesada@salud.madrid.org. ORCID: <https://orcid.org/0000-0002-4420-5214>. 7. José Miguel García Sagredo. Profesor honorífico, Facultad de Medicina. Universidad de Alcalá. Alcalá de Henares, Madrid (España). Email: jgsagredo@salud.madrid.org. ORCID: <https://orcid.org/0009-0000-7355-8425>.



Copyright (c) 2025 Teresa Pàmpols, Pilar Nicolás Jiménez, María Fenollar-Cortés, Ignacio Blanco, Antonio Pérez Aytés, Juan Francisco Quesada Espinosa, José Miguel García Sagredo. Esta obra está bajo una licencia de Creative Commons Reconocimiento-No Comercial-Sin Obra Derivada 4.0 Internacional.

Resumen

La secuenciación genómica es ampliamente utilizada hoy en día en investigación y en la práctica clínica. Genera una ingente cantidad de datos crudos que, debidamente analizados e interpretados, se comunican a la persona secuenciada vinculados al asesoramiento genético. En el entorno asistencial el informe interpretativo se incluye en la historia clínica pero los datos crudos se archivan aparte. Sin embargo, cada vez con más frecuencia, un número creciente de personas solicita a las instituciones secuenciadoras acceder a sus datos personales crudos (sin interpretar) por razones muy diversas. Esta petición es objeto hoy en día de un intenso debate debido en gran parte a las múltiples posibilidades de reutilización. Para reflexionar sobre ello se contemplan en primer lugar unos conceptos básicos sobre el genoma personal y los datos genómicos crudos. A continuación, se analizan los factores que han contribuido a la creciente disponibilidad de información genómica y las posibilidades de reutilizar la secuencia cruda con fines adicionales, clínicos, de salud, de investigación o incluso recreativos. La solicitud de acceso a los datos genómicos plantea cuestiones éticas, legales y prácticas, y por ello es interesante revisar las actuales políticas de almacenamiento y acceso de las instituciones secuenciadoras europeas y americanas. Partiendo de estos conocimientos previos, en un segundo artículo (Acceso a la secuencia del genoma II Consideraciones éticas, legales y sociales) se analizan específicamente y en profundidad dichos temas.

Palabras clave: Acceso a la secuencia del genoma; datos genómicos crudos; solicitud de acceso; reutilización de la secuencia genómica cruda; prácticas y políticas de almacenamiento y entrega de la secuencia genómica cruda.

Resum

La seqüenciació genòmica és àmpliament utilitzada avui en recerca i en la pràctica clínica. Genera una ingest quantitat de dades crues que, degudament analitzades i interpretades, es comuniquen a la persona seqüenciada vinculats a l'assessorament genètic. En l'entorn assistencial l'informe interpretatiu s'inclou en la història clínica però les dades crues s'arxivien a part. No obstant això, cada vegada amb més freqüència, un nombre creixent de persones sol·licita a les institucions seqüenciadores accedir a les seves dades personals crues (sense interpretar) per raons molt diverses. Aquesta petició és objecte d'un intens debat degut en gran part a les múltiples possibilitats de reutilització. Per a reflexionar sobre això es contemplen, en primer lloc, uns conceptes bàsics sobre el genoma personal i les dades genòmiques crues. A continuació, s'analitzen els factors que han contribuït a la creixent disponibilitat d'informació genòmica i les possibilitats de reutilitzar la seqüència crua amb finalitats addicionals, clínics, de salut, de recerca o fins i tot recreatius. La sol·licitud d'accés a les dades genòmiques planteja qüestions ètiques, legals i pràctiques, i per això és interessant revisar les actuals polítiques d'emmagatzematge i accés de les institucions seqüenciadores europees i americanes. Partint d'aquests coneixements previs, en un segon article (Accés a la seqüència del genoma II Consideracions ètiques, legals i socials) s'analitzen específicament i en profunditat aquests temes.

Paraules clau: Accés a la seqüència del genoma; dades genòmiques crues; sol·licitud d'accés; reutilització de la seqüència genòmica crua; pràctiques i polítiques d'emmagatzematge i lliurament de la seqüència genòmica crua.

Abstract

Genome sequencing is nowadays widely used for research as well as in clinical practice. It generates a great amount of raw data that once analyzed and interpreted are delivered to the sequenced individuals linked to genetic counseling; in clinical practice, the laboratory report is included in the medical record. However, an increasing number of people request the sequencing institution the access to its personal raw (uninterpreted) data for very different reasons. This access is currently being subject to a deep debate due in large part to its multiple reuse possibilities. To reflect on it, there are considered, first of all, some basic concepts on personal genome and raw genomic data. Next, the factors that have contributed to the increasing availability of genomic information and the possible reuses of the raw genome sequence, which can then be used for additional clinic, health, research or even recreational purposes, are analyzed. Access upon request to raw genomic data pose ethical, legal and practical issues, it is therefore interesting to look at practices and policies of data storage and personal access to raw data of European and American sequencing institutions. Based on this prior knowledge, in a second article access to genome sequence II. Ethical, legal and social (ELSI) issues, these topics are specifically and in depth analyzed.

Keywords: Access to genome sequence; raw genomic data; access upon request; reuses of genome raw sequence; practices and policies for storage and delivery of genome raw sequence.

1. Introducción

En 2003, gracias al Proyecto Genoma Humano, se obtuvo la primera secuencia de un genoma humano promedio de referencia, aunque los primeros genomas individuales no se secuenciaron hasta 2007-2008. Quince años más tarde hay millones de personas cuyo genoma ha sido secuenciado debido a los rápidos progresos de las tecnologías secuenciadoras del ADN y al descenso espectacular de sus costes, lo que ha facilitado sus aplicaciones con fines diagnósticos y de investigación biomédica, y el desarrollo de la medicina genómica personalizada de precisión, que ha generado una larga serie de proyectos de investigación implicando a grandes cohortes de participantes con el objetivo de relacionar el genoma con la clínica y estilos de vida (Kovanda *et al.*, 2021; Wiews J *et al.*, 2024; Karaboue M & Lacaella GU, 2025). A la vez, también ha supuesto la aparición de grandes compañías comerciales DTC-GT (*Directed-to-consumer Genetic testing*) de venta directa de pruebas genéticas de todo tipo a los consumidores, incluyendo los populares estudios de genealogías y antepasados.

Esta creciente disponibilidad y expansión de la información genética y datos genómicos constituye un desafío que interpela a profesionales sanitarios, investigadores, pacientes, individuos, familias, instituciones, industria y gobiernos, y hace necesario revisar sus respectivos papeles y responsabilidades (Borry *et al.*, 2018).

Una de las consecuencias de esta disponibilidad ha sido la petición de acceso a los datos genómicos crudos por parte de los pacientes y los participantes en proyectos de investigación, lo cual plantea cuestiones éticas, legales y de operatividad.

En la asistencia sanitaria los pacientes tienen derecho a acceder a sus datos médicos y a su historia clínica, incluyendo los resultados de pruebas de todo tipo y datos personales. En lo que concierne a las pruebas genéticas, es un requisito ético y también de los estándares para la acreditación de los laboratorios, la entrega de informes interpretativos que deben ir acompañados del correspondiente asesoramiento genético. Pero los datos crudos sin interpretar del laboratorio no se considera tradicionalmente que sean parte de la historia clínica y no se incluyen en la misma, ya que la secuencia cruda de los más de 6.400 millones de pares de bases nucleotídicas del genoma de las células somáticas no tiene en sí misma un significado, es una secuencia repetitiva ilegible e ininterpretable si no se dispone de los recursos bioinformáticos y la experiencia adecuada.

Los datos genómicos crudos una vez entregados pueden, sin embargo, ser reanalizados, reinterpretados y reutilizados de nuevo con propósitos adicionales clínicos, de salud, de investigación o incluso creativos. La información que contienen los datos crudos tiene potencial para ser clínicamente significativa y alterar el proyecto vital del paciente y su familia, por lo que entregarla sin la suficiente preparación previa o soporte podría considerarse éticamente

cuestionable. Sin embargo, la curiosidad acerca del genoma personal no es inherentemente negativa y hay cada vez mayor consenso en que las personas tienen el derecho a acceder a su propia secuencia en bruto, en tanto que es algo que los hace únicos, los identifica y les pertenece (Prainsack, 2014). Los progenitores o tutores tienen también un derecho de acceso a la información sobre los menores de edad, actuando en su nombre y en su mejor interés, si bien este derecho es más evidente en el contexto asistencial que en el de la investigación.

La persona que recibe una copia de su genoma o del menor de edad, adquiere responsabilidades acerca de su custodia y futuros usos, ya que su derecho a recibirlas plantea posibles conflictos de derechos e intereses de terceras partes o personas: los profesionales sanitarios, las instituciones sanitarias, los miembros de la familia que comparten información genética y el menor cuando es el caso.

En el contexto de la investigación el análisis ético también sugiere que los participantes tienen derecho a recibir sus datos genómicos crudos. Además, esto favorece su compromiso con la participación ciudadana en la ciencia y los posiciona para una mejor administración de su salud y bienestar, sin embargo, no por ello dejan de suscitar cuestiones acerca de la sostenibilidad de los costes de la entrega, su posible impacto en la propiedad intelectual y el reconocimiento del derecho de los investigadores que generaron los datos a un adecuado reconocimiento científico de su producción, así como si podría frenar la libertad académica de los investigadores, además de dar lugar a riesgos para el participante ya que se están compartiendo datos experimentales crudos no certificados como los de un laboratorio clínico, que pueden ser reinterpretados o reutilizados por otros (Winkler EC *et al*, 2019; Schickhardt *et al*, 2020).

Estos potenciales conflictos de derechos e intereses hacen necesario orientar ética y legalmente cómo se hace efectivo en la práctica el derecho de acceso al genoma crudo.

El tema merece un análisis extenso y por ello se ha estructurado en dos artículos. El presente, "Acceso a la secuencia del genoma (I). interés en el acceso a la secuencia cruda y políticas de las instituciones secuenciadoras", en el que se exponen conceptos básicos sobre el genoma personal, así como el contexto que ha llevado a su creciente disponibilidad y a la consiguiente solicitud de acceso al genoma crudo, analizando sus posibles usos una vez obtenido, y revisando a continuación las actuales políticas de acceso de las instituciones secuenciadoras europeas y americanas en base a las regulaciones éticas y legales nacionales; y un segundo artículo "Acceso a la secuencia del genoma (II). Consideraciones sociales, éticas y legales", donde se abordarán con mayor profundidad los aspectos éticos, legales y sociales (AELS), a fin de que a partir de las conclusiones que se obtengan, puedan proponerse finalmente líneas de buena práctica que sean de utilidad para orientar la política de nuestras instituciones secuenciadoras.

2. El genoma humano: conceptos básicos

2.1. Qué es el genoma

Denominamos genoma a la secuencia ordenada de los pares de bases del ADN que nos identifica como especie. La secuencia que se encuentra en nuestros gametos tiene aproximadamente 3.200 millones de pares de bases, que identificamos con las letras A, T, G y C (adenina, timina, guanina y citosina). Nuestras células somáticas son diploides ya que proceden de la unión de un gameto materno y otro paterno, por lo que, al secuenciar lo que en realidad obtenemos observamos y analizamos es la secuencia de unos 6.400 millones de pares de bases de ADN, heredando la mitad de nuestro genoma de cada progenitor”.

Podemos secuenciar un panel de genes o secuenciación dirigida (*targeted sequencing*); la parte del genoma que se traduce a proteínas, es decir el exoma WES (*whole exome sequencing*) que representa un 2% del genoma; o el genoma completo WGS (*whole genome sequencing*).

El 99,9 % del ADN nos caracteriza como especie y lo compartimos todos los seres humanos. El 0,1% restante (unos sesenta y cuatro millones de pares de bases) es el que nos diferencia a los unos de los otros y nos hace únicos. Por eso decimos que el ADN de una persona es identificativo, aunque, de hecho, solo lo es si tenemos información previa para compararlo, por ejemplo, con otra muestra anterior, o bien conocemos el ADN de nuestros antecesores, descendientes o hermanos/as.

Si bien es cierto que nuestro ADN es personal y único, no solo me identifica a mí, sino que contiene información de nuestros familiares. Cualquier quiebra de la privacidad de la información genómica acerca del estado de salud de un individuo puede afectar potencialmente a otros miembros de la familia y, de hecho, se ha propuesto una nueva categoría de privacidad, la “privacidad de red familiar” (*familial network privacy concern*) (Chow-White, 2015).

2.2. Qué son los datos genómicos crudos

El término “datos genómicos crudos” se refiere al conjunto de la gran cantidad de datos genómicos generados por las *next generation sequencing technologies* (NGS) y *high-throughput sequencing technologies* (HTS), o de secuenciación masiva del genoma o del exoma durante las etapas primaria y secundaria del proceso informático de análisis de la secuencia para la interpretación final de su significado científico-clínico. Los datos crudos son claramente distintos de los datos específicamente interpretados y anotados como hallazgos relevantes para la salud del individuo o los objetivos de una investigación. Los datos crudos son también fundamentalmente diferentes

de lo que se entiende por resultados de la investigación cuyo retorno tiene requerimientos específicos (Jarvik GP *et al.*, 2014; Vearns DF *et al.*, 2023; Wolf S & Green R, 2023).

Los tipos y formatos de los datos crudos dependen de la plataforma secuenciadora, de los algoritmos informáticos empleados para procesar los datos y del tipo de texto en que se van a almacenar y leer.

Los formatos de archivo para almacenarlos son típicamente: FASTQ (formato de archivo basado en texto para almacenar la secuencia de nucleótidos junto con sus puntuaciones de calidad); BAM (de *Binary alignment map*), formato de archivo binario comprimido que contiene información de la secuencia después de haber sido alineada con un genoma de referencia) y SAM (de *Sequence Alignment Map*), para las etapas primarias del proceso informático y VCF (de *Variant call format*), que almacena información sobre las variaciones en la secuencia respecto a un genoma de referencia. En el estudio de tumores el término datos crudos incluye los datos diferenciales entre el genoma del tumor y el genoma constitutivo.

Otras veces los datos crudos no son la secuencia completa del genoma o del exoma, sino que son el resultado de un chip o micro matriz de varios miles de SNPs (*Single nucleotide polymorphisms o “snips”*), es decir, polimorfismos de un único nucleótido.

2.3. El espacio de almacenamiento de una copia del genoma

Para secuenciar un genoma, los millones de pares de bases del mismo se rompen en pequeños fragmentos de ADN que se secuencian repetidamente en paralelo. A continuación, se reconstruye o ensambla *“in silico”* estos fragmentos del ADN de la persona comparándolo o cartografiándolo con uno o varios genomas de referencia a fin de identificar variaciones o variantes que pueden afectar la salud.

Aunque la precisión de la secuenciación para cada nucleótido individual es alta, dado el elevado número de nucleótidos en el genoma, si solo se secuencia una vez puede haber un número significativo de errores, por lo que, para aumentar la precisión, una región del genoma se secuencia múltiples veces, unas 30 para el diagnóstico de enfermedades raras y unas 100 en caso de tumores. El número de veces que se secuencia se denomina profundidad de lectura o cobertura.

Una copia del genoma ocupa aproximadamente 3,2 Giga bits (GB) y el genoma humano diploide 6,4 GB, pero el número de datos que deben almacenarse por cada genoma a fin de corregir los errores que pueden ocurrir durante la secuenciación y para el análisis preliminar es de unos 120 GB o más. En el caso del cáncer el número de datos es mucho mayor porque se secuencia el

tumor y el genoma constitutivo. En los laboratorios de investigación se procesan Terabytes de información (1 Terabyte = 1024 GB).

Nuestra habilidad actual para secuenciar el ADN sobrepasa nuestra capacidad de descifrar la información que contiene, por lo que la ciencia de los datos genómicos será una vibrante área de investigación en los próximos años. Dada la gran cantidad de datos complejos asociados al genoma, la genómica se considera un área de “bigdata” con requerimientos computacionales específicos. Hay estimaciones que predicen que la investigación genómica generará entre 2 y 40 Hexa bits (HB) de datos en la próxima década (un HB equivale a mil millones de Giga bits (GB) o 1018 bits) y que almacenar y procesar datos genómicos excederá las necesidades computacionales de YouTube y Twitter. Las instituciones secuenciadoras podrían ser los mayores usuarios del mundo de almacenamiento de datos (Kahn, 2011; CheckHyden, 2015).

2.4. El coste del almacenamiento y entrega de la copia de la secuencia cruda

Calcular los costes de almacenamiento en la nube es complicado porque depende no solo del tamaño de los archivos sino del volumen de test que realiza el laboratorio, el nivel de almacenamiento, la compresión de los datos y el tiempo que se van a guardar. El coste de almacenar 1GB durante un mes puede oscilar entre 0,009 y 2,3 céntimos de \$ en servicios de nube como Amazon web, Google *cloud platform* o Azur (Krimm et al., 2020). A primera vista puede parecer poco, pero es un coste acumulativo y teniendo en cuenta el número de genomas que secuencie un laboratorio y el tiempo que estipule de almacenamiento, las cifras pueden ser importantes en su presupuesto, por lo que hace años que se exploran mejores sistemas de compresión y computación. La protección de la privacidad y la identidad son muy relevantes y el uso de la tecnología de cadena de bloques (*blockchain*) es una opción prometedora (Langmead & Nellore, 2018; Gürsoy et al., 2022).

El coste de entregar la secuencia cruda de un genoma depende del tamaño del fichero, pero dado que ocupará como mínimo unos 120 GB, habría que pensar en una unidad USB que pueda tener capacidad de hasta 1 terabyte. En cualquier caso, tanto el presupuesto de un laboratorio asistencial como el de un proyecto de investigación deben contemplar los costos de entregar los datos crudos a los pacientes o a los participantes.

3. La disponibilidad de información genómica y el interés por el acceso a la secuencia cruda

La primera secuencia del genoma humano se completó en 2003, en el marco del Proyecto Genoma Humano, a partir de una muestra compuesta de ADN procedente de voluntarios cuyo anonimato fue celosamente protegido (www.genome.gov). Pero el primer genoma individual no se secuenció hasta 2007 y fue el de Craig Venter (fundador de *Celera Genomics* que emprendió su propio Proyecto Genoma Humano en 1999 al margen del consorcio público) y el segundo, en 2008, el de James Watson (premio Nobel, codescubridor de la doble hélice del ADN y padre del Proyecto Genoma Humano) al cual se le entregó la secuencia en un DVD. Ambos hicieron público su genoma personal, Venter lo depositó en la base de datos pública, *GeneBank* y Watson autorizó su publicación en la revista *Nature* (Wheeler et al., 2008). Debido a ello se cuestionó por primera vez el hecho de hacer pública la secuencia en una base de datos abierta, sin tener en cuenta además de los riesgos propios, los concernientes a los familiares relacionados.

Veinte años más tarde del primer borrador hay millones de ciudadanos con su genoma secuenciado y a veces depositado en bases de datos de acceso público (Stephens et al., 2015). Han contribuido a esta disponibilidad:

3.1. El descenso de los costes de la secuenciación genómica

Secuenciar el genoma de James Watson en 2008 todavía costó 1 millón de dólares. A partir de entonces se produjo un punto de inflexión en la curva de descenso de los costos de las NGS, llegando a bajar hasta los 1.000 \$ en el 2015 y todavía sigue bajando debido en parte a las nuevas técnicas de secuenciación de tercera generación. Este descenso, junto con los avances en el desciframiento de la información que contiene, ha propiciado su uso creciente con finalidades clínicas de diagnóstico y de investigación biomédica.

3.2. Las expectativas del desarrollo de la medicina personalizada de precisión

Una parte importante de este desarrollo es el uso de la información genómica para conseguir diagnósticos mejores y más costo efectivos, prevenir riesgos y proporcionar mejores cuidados de salud, lo cual ha contribuido a que un número creciente de países inviertan esfuerzos en explorar el genoma humano y ha generado una larga serie de proyectos de investigación reclutando a grandes cohortes de participantes con el objetivo de relacionar el genoma con la clínica y estilos de vida.

En una revisión que investigó 240 países, se identificaron 86 con proyectos genómicos y/o bases de datos genómicos, de los cuales 41 están activos actualmente (Kovanda *et al.*, 2021). La revisión, al ser de 2021, no incluye la infraestructura española IMPaCT, impulsada y financiada por el Instituto de Salud Carlos III en 2022 y que es la base para la implementación de la medicina personalizada de precisión en el Sistema Nacional de Salud (<https://impact.isciii.es>).

España participa así mismo en la Iniciativa Europea *1+Million Genomes*, que está dirigida a facilitar acceso seguro a la genómica y los correspondientes datos clínicos a través de Europa para una mejor investigación, cuidados de salud personalizados y desarrollo de políticas sanitarias, participan instituciones académicas e industria de 25 países.

3.3. La aparición de grandes compañías comerciales de venta directa de pruebas genéticas a los consumidores (DTC) y la emergencia de portales TPI (*Third Party Interpretation*) o de interpretación por terceros

Las compañías DTC (*Directed to Consumer*) emergieron a principios del 2000 como un medio que permitía a los consumidores acceder a información genética sin la intervención de un médico o de los servicios de salud. Por aquel entonces las pruebas se enfocaron al infoentretenimiento no médico (rasgo de inserción del lóbulo de la oreja, propensión a sonrojarse al beber alcohol, etc.) y a la información étnica y familiar (genealogías, antepasados) pero a veces incluía también información sobre trastornos poligénicos y la oferta se ha ido haciendo cada vez más extensa incluyendo pruebas para trastornos mendelianos, nutrigenética, bienestar, buen estado físico o habilidades, personalidad y salud emocional (Allyse *et al.*, 2018).

El suministro de información relacionada con la salud ha suscitado controversia en el ámbito sanitario y un importante debate ético y legal sobre la validez y utilidad clínica de las pruebas, la privacidad de los consumidores, la honestidad de la propaganda y la ausencia de supervisión médica (Greely, 2020), llevando a numerosas organizaciones profesionales a manifestar su posición sobre el uso de las DTC-GT con propósitos de salud incluyendo la Asociación Española de Genética Humana (AEGH) (Pàmpols *et al.* 2019) y algunos países, como España, tienen legislación restrictiva al respecto (Romeo Malanda, 2013).

En 2010 la *Food and Drug Administration* (FDA) inició procesos reguladores y en 2013 envió cartas de advertencia a las compañías que ofrecían información genética relacionada con la salud por violaciones de la *Food, Drug and Cosmetic Act*, por lo que las compañías tuvieron que limitarse a proporcionar únicamente los resultados de antepasados, si bien daban acceso a los usuarios a los datos crudos del genoma. Esta situación hizo emerger portales TPI (*Third Party Interpretation*)

o de interpretación por terceros, que mediante análisis bioinformático de los datos crudos generan un informe sobre las variantes genómicas del usuario, incluyendo información clínica. A partir de 2015 hasta 2022, la FDA ha ido autorizando a 23andMe, que había ido adecuando su actividad, a la venta de diversas pruebas genéticas. Sin embargo, la autorización de la FDA advierte de las limitaciones de las pruebas, que el informe no sustituye la intervención de un profesional sanitario y que puede ser necesaria la verificación de resultados en un laboratorio clínico que tenga la certificación de la CLIA (*Clinical Laboratory Improvement Amendment*) y la acreditación del CAP (*College of American Pathologists*) ya que solo sus resultados pueden ser utilizados con fines diagnósticos o de tratamiento.

A principios de 2019 se estimaba que más de 26 millones de consumidores habían accedido a pruebas DTC de antepasados o por motivo de salud y que si se continuaba al mismo ritmo, serían 100 millones en 2021 (Regalado, 2019). No obstante, resulta relevante reflexionar sobre la evolución de la demanda de pruebas DTC, que ha disminuido en 2025, al mismo tiempo que la empresa emblemática 23andMe enfrenta el riesgo de colapso. A esto se suma la creciente preocupación expresada por sectores críticos respecto a la falta de fiabilidad de las pruebas, los problemas de privacidad y ciertas limitaciones como recurso para la investigación (Kwon, 2025; Rutherford A, 2025).

4. Usos de la secuencia cruda del genoma

Una vez obtenida una copia de la secuencia cruda puede utilizarse para su reanálisis y reinterpretación para tener una segunda opinión médica, pero también da acceso a todo tipo de servicios y productos que pueden ser muy atractivos para nuestra sociedad consumista, incluyendo las aplicaciones con fines recreativos o la genética como entretenimiento, como apps para el móvil o TV o incluso bufandas o chales personalizados con tu información genética (Khan & Mittelman, 2018). La compañía HLTH NFT se anuncia como el primer DNA NFT *marketplace* y DNA *verse* (dnaverse.io) y ya ha creado la primera colección de NFT (*Non fungible tokens*) con datos de ADN reales en el metaverso, ofreciéndose a inmortalizar tu unicidad con precisión científica en una obra de arte NFT (Jones, 2021).

Los portales TPI procesan los datos crudos basados en chips de SNPs y los datos crudos del genoma o del exoma completos (WGS o WES) y permiten cargar archivos FASTQ o BAM utilizando los servicios de nube de Amazon.

Algunos ofrecen interpretación de antepasados, a veces combinada con otras características: *Genomelink*; *GED match*; *Myheritage DNA*; *FamilyTree DNA*; *Living DNA*; *My True Ancestry*. Otras como *Genopalat*, *Athletigeny*, *DNA Fit*, dan información sobre nutrición, salud, bienestar y fitness. *DNA romance* utiliza algoritmos con datos de 120 marcadores incluyendo atracción química y personalidad para encontrar una pareja adecuada. *Promethease* permite comparar directamente los datos crudos de los chips de SNPs con la SNPedia, una “wiki” para genética humana que proporciona información sobre variantes genéticas y fenotipos (Cariasso et al., 2012).

PARABON ofrece su servicio *Snapshot Phenotyping Service*, para la predicción de aspecto físico a partir del ADN de una persona desconocida. Su *Snapshot F DNA Phenotyping Forensic* predice antepasados genéticos. Este recurso y el acceso a bases de datos públicas de genealogías se utilizan por la policía para resolver crímenes. El portal *DNAolves* invita directamente a resolver crímenes aportando tu DNA.

Sano Genetics promueve la investigación científica creando herramientas para los investigadores para analizar grandes bases de datos anónimos y reclutar participantes en la investigación. *DNA Land* y *Luna DNA* son portales académicos que dan la oportunidad de ayudar a la investigación científica e innovación aportando tu genoma.

Antes de cargar los datos crudos en un portal TPI es recomendable informarse. En una revisión de 29 portales TPI, se comparan 5 de ellos seleccionados por su constante mención en artículos sobre el genoma personal, blogs y fórmulas y que proporcionan información sobre la salud empleando datos genómicos basados en SNPs: *Codegen.eu*, *Interpretome*, *Live Wello*, *Enlis Personal* y *Promethease* (Badalato et al., 2017). Se analiza la validez y utilidad clínica de los resultados, los aspectos relacionados con la autonomía y consentimiento informado, la potencial interpretación errónea de los resultados sin supervisión médica, la propaganda y material promocional y la política de privacidad y uso de los datos. Las dudas sobre la validez y utilidad clínica de las DTC-GT, son extensivas a la de las predicciones de riesgo aportados por los TPI, que todavía requieren más investigación, y pueden ser difíciles de entender y ser malinterpretadas sin la adecuada supervisión médica; la información sobre riesgos y limitaciones de los resultados suele ser escasa mientras que se ensalzan los beneficios; la política de privacidad, el uso de datos y tiempo de almacenamiento no está además sometido a los estrictos estándares y regulaciones de las instituciones sanitarias o de investigación (Badalato et al., 2017). Reducir la información personal de salud a datos es éticamente problemático y es importante considerar el impacto que los resultados pueden tener en la concepción individual de la salud, identidad personal y estructuras familiares. La consulta de los resultados con un profesional sanitario antes de tomar decisiones en materia de salud es altamente recomendable ya que los falsos positivos son frecuentes (Moscarello et al., 2019).

Ceder los datos en bruto a portales TPI con fines de investigación, es una acción solidaria y altruista, pero es importante asegurarse de las garantías que ofrecen y considerar cómo protegen la privacidad, los riesgos de reidentificación y las implicaciones para los familiares, puesto que al ceder nuestros datos estamos cediendo también una parte importante de los suyos (Shahani&Borry, 2015). La protección total de la privacidad es poco realista en esta era de DTC-GT, dada, además, la creciente popularidad de compartir la información genética en las redes sociales (Knoppers, 2010) y la proliferación de bases de datos genómicos de consumidores que incrementa notablemente el riesgo de reidentificación (Gymrek *et al.*, 2013; Haeusserman *et al.*, 2017; Erlich *et al.*, 2018)

El gran valor del genoma personal y los datos de salud asociados para la investigación y el desarrollo de la industria farmacéutica y biotecnológica, han llevado a establecer acuerdos con las compañías DTC-GT para comprarles datos. En este contexto ha aparecido una nueva generación de *start-ups* que proponen, entre otros, dejar el control de los datos en manos de cada usuario. Son las llamadas “DNA marketplace” o plataformas de intercambio de DNA, que facilitan el contacto entre usuarios y compañías para compartir datos mediante una plataforma de cadena de bloques y recibir a cambio incentivos monetarios en forma de criptovalores, eligiendo cómo y con quién compartir datos y ser compensados por ello. Son ejemplos *NebulaGenomics*, *Luna DNA* y *EncryptoGen* (Ahmed *et al.*, 2019). La participación de los individuos en estas *start-ups* contribuyendo a proporcionar datos con ánimo de lucro suscita un debate ético sobre si una persona debe poder obtener dinero de sus recursos biológicos, similar a la venta de esperma o sangre, o sobre si el individuo es realmente el propietario de su genoma, dado que una buena parte del mismo es compartido con los familiares, así como cuestiones relativas a las estrategias de incentivos y su impacto sobre el consentimiento, si los familiares consienten en compartir y vender datos, si va a repercutir en la investigación con fondos públicos sin ánimo de lucro, la duración del consentimiento, la protección de la privacidad y seguridad y riesgos todavía desconocidos por la novedad de estas plataformas (Mozilla Foundation, 2019; Koplin *et al.*, 2022; Rosenbaum, 2022).

Las compañías de pruebas genéticas orientadas a los consumidores transforman el importante volumen de datos genómicos que reúnen en activos, creando beneficios financieros, y el capitalismo de datos y la bio-capitalización suscitan importantes aspectos éticos y sociales (Geiger & Gross, 2021).

5. Políticas de acceso de las instituciones secuenciadoras a la secuencia cruda del genoma

El Reglamento (UE) 2016/679 de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos, en adelante RGPD), reconoce el derecho de acceso a los datos personales, pero prevé que no afectará negativamente a derechos de terceros y, además, deja margen a los Estados para plantear excepciones en el ámbito de la investigación. La cuestión es, por un lado, si el derecho de acceso es extensivo a la secuencia en crudo y qué límites o excepciones se pueden aplicar (sobre esto se profundizará, en un segundo artículo Acceso a la secuencia del genoma (II) Consideraciones éticas, legales y sociales). La práctica habitual es que las instituciones entreguen una copia cuando se presenta una solicitud de acceso.

El primer estudio empírico acerca de las políticas y prácticas de acceso de las Instituciones secuenciadoras (IS) en los Estados miembros de la EU/EEA (Unión europea y Área económica europea) se publicó en 2020 (Narayanasany *et al.*, 2020). Accedieron a ser entrevistados 39 participantes de 35 instituciones de 21 países, que generan, procesan y almacenan datos de la secuenciación de genomas y exomas con propósitos clínicos o de investigación. En el periodo estudiado de 10 meses, se generaron 317.259 secuencias.

De cara a valorar la accesibilidad se preguntaba en qué tipo de formato de archivo estaban disponibles las secuencias y durante cuánto tiempo permanecían accesibles, factores que estaban condicionados por la capacidad de la estructura de almacenaje y las políticas de almacenamiento internas (institucional), nacionales o europeas. El formato más frecuente era FASTQ, seguido de BAM y VCF. Cinco instituciones utilizaban tecnologías de compresión de datos para ahorrar costos de almacenaje. Las IS empleaban medidas de seudonimización o encriptación para garantizar la confidencialidad y seguridad.

El tiempo de almacenamiento iba de tres meses a 115 años e indefinidamente. En total las retenían indefinidamente 284.522 conjuntos de datos genómicos en crudo. Respecto a la sostenibilidad de guardarlos indefinidamente, la mayoría lo consideraba inviable a largo plazo, por ejemplo, más de 10 años. A pesar de tener un marco regulador parecido, las políticas y prácticas de almacenamiento eran muy diversas. El 10% de las IS asumían la responsabilidad de almacenar los datos genómicos en cumplimiento de la legislación nacional para los datos clínicos, asumiendo que los datos genómicos eran considerados como tales.

La parte final del estudio era la experiencia en cuanto a solicitudes y entrega de datos crudos. Se documentaron solo 28 casos en 12 instituciones secuenciadoras de 10 países, una cifra muy baja (0,009%) dada la gran actividad secuenciadora. El formato de archivo entregado más a menudo fue VCF seguido de BAM y FASTQ, la mayoría fueron entregados en un dispositivo externo y unas pocas permitían acceso para descargarlos.

Las personas secuenciadas y las instituciones secuenciadoras no están en contacto directo, sino que la petición se cursa a través de un contacto intermediario, generalmente un profesional de la salud o un investigador principal. La autorización para garantizar el acceso es a discreción de dicho intermediario, en algunos casos era evaluada por un panel incluyendo personal de la IS. En resumen, estos contactos intermediarios actúan como “gatekeepers” o guardianes para que las personas secuenciadas accedan a sus datos crudos. Solo una institución tenía un procedimiento estandarizado interno para atender las peticiones.

En el estudio se solicitaba a las IS participantes que enviaran el modelo de consentimiento informado para la secuenciación. De los 9 sobre 35 participantes que se recibieron únicamente dos incluían información sobre la política de acceso a los datos crudos. Una de las respuestas resaltaba que la política de la organización era de no dar acceso a los datos a menores que no hubiesen alcanzado la edad legal.

Una consideración importante en la discusión sobre la política de almacenamiento y acceso era la distinción entre datos clínicos y datos de investigación, principalmente porque los datos de una investigación son experimentales y puede que no tengan suficiente calidad para ser reutilizados por el usuario y porque, según el RGPD, su acceso puede estar restringido (art. 89.2), mientras que el derecho de acceso en el entorno clínico está implantado de forma general. Se plantean dudas adicionales cuando se considera que, dado el cada vez más bajo coste de la secuenciación, el mantenimiento a largo plazo podría no ser y que, de acuerdo con el principio de limitación del plazo de conservación, solo deben ser conservados mientras sea necesario para alcanzar un propósito específico.

También se manifestaba preocupación por quién debía asumir los costes de la entrega (infraestructura, hardware, personal/administración). A pesar de todo ello, el estudio pone de manifiesto la abrumadora tendencia de las IS a garantizar el acceso a la secuencia cruda, así como las complejidades técnicas y la de los procedimientos a seguir.

Dos años más tarde de este estudio, la *European Society for Human Genetics* (ESHG) publicó sus recomendaciones actualizadas (Souche *et al.*, 2022) en el cual se menciona que hay países como Holanda guarda los archivos intermedios o temporales FASTQ, SAM o BAM solamente un año, mientras que los archivos VCF los guarda cinco años.

En EE.UU. el *American College of Medical Genetics and Genomics* (ACMG) y el *College of American Pathologists* recomendaron a los laboratorios clínicos, de acuerdo con las regulaciones de la CLIA (*Clinical Laboratory Improvement Amendment*), que guardasen por lo menos dos años las secuencias, garantizando el cumplimiento de la legislación local y federal sobre privacidad. (Rehm HI *et al.*, 2013; Aziz *et al.*, 2015). En documentos más recientes (Rehder C *et al.*, 2021; Moyer A *et al.*, 2025) siguen recomendando guardar dos años o más como mínimo todos los archivos (FASTQ, BAM y VCF), pero no hay un requerimiento universal para su almacenamiento indefinido y algunos laboratorios guardan menos tiempo los archivos primarios FASTQ y SAM o BAM de mayor tamaño que los archivos VCF, debiéndose advertir a los pacientes que pueden solicitar el tipo de archivo en el formato que consideren más útil de acuerdo con sus necesidades pero que tengan en cuenta estas y otras limitaciones, ofreciendo asesoramiento y ayuda al respecto. También son interesantes las recomendaciones acerca de la validación de procesos cuando el paciente solicita que se compartan los datos crudos con otro laboratorio clínico. El problema de guardar los archivos FASTQ, SAM o BAM son los elevados costes de almacenamiento, por lo que algunos autores han considerado si sería más efectivo volver a secuenciar al paciente si fuese necesario en el futuro. De todas formas, la genómica es una ciencia de “big data”, la situación es fluida y cambiante. El problema del almacenamiento de datos se tiene en muchos ámbitos, por lo que el desarrollo y evaluación de herramientas de compresión es un área en progreso que sin duda repercutirá en mejorar los costes de almacenamiento (Mcghee EJ & Milton S, 2023). La empresa Illumina, por ejemplo, ha desarrollado un método de compresión de dato genómico crudo (ficheros FASTQ) con una tasa de compresión de cuatro a seis veces mayor que el formato de compresión más utilizado, el zip estándar de GNU (gzip).

El *Canadian College of Medical Geneticists* también recomienda guardar las secuencias por lo menos dos años (Hume *et al.*, 2019).

En la referencia (Shahahni *et al.*, 2018) se contemplan las políticas de almacenamiento, el acceso a los datos en bruto y cómo enfocan el tema en el consentimiento informado para la prueba genética en seis laboratorios de EE.UU. Uno de ellos especifica que no proporcionará los datos crudos, pero que estará disponible un archivo VCF bajo petición y puede comportar costos adicionales. Tres laboratorios indican que el paciente o el médico en su nombre pueden solicitar el acceso, en principio para futuro reanálisis o validación. Uno especifica que puede ser solicitado

por una institución y deberán firmar la solicitud el investigador y el paciente y garantizarlo con un consentimiento informado. Dos solo indican que en el futuro los datos crudos en bruto puede que se integren en la historia clínica. Esta opción ha sido vista con reservas en algunas líneas guía o documentos de posición como la de la *UK PHG foundation* y la *European Society for Human Genetics* (ESHG) porque el almacenamiento de la historia clínica es a muy largo plazo y preocupa la protección de datos (Wright et al., 2011; van El et al., 2013).

Una alternativa sería guardar el formato VCF junto con el informe clínico final (Rehm HI et al., 2013; Aziz et al., 2015).

Las recomendaciones de las sociedades profesionales que se han citado se dirigen fundamentalmente a los laboratorios clínicos y al paciente que sea objeto de cualquier actuación sanitaria que implique la utilización de la secuenciación genómica, bien sea un proceso diagnóstico o un cribado poblacional genómico universal si llegase el caso, habiendo un claro consenso en que se le debe facilitar el acceso a sus datos crudos. En el contexto de la investigación el principio ético de reciprocidad por la contribución del participante en la investigación también apoya su entrega, pero tiene matizaciones propias. En los proyectos dirigidos a la medicina de precisión y la investigación translacional, la interacción entre investigación y cuidados clínicos genera nuevas cuestiones ya que son gobernados por principios éticos diferentes, por ejemplo, mientras que en la clínica la confidencialidad de los datos genómicos y el no compartirlos con terceros es esencial, en los grandes proyectos de la medicina de precisión, en cambio, es indispensable compartir los datos genómicos de cualquier tipo para compartir los beneficios de la investigación y promover el bien común (Winkler EC & Knoppers BM, 2022; Wiews J et al., 2024; Mc Cartney AM et al., 2024, Karaboue M & Lacaella GU, 2025) Este uso compartido de datos crudos tiene requerimientos propios y está fuera de los objetivos de esta publicación.

6. Conclusiones

La creciente disponibilidad y expansión de la información genética ha contribuido a que emerja el interés de las personas secuenciadas por acceder a la secuencia cruda (sin interpretar) de su genoma personal y hay cada vez mayor consenso en que tienen el derecho de hacerlo. Pero si bien es cierto que nuestro genoma es personal y único, también contiene información sobre nuestros familiares, y la persona que recibe una copia de la secuencia cruda de su genoma o de un menor de edad, adquiere responsabilidades acerca de su custodia y posibles usos, ya que la copia puede ser reanalizada con fines de salud, de investigación o incluso recreativos y también depositarse

posteriormente en bases de acceso público con lo que se podría perjudicar la autodeterminación informativa de los familiares, todo lo cual plantea cuestiones éticas y legales. Las políticas de acceso de las instituciones secuenciadoras de los países europeos no son uniformes a pesar de tener un marco regulador parecido. En general la autorización para garantizar el acceso es a discreción de un contacto intermediario, profesional de la salud, investigador principal o panel evaluador, pero no todas informan de su política de acceso en el consentimiento informado previo a la secuenciación. El formato del archivo para el almacenamiento y entrega no es uniforme y el periodo de almacenamiento varía entre 3 meses a 115 años e indefinidamente. Tanto en Europa como en EE. UU. hay una abrumadora tendencia a garantizar el acceso a la secuencia cruda, pero preocupan los costes y la sostenibilidad del almacenamiento a largo plazo, así como las complejidades técnicas y la de los procedimientos a seguir. Por eso es importante abordar con mayor profundidad los aspectos sociales y éticos, así como nuestro marco legal para las políticas de acceso de las instituciones secuenciadoras españolas lo cual se lleva a cabo en “Acceso a la secuencia del genoma (II). Consideraciones sociales, éticas y legales”, a fin de que puedan proponerse líneas de buena práctica, útiles para orientar su política al respecto.

Conflicto de intereses

Los autores declaran que no tienen conflictos de intereses.

Referencias bibliográficas

- ◆ Ahmed E & Shahani M. (2019) DNA Marketplace: An Analysis of the Ethical Concerns Regarding the Participation of the Individuals. *Front Genet* 10: 1107.
<https://doi.org/10.3389/fgene.2019.01107>.
- ◆ Allyse Ma, Robinson Dh, Ferber Mj& Sharp Rs. (2018) Directed-to-consumer testing 2.0 Emerging models of Direct-to-consumer genetic testing, *Mayo Clin Proc*. January; 93(1): 113 a 120.
- ◆ Aziz N, Zhao Q, Bry L, Driscoll Dk, Funke B, Gibson Js, Grody Ww, Hedge Mr, Hoeltge Ga, Leonard Dgb, Merker J D, Nagarajan R, Palicki La, Robetorye Rs, Schrijver I, Weck Ke & Voelkerding Kv. (2015) College of American Pathologists Laboratory standards for Next-Generation sequencing clinical tests. *ArchPatholLabMedApril*; 139(4): 481 a 493.
<https://doi.org/10.5858/srps.2014-0250-CP>.
- ◆ Badalato L, Kalokairinou L & Borry P. (2017) Third party interpretation of raw genetic data: an ethical exploration. *Eur J HumGenet*. 25: 1189-1194.
<https://doi.org/10.1038/ejhg.2017.126>.
- ◆ Borry P, Bentzen Hb, Budin-Ljosne I, Cornel Mc, Howard Hc, Feeney O, Jackson L, Mascalzoni D, Mendes A, Peterlin B, Riso B, Shabani M, Skirton H, Sterckx S, Vears D, Wijst M & Felzman H.

- (2018) The challenges of the expanded availability of genomic information: an agenda setting paper. *J Community Genet* 9: 103 a 116. <https://doi.org/10.1007/s12687-017-00331-7>.
- ◆ Cariasso M & Lennon G. (2012) SN pedia: a wiki supporting personal genome annotation, interpretation and analysis. *Nucleic acids Research*. Jan; 40 (Database issue): D1308-12. <https://doi.org/10.1093/nar/gkr798>.
 - ◆ Check Hayden E. (2015) Genome researchers raise alarm over big data. *Nature* 7 July. <https://doi.org/10.1038/nature.2015.17912>.
 - ◆ Chow-White, P.A., Mac Aulay, M., Charters, A. & Chow, P. (2015) From the bench to the bedside in the big data age: ethics and practices of consent and privacy for clinical genomics and personalized medicine. *Ethics and Information technology* 17(3): 189 a 200. <https://doi.org/10.1007/s10676-015-9373-x>.
 - ◆ Erlich Y, Shor T, Pe'er I & Carmi S. (2018) Identity inference of genomic data using long-range familial searches. *Science* Oct; 362: 690 a 694. <https://doi.org/10.1126/science.aau4832>.
 - ◆ Geiger S & Gross N. (2021) A tidal wave of inevitable data? Assetization in the consumer genomics testing industry. *Business & Society*. 60(3); 614 a 649. <https://doi.org/10.1177/0007650319826307>.
 - ◆ Greely Ht. (2020) The future of DTC Genomics and the law. *The Journal of Law, Medicine & Ethics*. Mar; 48 (1):151-160. <https://doi.org/10.1177/1073110520917003>.
 - ◆ Gürsoy G, Branson Ch M, Wagner S, Khanna A& Gerstein M. (2022) Storing and analyzing a genome on a blockchain. *Genome Biology* 23:134. <https://doi.org/10.1186/s13059-022-22699-7>.
 - ◆ Gymrek M, McGuire Al, Golan D, Halperin E & Erlich Y. (2013) Identifying personal genomes by surname inference. *Science* Jan 18; 339: 321 a 324. <https://doi.org/10.1126/science.1229566>.
 - ◆ Haeusermann T, Greshake B, Blasimme A, Irdam D, Richards M & Vayena Eff. (2017) Open sharing of genomic data: who does it and why. *PLoS ONE* 12(5): e 0177158. <https://doi.org/10.1371/journal.pone.0177158>.
 - ◆ Hume S, Nelson Tm, Spevak M, Mccready E, Agatep R, Feilotter H, Parboosinh J, Stavropoulos Dj, Taylor S & Atcockley Tl. (2019) CCMG practice guidelines: Laboratory guidelines for next-generation sequencing. *J Med Genet* 56: 792 a 800. <https://doi.org/10.1130/jmedgenet-2019-106152>.
 - ◆ Jones N. (2021) Scientists embrace digital crazefor non-fungible tokens. *Nature*, 594:481 a 482. <https://doi.org/10.1038/d41586-021-01642-3>.
 - ◆ Karaboue M & Lacasella Gu. (2025) Medicalisation in the post-genomic era. Transformation in identity, autonomy and ethics. *Clin Ter*; 176 (1). 109 a 113. <https://doi.org/10.7417/CT.2025.5173>.
 - ◆ Khan R & Mittelman D. (2018) Consumer genomics will change your life, whether you get tested or not. *Genome Biology* 19:120. <https://doi.org/10.1186/s13059-018-1506-1>.
 - ◆ Kahn SD. (2011) On the future of genomic data. *Science* Feb11; 331(6018): 728-9. <https://doi.org/10.1126/science.1197891>.

- ◆ Knoppers BM. (2010) Consentto “personal genomics” and privacy. Direct- to-consumer genetic tests and population genome research challenge traditional notions of privacy and consent. *EmboReports* 16(6): 416 a 419. <https://doi.org/10.1038/embor.2010.69>.
- ◆ Koplin Jj, Skeggs J & Gyngell Ch. (2022) Ethics of buying DNA. *Bioethical Inquiry*. 19: 395 a 406. <https://doi.org/10.1007/a11673-022-10192-w>.
- ◆ Kovanda K, Zimani An & Peterlin B. (2021) How to design a national genomic project-a systematic review of active projects. *Human genomics* 15:20. <https://doi.org/10.1186/s40246-021-00315-6>.
- ◆ Krum N & Hoffman N. (2020) Practical estimation of cloud storage costs for clinical genomic data. *Practical Laboratory Medicine* May 15: 20. <https://doi.org/10.1186/s40246-021-00315-6>.
- ◆ Kwon D. (2025) What went wrong at 23andMe? Why the genetic data giant risk collapse. *News Explainer* 23 January. <https://doi.org.bucm.idm.oclc.org/10.1038/d41586-025-00118-y>.
- ◆ Langmead B & Nellore A. (2018) Cloud computing as a platform for genomic data analysis and collaboration. *Nat Rev Genet* April; 19(4): 208 a 219. <https://doi.org/10.1038/nrg.2017.113>.
- ◆ Mc Cartney AM et al. (2024) Benefit sharing by design: A call to action for human genomics research. *Annu Rev Genom Hum Genet*; 25: 369 a 395. <https://doi.org/10.1140/annurev-genom-021623-104241>.
- ◆ Mcghee Ej & Milton S. (2023) FASTQ File compression benchmarking using lossless general purpose. *PEARC'23: practice and experience in advanced research computing 2023: Computing for the common good*: 315 a 319. <https://doi.org/10.1145/3569951.3597577>.
- ◆ Moscarello T, Murray B, Reuter C. M. & Demo E. (2019) Direct-to-consumer raw genetic data and third-party interpretation services: more burden than bargain? *Genetics in Medicine* Mar 21(3): 539 a 541. <https://doi.org/10.138/s41436-018-0097-2>.
- ◆ Moyer A et al. (2025) CAP laboratory improvement program. Laboratory considerations for releasing next generation sequencing data to patients. *Arch Pathol Lab Med* 149:152-158. <https://doi.org.10.5858/arpa.2023-0419-CP>.
- ◆ Mozilla Foundation. (2019) 23 reasons not to reveal your DNA. *Internet Health Report* pp. 31 a 35.
- ◆ Narayanasamy S, Markina V, Thorogood A, Blazkova A, Shahari M, Knoppers Bm, Prainsack B & Koesters R. (2020) Genomic sequencing capacity, data retention and personal access to raw data in Europe. *Front genet* May 16; 11:303. <https://doi.org/10.3389/fgene.2020.00303>.
- ◆ Pàmpols Ros, T, García Sagredo Jm, Pérez Aytés A, Díaz De Bustamante A. (2019) Pruebas genéticas de venta directa a los consumidores. Perspectiva de la Comisión de Ética de la Asociación Española de Genética Humana. *MedClin (Barc)* 153(1): 35 a 40. <https://doi.org/10.1016/j.medcli.2019.01.028>.
- ◆ Prainsack, B (2014) DIY Genetics: The right to know your own genome. In: Chadwick, R., Levitt, M. & Shickle, D. (eds). *The Right to Know and the Right Not to Know: Genetic Privacy and Responsibility*. Cambridge University Press pp. 100 a 115. <https://doi.org/10.1017/CBO9781139875981.010>.

- ◆ Rehder C et al. (2021) Next generation sequencing for constitutional variants in the clinical laboratory, 2021 revision. A technical standard of the American College of Medical genetics and Genomics (ACMG) and its supplementary material. *Genetics in Medicine* 23: 1399 a 1415. <https://doi.org/10.1038/s41436-021-01139-4>.
- ◆ Rehm Hl, Bale Sj, Bayrak-Toydemir P, Berg Js, Brown Kk, Deignan Jl, Friez Mj, Funke Bh, Hedge Mr & Lyon E. (2013) ACMG clinical laboratory standards for next-generation sequencing. *Genet Med Sep*; 15(9): 733 a 747. <https://doi.org/10.1038/gim.2013.92>.
- ◆ Regalado A. (2019) More than 26 million people have taken an at-home ancestry test. *MIT Technology Review*. February 11.
- ◆ Romeo Malanda, S. (2013) Análisis genéticos directos al consumidor: su régimen Jurídico en el ordenamiento jurídico español y propuestas de actuación. En Carlos María Romeo Casabona (Ed) *Hacia una nueva medicina: consejo genético*. Publicaciones de la Cátedra universitaria, Diputación foral de Bizkaia de derecho y genoma humano y Editorial Comares. Bilbao – Granada.
- ◆ Rosenbaum E. (2018) 5 biggest risks of sharing your DNA with consumer genetic-testing companies. *CNBC Disruptor/50*.
- ◆ Rutherford A. (2025) As a geneticist I will not mourn 23andMe and its jumble of useless health information. *The guardian* 27 Mar.
- ◆ Schickhardt C, Fleischer H & Winkler Ec. (2020) Do patients and research subjects have a right to receive their genomic raw data? An ethical and legal analysis. *BMC Medical Ethics*. 21:7. <https://doi.org/10.1186/s12910-020-0446-y>.
- ◆ Shahani M & Borry P. (2015) Challenges of web based personal genome data sharing. *Life Sciences Society and Policy* 11:39. <https://doi.org/10.1186/s40504-014-0022-7>.
- ◆ Shahani M, Vears D & Borry P. (2018) Raw genomic data: Storage, Acces and Sharing. *Trends in genetics January*; 34(1): 8 a 10. <https://doi.org/10.1016/j.tig.2017.10.004>.
- ◆ Souche E et al. (2022) Recommendations for whole genome sequencing in diagnosis for rare diseases. *Eur J Hum Genet* 30: 1017 a 1021. <https://doi.org/10.1038/s41431-022-01113-x>.
- ◆ Stephens Zd, Lee Sy, Faghri F, Campbell Rh, Zhai C, Efrom Ms, Lier R, Schatz Mc, Sinha S & Robinson E. (2015) Big data: Astronomical or Genomical? *PLOSBiology July* 7. <https://doi.org/10.1371/journal.pbio.1002195>.
- ◆ Van El Cg, Cornel Cc, Borry P, Hastings Rj, Fellmann F, Hodgson Sv, Howard Hc, Cambon-Thomsen A, Knopppers Bm, Meijers-Heijboer H, Scheffer M, Tranebjærg L, Dondorp W & De Wert Gmwr (2013) Whole genome sequencing in health care. *Eur J Hum Genet* 2013; 21: 580 a 584. <https://doi.org/10.1038/ejhg.2013.46>.
- ◆ Vears Df et al. (2023). A practical Check list for return of results from genomic research in the European context. *Eur J Hum Genet*; 31. 687 a 695. <https://doi.org/10.1038/s41431-023-01328-6>.
- ◆ Wheeler Da, Srinivasan M, Egholm M, Shen Y, Chen L, Mc Guire A, He W, Chen Y, Makhijara V, Roth Gt, Gomes X, Tartaro K, Niuzi F, Turcotte Cl, Irzyk Gp, Lupski Ja, Chinault C, Song X, Liu Y & Rothberg Jm. (2008) The complete genome of an individual by massively parallel DNA sequencing. *Nature*; 452: 872 a 877. <https://doi.org/10.1038/nature06884>.

- ◆ Wiews J *et al.* (2024). Toward realizing the promise of AI in precision health across the spectrum of care. *Annu Rev Genom Hum Genet*; 25: 141 a 159. <https://doi.org/10.1146/annurev-genom-010323-010230>.
- ◆ Winkler Ec, Kai C, Idler I, *et al.* (2019). On the release of raw genomic data to patients and study participants. EURAT project group “Ethical aspects of whole genome sequencing”. Position paper on the release of raw genomic data to patients and study participants. Heidelberg. December 2019.
- ◆ Winkler E C & Knoppers Bm. (2022). Ethical challenges of precision cancer medicine. *Seminars in cancer biology*; 84: 263 a 270. <http://doi.org/10.1016/j.semcan.2020.09.009>.
- ◆ Wolf S & Green R. (2023) Return of results of genomic research using large scale or whole genome sequencing toward a new normal. *Ann Rev Genomics Hum Genet*. 2024:393 a 414. doi:10.1146/annurev-genom-101122-103209.
- ◆ Wright C, Burton H, Hall A, Moorthie S, Pokorska-Bocci A, Sagoo G, Sanderson S & Skinner R. (2011) Next steps in the sequence. The implications of whole genome sequencing for health in UK. PHG Foundation.

Fecha de recepción: 10 de febrero de 2025.

Fecha de aceptación: 29 de mayo de 2025.

Fecha de publicación: 16 de junio de 2025.