



UNIVERSIDAD DE JAÉN
Facultad de Ciencias Experimentales

Trabajo Fin de Grado

El ADN antiguo y su aplicación al estudio del pasado humano

Alumno: Daniel Vega Meco

Jaén, Octubre, 2021



**UNIVERSIDAD DE
JAÉN**



Trabajo Fin de Grado

El ADN antiguo y su aplicación al estudio del pasado humano

Firma del alumno:

Alumno: Daniel Vega Meco

Jaén, Octubre, 2021

Índice	Páginas
RESUMEN	1
ABSTRACT	1
1.INTRODUCCIÓN	2
1.1. <i>La molécula de ADN: características</i>	2
1.2. <i>Genoma Humano</i>	4
1.2.1. <i>Proyecto Genoma Humano: resultados (genes, regiones que no son genes, SINEs, LINES)</i>	4
1.2.2. <i>Variación genética entre individuos: SNPs, CNVs</i>	6
2. OBJETIVOS	9
3. MATERIAL Y MÉTODOS	10
4. RESULTADOS	11
4.1. <i>ADN antiguo: características</i>	11
4.1.1. <i>¿Qué es el ADN antiguo?</i>	12
4.1.2. <i>Problemáticas en el estudio del ADN</i>	12
4.1.2.1 <i>Degradación del ADN</i>	12
4.1.2.2. <i>Contaminación del ADN</i>	13
4.1.2.3. <i>Inhibición por PCR</i>	14
4.2. <i>Metodologías de secuenciación del ADN antiguo</i>	16
4.3. <i>Aplicaciones de la ADN</i>	19
4.3.1. <i>Aplicación del ADN antiguo en el estudio del pasado humano remoto: Paleogenética.</i>	19
4.3.2. <i>Aplicación del ADN antiguo en el estudio del pasado humano reciente: Arqueogenética</i>	27
5. CONCLUSIONES	32
6. BIBLIOGRAFÍA	33

RESUMEN

La genética es una rama esencial para el estudio de nuestros antepasados y que nos ayuda a entender los procesos que se han producido a lo largo de la historia. Gracias al Proyecto de Genoma Humano actualmente podemos hacer comparaciones de los genomas de los humanos modernos actuales, con los genomas de las poblaciones antiguas, es decir, con aquellas que se van encontrando con cada expedición. El ADN antiguo es un componente esencial del cual se extrae la información genética necesaria que se necesita para el emprendimiento de los aquellos estudios que están relacionados con la Arqueogenética y la Paleogenética. Por otra parte, tenemos que resaltar lo frágil que es este componente y la dificultad que tiene trabajar con él bajo diferentes condiciones. En esta revisión se analizarán y se explicarán los resultados de los estudios que hayan sido relevantes para la historia y la evolución, así como la notable función que tiene el ADN antiguo.

Palabras clave: ADN antiguo, Arqueogenética, Genoma, Información genética, Paleogenética, Proyecto de Genoma Humano.

ABSTRACT

Genetics is an essential branch (of biology) for the study of our ancestors, therefore it helps us to understand the processes that have occurred throughout history. Currently, the Human Genome Project allows us to draw comparisons between genomes of contemporary and ancient populations, in other words, with those that are encountered with each expedition. Ancient DNA is an essential component to extract required information for studies related to Archeogenetics and Paleogenetics. On the other hand, we have to highlight how fragile this component is and how difficult is to work with it under different conditions. This review will analyze and explain the results of studies that have been relevant to history and evolution as well as the remarkable role of ancient DNA.

Keywords: Ancient DNA, Archaeogenetics, Genetic Information, Genome, Human Genome Project, Paleogenetics.

1. INTRODUCCIÓN

La genética es una ciencia que se ocupa de la herencia y de aspectos relacionados con la misma. Para realizar cualquier investigación genética, el conocimiento de las moléculas responsables de la herencia es esencial. Hoy en día, ya sabemos que la molécula de ADN (ácido desoxirribonucleico) (DNA: Desoxirribonucleic acid), la encontramos integrada en el núcleo celular (ADN nuclear) y también en algunos orgánulos celulares como es el de las mitocondrias (ADN mitocondrial). La estructura del ADN está dividida en varios niveles de complejidad. La estructura primaria de este consta de una sucesión de unidades que están repetidas linealmente, las cuales se denominan nucleótidos. Cada nucleótido consta de tres componentes fundamentales: un azúcar denominado desoxirribosa, un grupo fosfato y una base nitrogenada. En cada nucleótido los dos primeros componentes mencionados anteriormente se mantienen, pero en el caso de la base nitrogenada varía, pudiéndose encontrar cuatro bases nitrogenadas diferentes, las cuales se denominan: Adenina (A), Guanina (G), Citosina (C) y Timina (T). Una secuencia de ADN viene determinada por su sucesión de bases y es esa secuencia de nucleótidos lo que va a definir el genotipo del individuo, que, junto con otros factores externos, formarán el fenotipo, Peña-Castro, J. M. *et al.*, (2013).

1.1. La molécula de ADN: características

A continuación, se describe de forma breve el modelo simple del ADN en el cual es posible visualizar varios hechos importantes [Fig.1]:

- La estructura en forma de doble hélice del ADN es muy regular y se extiende por toda la longitud de la molécula, pudiendo llegar hasta los 10 cm.
- Está formada por dos hebras complementarias una de la otra en disposición antiparalela, además ambas pueden portar codificación genética, una sirve de molde para la transcripción, a esta hebra se le denomina “hebra molde” o “antisentido”, mientras que a la otra se le denomina “de sentido”, cuya secuencia sería igual al del ARNm.
- Las dos hebras se unen mediante los denominados “puentes de hidrógenos” producidos por las bases nitrogenadas de cada hélice.

Estos puentes de hidrógeno son débiles, pero juntos y con un gran número de ellos se consigue unir ambas hebras. También cabe destacar que la inestabilidad de estos ayuda a que se puedan separar fácilmente las hebras.

- El esqueleto de las hebras está formado por iones fosfatos que estos a su vez forman uniones con dos desoxirribosas contiguas.
- Las dos cadenas se encuentran próximas entre sí a lo largo del ADN, por lo cual forman un surco grande o también conocido como hendidura mayor con un tamaño de 22 Å y un surco pequeño o hendidura menor de 12 Å. Gracias a estos surcos las hebras se pueden comunicar con las proteínas que regulan tanto su transcripción como el reconocimiento de señales. Esta molécula tiene una estructura uniforme, es decir, por cada vuelta completa hay el mismo número de pares de bases, que son 10,5. La distancia entre pares de bases es de 0,34 nm y la vuelta completa mide 3,4 nm o lo que es lo mismo, 34 Å.
- Finalmente, el grado de torsión de las hélices no es absolutamente fijo, varía según la secuencia de bases que forman cada segmento. Esto es importante a la hora de influir sobre la capacidad transcripcional, Crick y Watson, (1954).

Todas estas propiedades hacen referencia a la denominada “forma B” del ADN.

La estructura se compacta y alcanza una mayor complejidad, en forma de cromatina, un complejo formado por ADN y por un octámero de proteínas histonas, estas se unen al ADN, ayudan a dar su forma a los cromosomas y ayudan a controlar la actividad de los genes, todo esto constituye la estructura terciaria. La cromatina se condensa más aún, hasta que llega a un nivel máximo de compactación en forma de cromosoma, Romero R. G *et al.*, (2011).

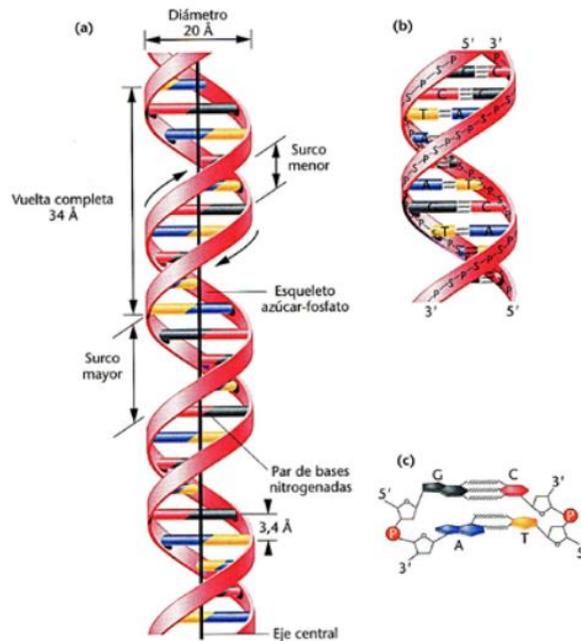


Figura 1. Esquema de la forma clásica del ADN. (Imagen recuperada de María Victoria, A. L., 2014)

1.2. Genoma Humano

1.2.1. Proyecto Genoma Humano: resultados (genes, regiones que no son genes, SINES, LINES)

En 2003, gracias al Proyecto del Genoma Humano se pudo “leer” la secuencia concreta de nucleótidos que componen el ADN humano. Aún quedan muchos interrogantes por resolver. Este proyecto tuvo como objetivo el de identificar la totalidad o el máximo número de genes, anotar su localización y averiguar qué función llevaban a cabo. Este proyecto tan importante se inició en 1990, en la cual participaron 18 países y fué financiado con fondos públicos. En él participan científicos de Gran Bretaña, Estados Unidos, Japón, Alemania, China y Francia principalmente, para obtener el objetivo mencionado anteriormente, la secuenciación del genoma también se llevó a cabo por otra empresa privada estadounidense denominada Celera Genomics, esta contaba con una importante aportación económica y en avances científicos. Ambas empresas realizaron las secuenciaciones se hicieron por el método de Sanger, Celera tardó solo tres años en realizar dicha secuenciación con un costo bastante menor que el consorcio público, ya que también disponía de los datos del consorcio público que tenía que publicarlos en menos de 24h. El consorcio público dividió su proyecto en tres fases: Fase I, crear genotecas genómicas, Fase II, mapeo

de las genotecas BAC y Fase III, secuenciación y ensamblaje, Venter, J. C *et al.*, (2001). En el caso de Celera lo que hizo fue saltarse la fase II, por lo que no realizó el mapeo físico de los clones BACs, lo que suponía el desconocimiento de que parte del genoma estaban secuenciando, se fragmentó el ADN genómico completo y secuenció al azar todos los fragmentos. Una vez que los secuenciaron, los ensamblaron (whole-genome shot gun sequencing) [Fig.2], Pennisi, E. (2001).

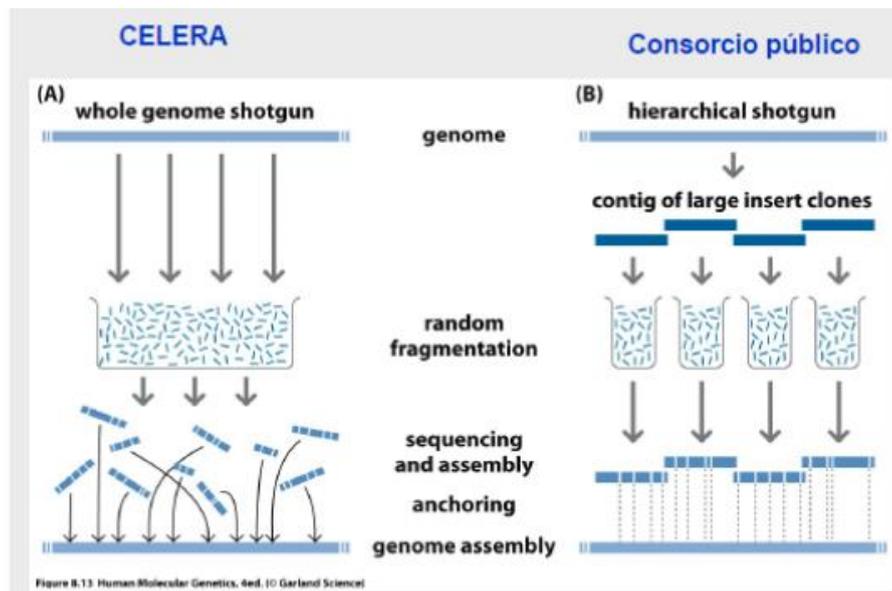


Figura 2. Ejecución de la secuenciación del Genoma Humano, a la izquierda se puede observar cómo lo realizó Celera Genomics y a la derecha como lo realizó el consorcio público. (Tomado del libro "Human Molecular Genetic", 4ed, Andrew P. Rea, 2010)

El genoma humano con sus 46 cromosomas, inicialmente se pensó que estaba integrado por 100,000 genes codificadores de proteínas, enzimas, hormonas y en general reguladores de todos los procesos vitales y características del ser humano. El cromosoma 22 fue el primero en ser identificado completamente contando con 545 genes activos, seguido posteriormente por el mapeo o secuenciación genética del cromosoma 21 con solo 225 genes activos, el cual abrió las posibilidades para desarrollar tratamientos genéticos en enfermedades relacionadas, como el Síndrome de Down, una forma de Alzheimer y varios tipos de cáncer, Henríquez, O. (2001).

Se considera que el número de genes por el que está compuesto el genoma humano es de unos 2.800, y el rango de error se vuelve cada vez menor. Se ha podido identificar unos 26.383 genes y se ha observado que la distribución es asimétrica

El genoma humano contiene más del 50% de secuencias repetidas, entre ellas incluimos: repeticiones intercaladas derivadas de elementos transponibles, y largas regiones genómicas que se han duplicado en tándem, palindrómicas o dispersas (se hablarán de ellas más adelante). Entre las secuencias repetidas mencionadas anteriormente, también se encuentran grandes segmentos duplicados de unos 50-500 kb con un porcentaje de 98-99,9% parentesco entre secuencias, en los que el desajuste durante la recombinación crea deleciones responsables de síndromes / enfermedades genéticas. Estas características son las responsables de que no se pueda dar un buen ensamblaje de una secuencia genómica correcta y acabada, además de complicar el proceso de ensamblaje, Lander E. S *et al.*, (2001).

Por otro lado, sabemos que el genoma humano es un mosaico de isocoros, un isocoro, región de ADN con un tamaño mayor de 300 kb y que son homogéneos a su composición de G (guanina) y C (citosina). Se sabe que actualmente el tamaño del genoma humano oscila entre 3.400 Mb (3 x 10⁹ pb), distribuidos en 23 cromosomas y codifica para aproximadamente 20.000 o 25.000 genes, está formado por 35.000 genes y tiene una densidad génica de 10 Mb, Lamolle G *et al.*, (2018).

Otra característica del genoma humano, son las regiones conocidas como “desiertos génicos”, que como su nombre indica son pobres en genes, es decir, contienen segmentos de ADN de más de media megabase en promedio dónde no se localizan genes. En el genoma humano estos desiertos génicos suponen un 20% del total (unas 605 megabases), y se localizan principalmente en los cromosomas 4, 13, 18 y X, Solari, A. J., (2004).

1.2.2. Variación genética entre individuos: SNPs, CNVs

La genética humana se ocupa del estudio de la especie *Homo sapiens*. Gracias a esta rama de la genética aprendemos sobre la variación genética humana y la transmisión de la variabilidad que contribuye inevitablemente a una mejor comprensión de la genética en general, del mismo modo que el estudio de la variación en otras especies nos ayuda a comprender la nuestra. Los análisis realizados hasta la fecha indican que en promedio la secuencia de nucleótidos del genoma de dos individuos cualesquiera es idéntica al 99.9%. Esto nos indica que el genoma humano presenta una variabilidad

muy baja entre individuos, así las variaciones en la secuencia entre las personas analizadas son menores a una base por cada mil, 1000 Genomes Project Consortium (2010).

Si cogiéramos a dos seres humanos sin parentesco, observaremos que la variación genética entre ambos sería de aproximadamente el 0,1%. Este resultado nos indica que aproximadamente un par de bases de cada 1.000 será diferente entre dos individuos, National Institutes of Health (US) *et al.*, (2007). Debemos tener en cuenta que las variaciones en el genoma pueden ser transmitidas a la siguiente generación si afectan a las líneas germinales, sin embargo, si afectan a las líneas somáticas no son transmitidas, Francisco Javier N. (2007). Existen de varios tipos de variaciones, por su importancia comentaremos aquí las dos principales: SNPs (Single Nucleotide Polymorphisms), polimorfismos de un solo nucleótido, y CNVs (Copy Number Variations), regiones en número variable de copias.

En relación con los SNPs, los estudios de genotipado de estos a gran escala se basan en una evaluación inicial de la variación de nucleótidos para identificar los sitios de la secuencia de ADN que albergan variación entre los individuos, ya que los SNPs son un tipo de variación más simple respecto al cambio, que es de un solo nucleótido, y se ha comprobado que las propiedades de los SNP que se encuentran están influidas por el esfuerzo de muestreo de descubrimiento, Clark, A. G *et al.*, (2005). Dado que somos organismos diploides, y que para la mayoría de los SNPs solo hay dos alelos (C ó T, por ejemplo), eso implica que hay tres posibles genotipos para cada uno de ellos. La frecuencia promedio en el genoma es de 1 por cada 1.200 nucleótidos. Éstos no se encuentran distribuidos al azar, la mayoría de estos se encuentran en intrones, ya que ocupan más espacio que las regiones codificantes. También podemos encontrarlos en regiones de ADN intergénicas y, también en las regiones codificantes. En estas últimas zonas, los SNP pueden provocar variaciones en la secuencia de los aminoácidos codificados, por tanto, los SNP son útiles para comparar variaciones genéticas entre individuos y en poblaciones, Fredman, D *et al.*, (2004).

Cabe destacar que en el genoma humano podemos encontrarnos grandes cantidades de los denominados elementos genéticos móviles, estos están distribuidos al azar y ocupan el 35% del genoma, son secuencias de ADN discretas, que como su nombre indica tienen una alta capacidad de transportarse o duplicarse en diferentes regiones del genoma. Este proceso se ha producido en la mayoría de los organismos desde hace millones de años. Gracias a su manera única de desplazarse en el genoma, los

elementos móviles son muy útiles para rastrear las relaciones entre los individuos, las poblaciones y las especies. Existen dos tipos: las cortas denominadas SINES (Short Interspersed Nuclear Elements), estas ocupan un total de 288 megabases, (10% del genoma), se forman como resultado de retroposiciones de un ARN especial. El SINE principal es la familia de elementos *Alu*, es un elemento rico en Guanina y Citosina, se suele localizar en las bandas R de los cromosomas humanos, Francisco Javier N. (2007), y las largas, conocidas como LINE (Long Interspersed Nuclear Elements) y ocupan 446 megabases (20% del genoma) algunos de estos se han formado por duplicaciones sucesivas a partir de un único gen, Xing, J *et al.*, (2007)

Existen también secuencias que son prácticamente idénticas a genes codificantes, pero carecen de funcionalidad, es decir, no codifican o expresan ningún fenotipo. Se les denomina pseudogenes y son difíciles de diferenciarlas de aquellas que sí son funcionales. Los pseudogenes se originan a partir de la descomposición de genes que se originaron por duplicación a través de la evolución, su replicación a lo largo de generaciones es un proceso bioquímico muy costoso a nivel de energía, pero la selección natural permite esta replicación, ya que las duplicaciones de genes hacen la divergencia funcional y esto conlleva a que se generen nuevos genes. Los pseudogenes no procesados tienen intrones y secuencias reguladoras, y su expresión está paralizada por codones de parada. Las copias extra de los genes funcionales acumulan mutaciones, y esto mantiene el gen original funcional. La duplicación del gen puede dar lugar a un nuevo gen con una función completamente diferente y esto lleva a un aspecto interesante sobre que algunas de ellas sí podrían tener funciones de regulación de la actividad de otros genes. Un ejemplo de regiones duplicadas durante la evolución son los denominados genes HOX, que participan en el desarrollo embrionario, estos provienen de un gen original denominado "parálogos", Tutar, Y. (2012).

El que no se encuentren muchos SNPs en los exones se puede explicar, ya que la selección natural tiende a eliminar estos polimorfismos de los exones cuando alteran de manera negativa la función génica. Se analizaron los SNPs y las variaciones de secuencias repetidas para constatar la identidad genética de distintas poblaciones humanas, en las que ya existen evidencias de diferencias en la frecuencia de determinados alelos de algunos genes como, por ejemplo, los grupos sanguíneos. Sin embargo, la especie humana es notablemente homogénea en lo que respecta a su genoma, Solari, A. J., (2004). Por tanto, la gran ventaja que pueden ofrecernos los

SNPs sobre los demás marcadores, es que además de ser muy abundantes y estar repartidos de manera uniforme por el genoma, es su análisis a gran escala mediante la utilización de los microarrays, Francisco Javier N. (2007).

En relación con los polimorfismos CNVs (Variaciones en número de copia) es de gran importancia, ya que contribuye a la variabilidad del genoma entre los diferentes individuos por sus deleciones y duplicaciones. Las CNVs se definen como secuencias de ADN mayores a 1000 pares de bases que se encuentran repetidas en tándem, y para las que existen diferencias entre individuos en relación al número de copias presentes. La existencia de estas variaciones a nivel de individuo o de poblaciones podrían tener impacto fenotípico de importancia, dado que pueden contener genes en su interior, lo que daría lugar a cambios en los niveles de expresión de los mismos. Las regiones que componen las CNVs se estima que ocupan hasta un 10% en el genoma. Estas se distribuyen de forma desigual en el genoma; las regiones pericentroméricas y subteloméricas de los cromosomas muestran una tasa de variación especialmente alta. Al contrario de los SNPs que abundan a lo largo del genoma, los CNVs son menos numerosos, esto se debe porque para considerarlos CNVs como tal se requieren que estas repeticiones sean de varios miles de pares de bases. El conjunto de CNVs descritas hasta la fecha en el genoma humano están recopiladas en una base de datos - "The Database of Genomic Variants" - que contienen datos de cerca de 4.878 loci, comprendiendo alrededor de 11.748 diferentes CNVs que han sido identificadas a lo largo del genoma. En la actualidad se está estudiando la asociación entre la variación de CNVs con características fenotípicas o patologías. La afectación del genoma por las CNVs se relaciona con la dosis génica y las interacciones entre genes, Rosado *et al.*, (2019).

2. OBJETIVOS

- Indagar sobre el conocimiento científico descrito acerca del ADN antiguo.
- Investigar su aplicación tanto en la Paleogenética como en la Arqueogenética.
- Revisar las diferentes investigaciones que se han realizado con la utilización del ADN antiguo, así como sus resultados y sus conclusiones.
- Ampliar el conocimiento de la Genética Humana en base a las investigaciones realizadas recientemente.

3. MATERIAL Y MÉTODOS

Para encontrar la información y poder llevar a cabo la realización de este trabajo, ha sido necesaria la consulta de páginas web científicas, así como de bibliografías de asignaturas específicas del grado, de artículos de investigación y de libros proporcionados por la propia universidad. Estos contienen gran información sobre las características del ADN antiguo, así como del Genoma Humano y la aplicación de ambos en las distintas investigaciones que se han realizado y su utilidad tanto en el pasado, presente y futuro. Los métodos de implementación incluyen, por otro lado, la selección de artículos científicos a través del motor de búsqueda como:

- **Google académico “Google Scholar”**: Este es un motor de búsqueda especializado que pertenece a Google, aquí podremos encontrar una gran cantidad de documentos académicos, así como revistas científicas, resúmenes y una gran cantidad de citas correspondientes a muchas disciplinas y fuentes, a su vez este permite el acceso a varias fuentes desde un solo lugar. El enlace para poder acceder a este buscador: [Scholar.google.es](https://scholar.google.es).

- **Pubmed**: Este sistema de búsqueda es un proyecto desarrollado por el Centro Nacional de información Biotecnológica (NCBI) de la Biblioteca Nacional de Medicina (NLM). Permite el acceso a la base de datos bibliográficos por NLM. También nos ha proporcionado citas y resúmenes de artículos científicos de nuestro interés, junto a bibliografías y actas de congresos. Otra manera de acceder a este sistema de información es mediante el siguiente enlace: pubmed.ncbi.nlm.nih.gov.

Con la palabra clave «Ancient DNA» encontramos aproximadamente 1.270.000 resultados, filtramos estos resultados, cambiando la búsqueda a «evolution in Ancient DNA» y seleccionamos que los artículos fueran desde el 2.000 hasta la actualidad y 452.000 obtuvimos resultados, tuvimos que afinar aún más la búsqueda, por tanto, esta última vez iniciamos la búsqueda con «human ancient DNA paleogenomics», con el resultado de 71 artículos. Debido a la cantidad de artículos que aparecen en la búsqueda bibliográfica, hemos tenido que seleccionar aquellos artículos en revistas

de alto impacto que consideramos que aportan una información valiosa para comprender nuestro genoma, así como su evolución y cambios que se han producido en ella y apreciar la valiosa utilidad del ADN antiguo.

4. RESULTADOS

4.1 ADN antiguo: características

Una vez mencionado las variaciones genéticas entre individuos, cabe mencionar que el ADN antiguo es una fuente importante de información sobre nuestro pasado, es necesario conocer sus mecanismos y funcionamiento para poder usarlo, y también tener en cuenta sus limitaciones descritas de forma abreviada anteriormente, Díez *et al.*, (2012). Una definición de ADN antiguo es: ADN que se ha recuperado de muestras biológicas que no se han conservado de manera específica para realizarles análisis a posteriori de ADN. En este campo cabe destacar a Svante Pääbo considerado como uno de los fundadores de la paleogenética, Saiz, M *et al.*, (2012).

También se puede emplear el ADN antiguo para el seguimiento de migraciones históricas o prehistóricas, conocida como “arqueogenética” en la que se aplican estudios a restos humanos de finales del Paleolítico, Neolítico y época antigua (10.000-15.000 años hasta la actualidad). Estos estudios nos están permitiendo conocer aspectos nuevos o diferentes sobre las diversas culturas antiguas, las cuales chocan en muchos casos con las teorías que habían sido propuestas desde el punto de vista de la arqueología. Un caso a destacar es la civilización de los Yamnaya y su gran expansión por Europa, por lo que se conoce como uno de los más espectaculares.

La genética nos permite conocer cómo de parecidos o diferentes eran los individuos de las poblaciones y cómo se desplazaron por el planeta, ya que se pensaba que muchas de estas civilizaciones permanecieron sedentarias en un mismo lugar, lo que se demostró que en realidad fueron reemplazadas por otras, las cuales asumieron gran parte de sus ritos y culturas. Al tratarse de un campo bastante reciente, hay muchos conflictos entre los arqueólogos y los genetistas, aunque también hay grupos nuevos de investigación multidisciplinarios que están dando grandes avances, además cabe destacar que hay diversos estudios que se han enfocado en la Península ibérica, Turbón, D *et al.*, (2003).

4.1.1. ¿Qué es el ADN antiguo?

El “ADN antiguo”, también conocido como ADN_a, de forma abreviada, debido a su interés antropológico, arqueológico y público por los restos humanos, es uno de los tipos de muestra más estudiados con diferentes finalidades, como la búsqueda de nuevas evidencias sobre el origen del hombre, o el estudio de su migración, mezclas de poblaciones, costumbres, etc. Hay quien entiende esto como ácido desoxirribonucleico que ha sido copiado y secuenciado a partir de especies extintas o de especímenes de poblaciones, Fraile, I. R., (2007).

Algunos ejemplos incluyen el análisis de ADN recuperado de material esquelético arqueológico o histórico, tejidos momificados, muestras de colección no congelados, muestras médicas, especímenes de herbarios y restos congelados en el hielo y “permafrost”, Fraile, I. R., (2007).

Como sabemos en los últimos años, el estudio de este ADN en restos humanos se ha convertido en un importante campo de investigación, este aporta evidencias que apoyan la teoría evolutiva de nuestra especie, Schmitz *et al.*, (2002); Caramelli *et al.*, (2003) o cómo descifrar la secuencia de nucleótidos Neanderthal.

Estos estudios basados en este ADN apoyaron la reconstrucción de las relaciones humanas y establecieron un flujo genético desde África a las colonias del Nuevo Mundo, también demostraron que el ADN que sobrevivía en los restos antiguos era a menudo principalmente de origen microbiano o fúngico y que el ADN endógeno se limitaba generalmente a concentraciones muy bajas de fragmentos cortos y contenían daños en locis como el ADN mitocondrial (ADN_{mt}), Willerslev *et al.*, (2005).

4.1.2. Problemáticas en el estudio del ADN_a

Como sabemos el ADN_a tiene una serie de características físicas y químicas que lo distinguen de los organismos vivos, la suma de estas impone ciertas limitaciones técnicas a la hora de recuperar ADN de restos antiguos.

4.1.2.1 Degradación del ADN

Este factor depende de los cambios a la exposición a la luz, la humedad y la temperatura. La contaminación fúngica y bacteriana provocaría la degradación física, química y bioquímica de este. En las células vivas, la integridad del ADN se mantiene

mediante procesos de reparación enzimática. Después de que el organismo muere, el comportamiento de las células que contienen la enzima catalítica se desintegrará. Como resultado, estas enzimas degradan rápidamente al ADN. Además, el ADN está sujeto a procesos de degradación provocados por bacterias, hongos e insectos. La exposición al agua puede ser la fuerza más dañina para el ADN.

Se ha demostrado que el agua inicia la rotura de la cadena de ADN al formarse enlaces glicosídicos. Por tanto, el paso clave para la conservación del ADN es la rápida deshidratación de los tejidos. También cabe mencionar que el pH del medio ambiente juega un papel muy importante en la preservación a largo plazo de esta molécula. La exposición a la luz ultravioleta puede dañar gravemente el ADN y degradarlo, limitar la exposición de las muestras a la luz ultravioleta (como un entierro rápido) es importante para minimizar las consecuencias del daño ultravioleta, por estos motivos existe una gran dificultad de encontrar muestras de ADN de calidad en aquellas zonas tropicales húmedas o con excesiva radiación como es el desierto, Saiz, M *et al.*, (2012).

4.1.2.2. Contaminación del ADN

Hay tres potenciales de contaminación: muestras de ADN contaminadas por el medio ambiente, contaminación debida a la transferencia de ADN entre muestras mediante la preparación de la muestra y, la causada por productos del ADN amplificado. La primera depende principalmente de la recogida de la muestra y del cuidado que haya tenido el equipo que la ha tomado. La contaminación debida a la transferencia de ADN entre muestras mediante la preparación de la muestra y la causada por productos del ADN amplificado pueden ser fácilmente controladas mediante procedimientos de laboratorios apropiados y la separación de áreas de trabajo. Además, es muy importante intentar minimizar la contaminación de la muestra en el momento de su toma por parte del personal arqueológico. Igualmente, la posible contaminación ha de ser monitorizada mediante controles negativos (muestras que contienen todos los reactivos excepto ADN) durante la extracción y la amplificación que se procesan del mismo modo que los indicios de ADN, Turbón D, *et al.*, (2003).

Dado que la cantidad de ADN que encontramos en la muestra es limitada, habrá complicaciones en las investigaciones de cualquier tipo de estudio con ADN. La

fuerza de contaminación varía mucho según el tipo de material arqueológico y la investigación a realizar, pudiendo extraerse ADN exógeno y ADN contaminante, Turbón D, *et al.*, (2003).

También es probable que ambos puedan ser coamplificados o en el peor de los casos amplificado, preferentemente el ADN contaminante. Si no se toman medidas especiales al manipular muestras dañadas, la alta sensibilidad de la PCR y su capacidad para amplificar pequeñas cantidades de ADN puede causar problemas. Se debe seguir un protocolo validado en el laboratorio para evitar la contaminación de muestras con grandes cantidades de ADN. La contaminación implica la transferencia accidental de ADN, Turbón D, *et al.*, (2003).

4.1.2.3. Inhibición por PCR

En los primeros estudios con ADN se observó que la fuente para conseguir muestras con las cuales se pudieran trabajar bien procedían de tejidos blandos, dientes y huesos. Más tarde, con el uso de la PCR en relación con el ADN, se pudo observar una relación inicial con la presencia (algunos de estos extractos) de sustancias que inhiben esta acción, Turbón D, *et al.*, (2003).

Existe un protocolo de extracción convencional del ADN denominado “método Fenol-Cloroformo”, Turbón D, *et al.*, (2003). Este método es laborioso y lento, ya que es la mejor opción para que a la hora de extraer nuestra muestra sufra el menor daño posible y podamos trabajar adecuadamente con él.

Existen estrategias para poder eliminar o mitigar la inhibición, ya sea aumentando la pureza del ADN extraído o bien tratando de contrarrestar este efecto inhibitorio durante la reacción de amplificación. La mayoría de estas medidas resultan en un aumento del riesgo de contaminación con ADN exógeno de la muestra, ya que aumentan la manipulación por parte del investigador, Turbón D, *et al.*, (2003).

No obstante, existen otras que producen mejoras sustanciales en la amplificación y que no presentan esta desventaja, como es el caso de almacenar los extractos en frío durante varios días anticipadamente a su amplificación. Este procedimiento produce la formación de un precipitado blanquecino que queda adherido a la pared del tubo y el cual se elimina por precipitación que consigue revertir la capacidad de amplificación,

Turbón D, *et al.*, (2003).

La naturaleza del inhibidor aún está por precisar, aunque se han propuesto un grupo de moléculas candidatas, fundamentalmente compuestos del suelo o productos degradados del organismo. A continuación, citaremos los más destacados, Turbón D, *et al.*, (2003):

- **Ácidos húmicos y fúlvicos:** La primera persona en hablar sobre la acción inhibidora de estas sustancias fue Svante Pääbo. Son una familia de moléculas que se encuentran con abundancia en el suelo y que pueden estar acompañando al ADN en el proceso de purificación y que son inhibidores muy potentes de la reacción de PCR.

- **Residuos de porfirinas o productos derivados de su degradación:** La acción inhibidora de estas moléculas es debida a su capacidad para secuestrar iones metálicos, debido a estructuras similares al grupo hemo. Las porfirinas se encuentran en diferentes estructuras como en la sangre, tejidos blandos y en hojas vegetales. Las porfirinas que más se encuentran son las protoporfirinas, en la que destacamos la protoporfirina IX, presente en la naturaleza. Esta protoporfirina tiene capacidad de quelar iones magnesio, los cuales se necesitan para el funcionamiento de la Taq polimerasa, por lo que hace que presente un gran poder inhibidor de la acción de PCR potencial. Igualmente, los derivados porfirínicos suelen formar agregados que quedan retenidos junto al ADN en la membrana de muchos procedimientos de filtración, como en el caso del Centriplus 30000, que se emplea durante la extracción de ADN antiguo.

- **Productos de Maillard:** consiste en una reacción de un grupo carbonilo con uno amino para dar compuestos carbonílicos y deoxisonas, en el caso de estos últimos cuando hay azúcares que están involucrados en la reacción. La coloración marrón que tiene y su prominente peso molecular ha estado relacionada por muchos autores con la capacidad de inhibición de algunos extractos de ADN. El propio daño molecular en el ADN puede provocar la producción de reacciones de Maillard.

Productos de degradación del ADN: La degradación del ADN, produce ciertos productos que son capaces de provocar la inhibición directa de su misma amplificación (No se debe confundir el poder inhibitorio de un extracto de ADN con la incapacidad de su amplificación). Algunos autores han estudiado el efecto de este daño postmortem en los resultados de secuencias de ADN y se han podido localizar de forma directa transiciones denominadas de “tipo 1” (A→G o T→C) y de “tipo 2” (C→T o G→A) que pueden encubrir como mutaciones lo que no es más que un deterioro molecular, este tipo de transiciones suelen situarse en los extremos. La congelación de muestras puede ser un factor importante y decisivo a la hora de conservar un ADN durante un periodo de tiempo largo, ya que la reducción de 20° C en la temperatura puede conllevar un descenso de entre 10-25 veces en la tasa de reacciones químicas. En el caso de muestras de gran antigüedad, datado entre 40000-50000 años, las bajas temperaturas parece que son decisivas a la hora de ralentizar estos procesos que son responsables de la degradación de los ácidos nucleicos durante el periodo postmortem. El daño postmortem más comúnmente asociado con el ADN es el de citosina a uracilo (timina, si hablamos de la 5-metilcitosina). También hay que tener en cuenta que nuestro ADN molde pueda estar contaminado de otro ADN. Dicha desaminación da lugar a residuos modificados que son leídos como residuos de desoxitimidina por la Taq polimerasa y, por tanto, dan lugar a incorporaciones erróneas de G→A.

Para descubrir si estas desaminaciones son las responsables de los errores observados, tratamos nuestro ADN molde con una enzima procedente de E.Coli denominada ADN-Uracil-Glicosilasa (Smith, C. I., *et al* 2003), la enzima elimina el uracilo del ADN y el hueco resultante se hidroliza posteriormente por β-eliminación con la finalidad de que la cadena termine separándose. Con este método se quiere conseguir que el ADN molde reduzca el número de sustituciones G/C → A/T si éstas se deben a la desaminación de la citosina, Hofreiter M *et al.*, (2001).

4.2. Metodologías de secuenciación del ADN antiguo

Los primeros estudios pioneros de ADN se basaron en el uso de la técnica de

secuenciación Sanger. Este método fue el mismo que se empleó en la consecución del Proyecto Genoma Humano. Sin embargo, el desarrollo tecnológico y bioinformático ha propiciado el surgimiento y la expansión de las técnicas de secuenciación masiva de ácidos nucleicos conocidas como “secuenciación de próxima generación” (NGS en inglés. “Next Generation Sequencing”). Este conjunto de técnicas ha desplazado por completo a la metodología Sanger en los proyectos de genómica que se llevan a cabo en la actualidad, y también son empleadas mayoritariamente en los estudios de ADN, Martín, J *et al.*, (2020).

Los métodos NGS son técnicas que secuencian de forma más rápida que la técnica de Sanger. Su desarrollo ha marcado el comienzo de una nueva etapa en el campo de la investigación, dado que proporcionan datos de mayor rendimiento con menor coste y permite también que se puedan realizar investigaciones genómicas a gran escala poblacional. Esta tecnología tiene tres mejoras en comparación del Sanger, Van Dijk *et al.*, (2014), las cuales son: no requieren un procedimiento de clonación bacteriana, ya que preparan genotecas para la secuenciación en un sistema libre de células, procesan millones de reacciones de secuenciación en paralelo y al mismo tiempo, y por último la detección de bases se realiza cíclicamente y en paralelo. Todo esto permitió a los científicos poder procesar la secuenciación de genomas enteros por un coste muchísimo menor comparado con el método Sanger y, sobre todo, en un periodo de tiempo mucho más corto, Van Dijk *et al.*, (2014). Sin embargo, los métodos NGS necesitan del uso de algoritmos para analizar el volumen de datos que se obtienen, y así alinear y mapear genomas a partir de lecturas cortas de menos de 200 pb. Hay dos conceptos que son fundamentales para entender el proceso y los resultados de las pruebas basadas en tecnologías NGS: cobertura y profundidad. La cobertura (*coverage* o *breadth of coverage*, en inglés) se refiere al porcentaje de bases del genoma de referencia que están siendo secuenciadas una cantidad determinada de veces, Sims D *et al.*, (2014). Por otro lado, la profundidad (*depth* o *depth of coverage*) representa el número promedio de veces que cada base en el genoma es secuenciada en los fragmentos de ADN, Rubio S *et al.*, (2020).

Dentro del campo del ADN, el método NGS preferido es Illumina, conocido como secuenciación por síntesis (SBS), donde los reactivos de secuenciación incluyendo los dNTPs marcados con fluorescencia, se añaden a la cadena de ADN. Cuando se obtienen las imágenes de la celda de flujo, cada uno de los dNTPs emite una

intensidad de longitud de onda única que la hará identificable [Fig.3].

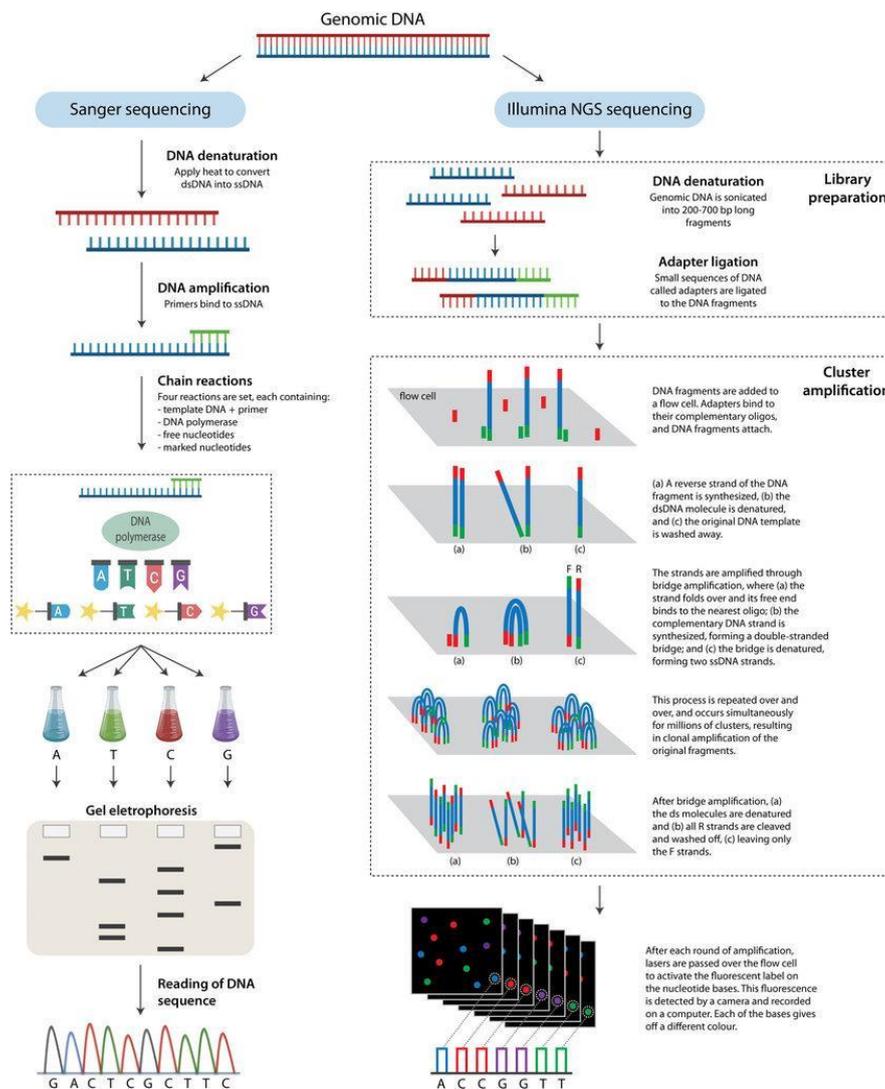


Figura 3. Comparación del método Sanger (izquierda) con el método Illumina (derecha). (Imagen recuperada de Young, A. D., 2020)

Las metodologías NGS aplicadas al ADN están adaptadas para poder identificar y evitar posibles contaminaciones en la muestra. Para ello, lo que primero se realiza es una segmentación del ADN, modificando los extremos de éste a través de la unión de secuencias adaptadoras.

Cabe mencionar que, si se está realizando reacciones múltiples de extremos emparejados, se tendrían que utilizar de 1 a 12 cebadores modificados, Son *et al.*, (2011).

En la secuenciación de los extremos, que es el método que nos interesa realizar para comprobar si existe algunas transiciones mencionadas, utilizamos como hemos

mencionado, un par de oligos o cebadores para marcar tanto el extremo 3' y 5' de cada fragmento de ADN de las genotecas monocatenarias que haremos que hibriden con oligonucleótidos complementarios a nuestra muestra, esto es parecido a multiplexación de PCR en tiempo real. Por tanto, utilizamos 12 cebadores para reconocer las 12 secuencias diferentes de las bibliotecas de ADN preparadas individualmente, lo utilizaremos para separar todas las lecturas generadas desde el secuenciador, Son *et al.*, (2011).

Por último, para la realización de los estudios que vamos a mencionar más adelante se utilizaba el método Sanger (método clásico), pero actualmente se usa el método NGS (Next Generation Sequencing) en español conocido como Secuenciación de Nueva Generación, concretamente hablaremos del NGS Illumina, que es el que más se utiliza para secuenciar ADN antiguo.

4.3. Aplicaciones de la ADNa

4.3.1. Aplicación del ADN antiguo en el estudio del pasado humano remoto: Paleogenética.

Uno de los hitos más importantes de la Paleogenética ha sido el desciframiento del genoma Neanderthal. Esta especie era el pariente evolutivo más cercano de los humanos actuales (también llamados “modernos” en la jerga científica), vivió en gran parte de Europa y Asia occidental antes de desaparecer hace 30.000 años. Estos son por así decir el “grupo hermano” de todos los seres humanos actuales. El desarrollo cultural de este grupo se dio durante el paleolítico medio, su distribución geográfica fue muy amplia y se adaptaron a multitud de hábitats, desde las costas de Portugal hasta el Oriente Próximo, Siberia occidental; y desde las llanuras noreuropeas hasta las penínsulas del sur de Europa, en su extremo más meridional, Hofreiter, M. (2011).

Una serie de rasgos morfológicos distinguían a esta especie en particular. Así poseían una complexión ancha y fornida, grandes crestas óseas, hombros anchos, clavículas largas y extremidades cortas, Bastir, M., *et al* (2011). La primera vez que se trabajó con secuencias de ADN de este homínido fue en 1997, que se obtuvo a partir de ADNmt y solo se consiguió obtener unas 370 pb, Hofreiter, M (2011). El ADN mitocondrial de los neandertales es diferente del de los humanos actuales y muy parecido entre sí, incluso entre individuos separados por miles de kilómetros. Todo

esto indicaba que los neandertales no contribuyeron al ADN mitocondrial de los humanos modernos y que eran una especie con muy poca diversidad, y, por tanto, con unos efectivos demográficos muy bajos, Lalueza-Fox., (2013). Por otro lado, otros estudios afirmaban que entre los neandertales y los ancestros de los humanos contemporáneos existe un ancestro común de hace unos 500.000 años, con esto podríamos afirmar que los neandertales pueden estar relacionados con algunos humanos modernos, ya que, por lo tanto, incluso si no se produjo ningún flujo de genes, en muchos segmentos del genoma, se espera que estos estén más estrechamente relacionados con algunos humanos actuales que entre sí, Pääbo, S., (1999). Esto es curioso, ya que durante miles de años los humanos han estado separados de los neandertales, por lo que debería ser ventajoso a la hora de identificar genes afectados por la selección positiva, Hofreiter, M. (2011).

En 2010 se pudo secuenciar el genoma Neanderthal en una especie de homínido euroasiático, estos investigadores se encontraron con los problemas que hemos mencionado anteriormente cuando se trabaja con ADN [Fig.4].

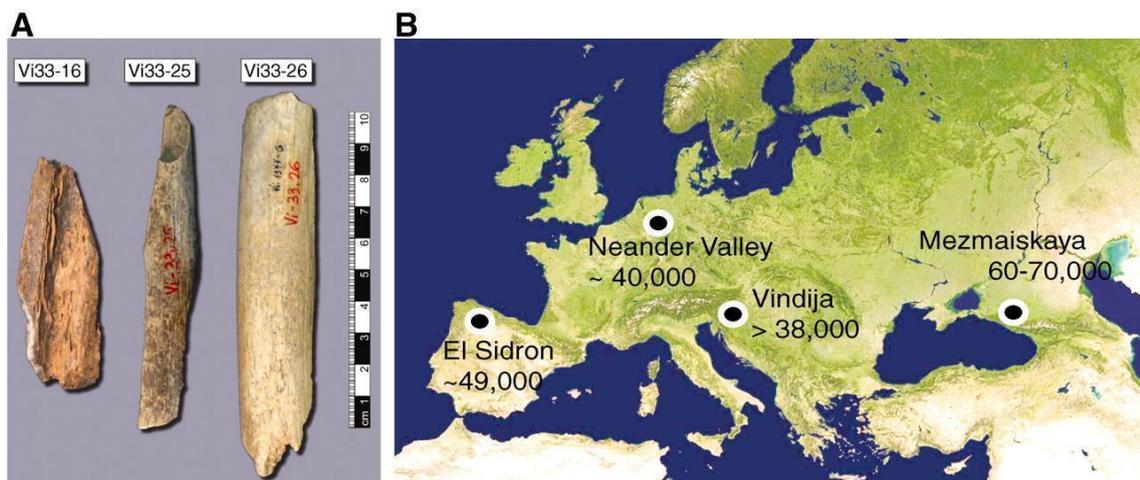


Figura 4. Imagen A: muestras de donde se recuperó el ADN (de un total de 21 huesos). Imagen B: Localización de las muestras utilizadas y la señalada es el ADN que secuenciaron (Imagen recuperada de Hofreiter, M., 2011)

Gracias a las secuencias de los ADNmt se averiguó que Vi33.16 y Vi33.25 (dos de las tres muestras que se estudiaron en esta investigación) difieren en 10 posiciones, por tanto, dedujeron que estos dos huesos provenían de diferentes individuos neandertales.

En el proyecto genoma Neanderthal se empleó el método NGS 454 (Roche). Para enriquecer las muestras y distinguir contaminaciones con ADN humano y otras especies (bacterias), deterioro químico o de ADNmt, crearon unas genotecas las cuales se amplificaron con los cebadores utilizados en la PCR en emulsión (un tipo de PCR). Una vez hecho esto, realizaron un mapeo (para averiguar la cantidad de ADN Neanderthal), mediante la plataforma 454 Life Sciences GS FLX, utilizando información de secuencias de nucleótidos en GenBank, cogiendo genomas de humanos, chimpancés, Rhesus y ratones.

Al comparar los resultados con diferentes poblaciones encontraron similitudes con humanos euroasiáticos, pero no con poblaciones africanas. Las estimaciones sugieren que el 1 – 4 % del genoma de euroasiáticos proviene de un origen Neanderthal. Todo esto apoyaría teoría del flujo genético, Fu, Q *et al.*, (2015), que incluye tres escenarios evolutivos principales, el primero, (i) el flujo de genes se produce entre poblaciones que se adaptan a las condiciones locales; (ii) el flujo de genes se restablece entre poblaciones diferenciadas tras un contacto secundario. También puede aplicarse a (iii) una única población en la que se mantiene un polimorfismo adaptativo estable a pesar del libre entrecruzamiento. La teoría propone un papel multifacético para el flujo de genes, Tigano, *et al.*, (2016). Por tanto, con esto último que hemos mencionado explicaría que se produjo una hibridación entre ambas especies.

Analizando un genoma humano y comparando los haplotipos, se pudo detectar haplotipos neandertales no presentes en el genoma de africanos y estimar cuándo ocurrió el intercambio génico entre ambas especies. Si este se cruza con un Sapiens, la descendencia tendrá la mitad del genoma Neanderthal, y si este híbrido de primera generación se cruza posteriormente sólo con humanos modernos, entonces el paso de varias generaciones (y las consiguientes recombinaciones) irán quedando segmentos de genoma Neanderthal cada vez más reducidos en el interior del genoma de humanos modernos, Fu, Q *et al.*, (2015).

Sabiendo el número de recombinaciones producidas por generación, se pudo estimar también que el primer intercambio se dio hace aproximadamente 50.000 años, justo cuando se estima que hubo una primera oleada de salida de humano moderno desde África hacia Eurasia, esto tiene sentido ya que ese intercambio concuerda con la salida

de estos homínidos de África. Estudios posteriores han analizado el ADN de un *Homo sapiens* que vivió hace 42.000 años en la región de Europa del Este, y se observó que hasta el 6 – 9% de su genoma era Neanderthal. Incluso se encontraron haplotipos de origen Neanderthal de más de 50 Mb de longitud en su interior, y se estimó que pudo tener un abuelo Neanderthal. Estos datos genéticos confirman la existencia de flujo genético entre ambas especies de homínidos, Fu, Q *et al.*, (2015).

La posibilidad de comparar el genoma Neanderthal con el genoma del *Homo sapiens* permite buscar si existe en este último algunas características únicas que ayuden a entender mejor nuestra evolución.

Para conseguir esto se tiene que realizar alineamientos del genoma completo, y de ahí comprobar en qué lugares del genoma humano de referencia no existe coincidencia con el chimpancé, orangután y el macaco Rhesus. De aquí obtenemos que aproximadamente 10.535.445 sustituciones y 479.863 indels (inserciones y deleciones) se han dado en el linaje humano. Sin embargo, la gran mayoría de esas sustituciones e indels (87,9% y 87,3%, respectivamente) ocurrieron antes de la divergencia Neanderthal de los humanos modernos, Green, R. *et al.*, (2010). Se comparó el genoma Neanderthal con el genoma de 5 humanos actuales, estas muestras pertenecían a personas de diferentes partes del mundo, trabajaron con el ADNmt y empezaron a realizar análisis para conocer la divergencia entre las muestras que poseían. Primeramente, secuenciaron los genomas de los humanos modernos, (eran, un individuo de África Meridional, un yoruba de África occidental, un papua Nueva Guinea, un Oriental y un Europeo Occidental). Tras secuenciarlas se realizaron alineaciones por 3 vías entre genoma de Neanderthal, humanos modernos y chimpancés, utilizando el genoma inferido del ancestro común de humanos y chimpancés y se analizaron, donde obtuvieron una divergencia con los neandertales entre un 8,2% y un 10,3%, Green, R. *et al.*, (2010).

Las características que compartimos todos los seres humanos actuales respecto a un Neanderthal son: 78 sustituciones de nucleótidos en genes codificantes para proteína, estas sustituciones en los humanos modernos están fijas, y en los neandertales procede de un ancestral (similar al chimpancé). Encontramos solo cinco genes con más de una sustitución fija que cambia la estructura primaria de las proteínas codificadas, que son: SPAG17, PCD16, TTF1, CAN15 y RPTN. Es decir, en el

genoma humano actual existen cambios de aminoácidos que están fijados, pero no en los neandertales. Una de las sustituciones en RPTN crea un codón de stop que hace que la proteína humana contenga 784 en lugar de 892 aminoácidos. Como hemos mencionado anteriormente, también se identifican sustituciones regulatorias que son fijas en los humanos actuales, pero no en los neandertales. Específicamente, encontramos 42 sustituciones y tres indels en las regiones 5' UTR de los genes, y 190 sustituciones y 33 indels en las regiones 3' UTR de los genes que se han vuelto fijas en los humanos desde que divergieron de los neandertales. Son de especial interés los microARN (miARN). Encontramos un miARN donde los humanos llevan una sustitución fija en una posición que era ancestral en los neandertales (hsa-mir-1304) y un caso de inserción de un solo nucleótido fijo donde el Neanderthal es ancestral (AC109351.3), Green, R. *et al.*, (2010). La gran pregunta que sigue sin poder responderse aún es si estos cambios génicos, exclusivos de la especie humana, han tenido importancia directa en nuestra evolución como especie.

También era de interés, aparte de comparar los genes codificantes, identificar regiones del genoma humano de evolución aceleradas (Human Accelerated Regions). Se definen como regiones del genoma que se conservan a lo largo de la evolución de los vertebrados, pero que cambiaron radicalmente desde que los humanos y los chimpancés se separaron de su ancestro común. Descubrieron que en una parte del genoma Neanderthal había unos 3259 cambios específicos pertenecientes a los humanos actuales. Los neandertales llevan un derivado del 91,4% de estos, significativamente más que para otras sustituciones e indels específicos de humanos (87,9%). Con estos datos sacaron la conclusión de que estos cambios en estas regiones se dieron antes de que se produjera la división entre los neandertales y los humanos actuales. Por otra parte, también identificaron 51 posiciones en 45 HAR donde los neandertales llevan la versión ancestral, mientras que todos los humanos actuales conocidos llevan la versión derivada, es decir, la más reciente. Todo esto indicó que en el genoma se produjeron modificaciones recientes que pueden ser particularmente interesantes en el estudio de estas en un futuro, Green, R. *et al.*, (2010).

Además del Neanderthal, se ha descrito recientemente un nuevo grupo homínido que compartió tiempo y espacio con éstos y también con los humanos modernos, denominados Denisovanos. Su descubrimiento se basa exclusivamente en el estudio

de su ADN, ya que por el momento no se han descubierto cráneos ni el hueso de la cadera, claves para una identificación morfológica clásica. A partir de muestras de una falange manual distal y dos molares, todos ellos excavados en la cueva de Denisova, en los montes Altai (sur de Siberia), se obtuvo un genoma que no correspondía ni a Neanderthal ni a humanos modernos, pero con mucho parecido a ambos. Las dos secuencias del genoma arcaico publicadas son de baja cobertura: 1,9 veces la cobertura genómica de la falange denisovana, esta muestra contenía cerca de un 70% de ADN endógeno. Como sabemos, cuando se trabaja con ADN muchas posiciones de los genomas están afectadas por errores de secuenciación o desincorporaciones de nucleótidos causadas por daños en este. Los intentos anteriores de generar una secuencia genómica de alta cobertura a partir de un humano arcaico se han visto obstaculizados por los altos niveles de contaminación ambiental, Meyer, M *et al.*, (2012).

La cobertura del genoma es bastante uniforme con el 99,93% de las posiciones "cartografiadas" cubiertas por al menos una, el 99,43% por al menos diez y el 92,93% por al menos 20 secuencias de ADN independientes. Se pudieron determinar genotipos de alta calidad (calidad del genotipo ≥ 40) para el 97,64% de las posiciones. Si bien la cobertura en bibliotecas preparadas a partir de muestras antiguas con métodos anteriores está sesgada hacia secuencias ricas en GC, la cobertura de las bibliotecas preparadas con el método monocatenario del individuo denisovano es similar a los once genomas humanos actuales (preparados de ADNdc) en que la cobertura se correlaciona positivamente con el contenido de AT, Meyer, M *et al.*, (2012).

Para sacar la secuencia un estudio realizó lo siguiente, primeramente, se secuenció estas genotecas [Fig.5] desde ambos extremos utilizando el Genome Analyzer Ix de Illumina que incluyeron lecturas para dos índices, para excluir la posibilidad de contaminación descendente con bibliotecas de ADN modernas, Green, R. E., *et al.*, (2010). Las secuencias de más de 35 pb se alinearon con el genoma humano de referencia (GRCh37/1000 Genome project release) y el genoma del chimpancé (CGSC 2.1/panTro2) con el Burrows-Wheeler Aligner. Después de eliminar los duplicados de la reacción en cadena de la polimerasa, los genotipos fueron llamados con el Genome Analysis Toolkit, Gross, W. L *et al.*, (2010); McKenna, A *et al.*, (2010).

Las tres genotecas denisovanas produjeron 82,2 gigabases de secuencia no duplicada alineada con el genoma humano, Gross, W. L *et al.*, (2010). Junto con los datos anteriores, Reich, D *et al.*, (2010), esto proporciona una cobertura de aproximadamente 31 veces de los ~1,86 gigabases del genoma autosómico humano a los que se pueden asignar con seguridad las secuencias cortas, Gross, W. L *et al.*, (2010).

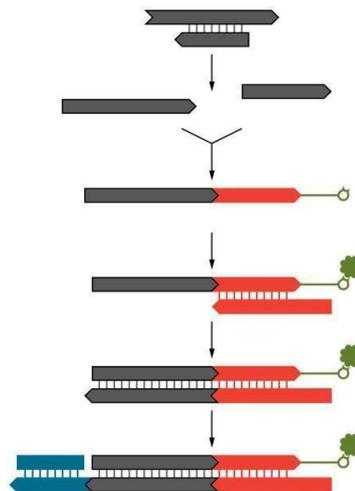


Figura 5. Para la preparación de las genotecas, se desfosforilan las moléculas de ADN antiguas y seguidamente, mediante calor la desnaturalizamos. Gracias a un adaptador previamente biotinilado podremos ligar el oligonucleótido por el extremo 3' de las moléculas, estas se inmovilizan en perlas recubiertas de estreptavidina, se empieza a crear copias cuando el cebador se une. Seguidamente se une, mediante ligación, el adaptador a la hebra recién sintetizada. Para terminar, las perlas se destruyen mediante calor para que éstas liberen aquellas moléculas que forman nuestra genoteca genómica.

Tras la realización de todo lo mencionado anteriormente, sacaron como conclusión que los denisovanos poseían los siguientes aspectos característicos en su genoma: Los simios poseen 24 pares de cromosomas, mientras que nosotros, los humanos, poseemos 23. Esto ocurrió gracias a una fusión de dos cromosomas acrocéntricos que formaron el cromosoma humano metacéntrico 2. Una diferencia en el cariotipo probablemente habría reducido la fertilidad de cualquier descendencia de denisovanos y humanos modernos, Meyer, M *et al.*, (2012).

Otro ejemplo de genomas humanos extintos son los denisovanos, en los que casi no hay datos fenotípicos, lo único que se sabe es que portan alelos asociados al color de ojos castaños, piel oscura y cabello castaño. Por otro lado, también conlleva cambios en determinadas posiciones de nucleótidos que los humanos modernos no portamos,

aunque, por otro lado, toda esta información se obtiene de un solo individuo, y se espera que toda la población también pueda compartir estas características, Meyer, M., *et al.*, (2012).

Para finalizar, cabe mencionar que gracias a la alta calidad del genoma denisovano se puede medir su heterocigosidad, con esto nos referimos al número de sitios diferentes que nos encontramos en los genomas tanto paternos como maternos. El valor de este rondaba alrededor del 0,022%, dentro de esta observamos que aproximadamente el 20% es procedente de africanos, entre el 26 - 33% de los euroasiáticos, y el 36 % de una población sudamericana llamada Karitiana. Al no encontrarse a lo largo del genoma denisovano tramos largos homocigóticos, y encontrásemos la existencia de la heterocigosidad de diferentes poblaciones, se descarta que se produjera una endogamia entre los antepasados de los denisovanos. Se llega a la conclusión de que la existencia de diversidad genética de la población a la que pertenecía el individuo denisovano era muy baja en comparación con los humanos actuales, Meyer, M *et al.*, (2012).

Para estimar cómo el tamaño del genoma de la población de los humanos actuales y de Denisovan ha cambiado con el tiempo, se utilizó un modelo conocido (y se aplicó a todos los genomas analizados), denominado: coalescente de Markov, Li, H., *et al* (2011). Cuyo resultado mostró que los genomas actuales comparten cambios similares en el tamaño del genoma de la población, en particular, se observó que este aumentó el doble que antes, es decir, es el doble que hace 125,000-250,000 años, Conrad, D. F *et al* (2011). Los denisovanos, por el contrario, muestran una drástica disminución de tamaño en el momento en que la población humana moderna comenzó a expandirse, Meyer, M *et al* (2012).

El hecho de que el tamaño del genoma de los denisovanos disminuyera pudo deberse a que la selección natural fue menos eficaz para eliminar mutaciones levemente perjudiciales. Esto quiere decir que la cantidad de sustituciones no sinónimas con respecto a las sustituciones sinónimas (aquellas que no producen un cambio en la secuencia de aminoácidos, al contrario de las no sinónimas) en los genomas analizados encontraron que esta relación era, en promedio, 1,5-2,5 veces mayor en denisovanos que en los humanos actuales, dependiendo de la clase de sitios y

poblaciones con las que se comparan los denisovanos, Meyer, M *et al.*, (2012).

4.3.2. *Aplicación del ADN antiguo en el estudio del pasado humano reciente: Arqueogenética*

Esta nueva perspectiva del conocimiento utiliza el ADN_a como material de investigación. La importancia de este tipo de análisis radica en dar respuesta a cuestiones arqueológicas, como establecer parentesco entre individuos, determinar el sexo de los restos o identificar la posibilidad de antepasados o descendientes entre los individuos analizados. Además, a un nivel superior de análisis, el componente de tiempo se puede incluir en el estudio de la variabilidad genética, de modo que se pueda realizar el análisis filogenético de poblaciones pasadas y especies extintas. De esta forma, se puede abrir una puerta para responder preguntas sobre: el proceso migratorio, la relación genética entre patrones residenciales y poblaciones antiguas, continuidad o reemplazo poblacional, reconstruyendo así la historia poblacional en el tiempo y el espacio, Destro-Bisol *et al.*, (2010); Crespo *et al.*, (2012). Sin embargo, es importante señalar que todos estos métodos son más robustos cuando se combinan con datos arqueológicos, etnográficos, históricos, etc., permitiendo que se coloquen en un marco teórico más completo, Cardozo, D. G. (2017).

Se puede emplear el ADN_a para el seguimiento de migraciones históricas o prehistóricas, conocida como “arqueogenética” en la que se aplican estudios a restos humanos de finales del Paleolítico, Neolítico y época antigua (10.000-15.000 años hasta la actualidad). Estos estudios nos están permitiendo conocer aspectos nuevos o diferentes sobre las diversas culturas antiguas, las cuales chocan en muchos casos con las teorías que habían sido propuestas desde el punto de vista de la arqueología. Un caso a destacar es la civilización de los Yamnaya y su gran expansión por Europa, por lo que se conoce como uno de los más espectaculares. El origen de Yamnaya parece ser encontrado en la cultura de Khvalnisk, que se desarrolló a orillas del río Volga, y en la cultura de Sredny, ubicada en el río Dnieper. Ambas culturas se formaron durante el V milenio a. C. y parte del IV. Mallory y Adams indican la movilidad y el intercambio comercial en la estepa póntica como causa de la génesis de la cultura de Yamnaya, gracias al uso de vehículos domésticos (como el caballo) y vehículos de ruedas tirados por bueyes, Myres, N *et al.*, (2011). Gimbutas ha identificado la cultura

de Yamna dentro de la "Cultura de Kurgans" más amplia, definiendo este último como una comunidad pastoral del V milenio a. C. hasta el III Milenio a. C. En su libro, la prehistoria de Europa del Este, parte I, la autenticidad enfatiza la diferencia física de los individuos de Yamnaya con los anteriores en el área ocupada, sosteniendo un esqueleto más elegante, con un índice robusto más bajo y una cara más estrecha. Genéticamente, un estudio realizado por Haak *et al.*, en 2015, reveló que los individuos que pertenecían a la cultura de Yamnaya eran una mezcla de tres poblaciones ancestrales. El elemento dominante pertenecía al grupo mesolítico oriental de cazadores-recolectores (Haplogrupos R1A y R1B, estos haplotipos son los más comunes de Europa, que proceden del cromosoma Y, Myres, N *et al.*, (2011)), seguidos por los cazadores recolectores que provienen del Cáucaso (Haplogrupo J) y los cazadores occidentales, Haak *et al.*, (2015). Posteriormente, el genoma ancestral Yamnaya se reconoce en R1B Haplogrupo (se define como un grupo grande de haplotipos), especialmente en el subclado R1B-M269, Bretos *et al.*, (2019).

El registro arqueológico da evidencia de importantes cambios culturales en Europa y Asia después del período Neolítico. Hacia los 3000 a.C., las culturas agrícolas neolíticas en Europa del Este, parecen ser reemplazadas principalmente por la cultura Yamnaya de la Edad de Bronce temprano, que se asocia con una percepción completamente nueva de la familia, la propiedad y la personalidad, que se extiende rápidamente desde Hungría hasta los Urales, Allentoft *et al.*, (2015). Hacia 2800 a. C. la nueva capacitación social y económica se desarrolló en la Europa templada, llamada de varias formas, culturas de cerámica de cruceros, una tumba única o una tapa de combate, posiblemente derivada del fondo de Yamnaya y se reemplaza culturalmente a los agricultores neolíticos restantes. En el oeste y centro de Asia, los coleccionistas de cazadores todavía dominaban la edad del bronce temprano, excepto en las montañas de Altai y en la cuenca del río Minusinsk, donde existía la cultura Afanasievo con afinidad cultural cerca de Yamnaya, Allentoft *et al.*, (2015).

Desde principios de 2000 a.C., una nueva clase de maestros artesanales conocida como la cultura de Sintastha ha aparecido en los Urales, creciendo y entrenando caballos y produciendo nuevas armas sofisticadas, Allentoft *et al.*, (2015). Para sacar todas estas conclusiones los investigadores recopilaron sus muestras del Cromosoma Y y del ADN mitocondrial, y cuyos resultados le llevaron a la conclusión de que todo lo anterior se produjo gracias a que se ejerció una gran influencia en distribución de alelos en las poblaciones mencionadas anteriormente, ya que obtuvieron un F_{st}

(conocido como índice de fijación, en inglés **proportion of the total genetic variance contained**, donde la S es de subpoblación y la T es la relación que existe entre la varianza genética total, que sirve para comprobar cómo de diferentes o iguales son dos inter-poblaciones siendo: 0 (las poblaciones son iguales) y 1 (las poblaciones son diferentes), Allentoft *et al.*, (2015).

Más baja en los individuos de la Edad de Bronce que en los individuos del Mesolítico y Neolítico. Un ejemplo, los cazadores- recolectores del mesolítico tenía un $F_{st} = 0,08$, tiene casi el mismo valor que los asiáticos orientales y europeos contemporáneos, Allentoft *et al.*, (2015), con todo esto podemos llegar a la conclusión de que los alelos después de la edad de Bronce gracias al crecimiento de la población y combinado con el flujo continuo de genes de poblaciones explicaría que los componentes genómicos ancestrales divergentes se difundieran aún más. En este artículo también se trabajó con SNPs, los cuales están muy estudiados y muy trabajados, concretamente con rs1426654 (pigmentación de la piel), rs4988235 (asociado con la intolerancia de la lactosa), rs16891982 (pigmentación de la piel, cabello, ojos) y rs12913832 (determinante de los ojos azules sobre marrones). Realizaron una comparación de estos SNPs entre las poblaciones antiguas y las poblaciones modernas obteniendo también las frecuencias de estos en cada uno de los grupos. Si nos fijamos en resultados sobre el SNP de la lactosa, vemos que su frecuencia en la población europea antigua es muy baja, con un valor de 5% de la derivación de alelos. Sin embargo, en la población europea de la moderna esta tolerancia es más alta. Puede deberse a la evolución convergente, esta suele producirse por cambios genéticos similares en especies independientes mediante un proceso conocido como evolución paralela. En algunas ocasiones, la evolución convergente es el resultado de la evolución de un cambio genético heredado de una población ancestral o de la hibridación entre especies. Por lo tanto, la evolución convergente a nivel genético puede ocurrir a través de estos tres procesos: primero, la evolución a través de mutaciones que ocurren de manera independiente en diferentes poblaciones o especies; segundo, la evolución de alelos polimórficos en poblaciones ancestrales compartidas; tercero, la evolución de alelos introducidos a partir de una población. a otro por hibridación se llama introgresión. Vale la pena distinguir estos escenarios porque cada uno de ellos proporciona evidencia de diferentes trayectorias evolutivas, Stern, D. L. (2013).

A continuación, vamos hablar de un reciente estudio en el cual, esta investigación

arqueogenética es de gran calado hasta la fecha. Se analizó el genoma de 271 humanos que vivieron entre el 4900 y el 1600 a.C. en el corazón de Europa, este estudio reveló cambios genéticos y procesos sociales sin precedente. El inicio de esta investigación se llevó a cabo por el descubrimiento de estos individuos en Bohemia, Papac, L *et al* (2021). Grandes migraciones se llevaron a cabo con el comienzo de la ascendencia "esteparia", y aproximadamente en el año ~2800 a.C. coexistieron tres grupos genéticos y culturalmente diferentes: Los Corded Ware (CW) aparecieron hacia el 2900 a.C., fueron en un principio genéticamente diversos y asimilaron mujeres de diversos orígenes y con ascendencia similar a los individuos Yamnaya, los Yamnaya ya mencionados anteriormente y los Bell Beaker. Tanto los grupos de Corded Ware como los de Bell Beaker (BB) sufrieron cambios dinámicos, que implican reducciones bruscas y sustituciones completas de la diversidad cromosómica Y en el ~2600 y ~2400 a.C., respectivamente, esta última acompañada de un aumento de la ascendencia neolítica. Este estudio arqueogenético se centra sobre todo en un enclave de tránsito de culturas entre el Neolítico tardío y la Edad de Bronce temprana.

Este estudio muestra que la CW temprana tiene una diversidad genética variada, algunas de las cuales son similares a Yamnaya, y algunas también están fuera de la diversidad genética neolítica previamente muestreada en Europa Central. Además, la utilización de muestras de pequeño tamaño para representar fenómenos arqueológicos suprarregionales, así como las interpretaciones histórico-culturales resultantes, han despertado críticas por parte de los arqueólogos. Existen unas series de cuestiones no resueltas, estas se refieren a los orígenes genéticos como geográficos de los individuos CW y BB su relación entre sí y con los individuos Yamnaya, así como el origen de los individuos Úntice de la Edad del Bronce Temprana (EBA). Para poder entender mejor estas transiciones, estos investigadores analizaron un transecto temporal arqueogenético de alta resolución de 271 individuos de la parte norte de Bohemia. A través de un denso muestreo genético de culturas arqueológicas que se solapan geográfica y temporalmente, los investigadores pretendieron: (i) abordar los cambios culturales en el Eneolítico y el EBA de Europa central, (ii) caracterizar la diversidad genética centroeuropea inmediatamente antes de la aparición del CW, (iii) datar cuándo aparecieron por primera vez en Europa central individuos con ascendencia esteparia de tipo Yamnaya y comprender su origen genético y su estructura social, (iv) averiguar el origen y el alcance del intercambio

biológico entre los "locales" y los "migrantes" tras la aparición del CW, y (v) identificar las transformaciones sociales vinculadas a los cambios genéticos y arqueológicos. Los perfiles genéticos altamente diversos (tanto nucleares como cromosómicos Y) de los primeros CW sugieren una organización social diferente a los últimos CW y BB, cuyo patrón cromosómico Y es indicativo de una patrilinealidad estricta. Esto quiere decir que estos grupos, además de ser de diferentes culturas, utilizan material y prácticas mortuorias, incluso tenían un patrón diferente de apareamiento y organización social. Esta teoría se ve apoyada por el hallazgo de una variación del cromosoma Y que no se solapa en absoluto entre los últimos CW y BB, parcialmente contemporáneos, lo que indica un gran grado de aislamiento de apareamiento paterno entre estos dos grupos, incluso cuando se encuentran en el mismo sitio (por ejemplo, el inicio del Úntice preclásico estuvo acompañado de una contribución de ADN nuclear del 40% y del 80% del cromosoma Y, que en última instancia se originó en el noreste y rompió las prácticas mortuorias diferenciadas por género y la estricta patrilinealidad del CW tardío y el BB.

Por lo tanto, los resultados que obtuvieron sugieren dos períodos principales (el primero de CW y el primero de Úntice) de influencia genética del noreste, aunque esta gran parte permanece sin muestrear en el registro arqueogenético europeo (por ejemplo, Bielorrusia). Los resultados revelan una historia compleja y altamente dinámica de la Europa central del Neolítico al EBA, durante la cual la migración y el movimiento de personas facilitaron cambios genéticos y sociales. La sociedad temprana de CW era diversa, apareciendo en fuertes transformaciones culturales y genéticas, involucrando a hombres y mujeres de diferentes orígenes y posibles razas. A pesar de la continuidad de la cultura material, se han producido cambios genéticos en las sociedades CW, BB y EBA. La afiliación cultural jugó un papel importante en el comportamiento social en el tercer milenio a.C. con el tiempo, con la afluencia de nuevas personas, el comportamiento social finalmente cambió. Si bien se puede observar la influencia de los procesos sociales en los patrones de diversidad genética, se necesita más investigación interdisciplinaria para caracterizar los factores impulsores de estos cambios, incluidos los niveles micro y macro regionales.

Para terminar este apartado vamos hablar de un estudio genético reciente titulado "*The genomic history of the Iberian Peninsula over the past 8000 years*", trabajan con

una muestra de 271 individuos del territorio peninsular: los cuales eran 4 individuos del mesolítico, 44 neolítico, 47 de la Edad del Cobre, 53 de la Edad del Bronce, 24 de la Edad del Hierro y 99 del periodo histórico de la península ibérica, Olalde *et al.*, (2019). Este estudio documenta que existió contacto entre la Península Ibérica y el norte de África comprendido entre los años 2500 a.C – 2000 a.C, ya que observan que se produce un reemplazo del casi el 40% de la descendencia de los íberos y observan también un reemplazo de casi el 100% en sus cromosomas Y. Los linajes parentales más comunes en la Edad del Cobre prácticamente desaparecieron durante la Edad del Bronce, siendo sustituidos por el linaje R1b-M269, el relacionado con los Yamnaya. Todo esto lo consiguieron, ya que realizaron una genoteca, estas fueron tratadas con la enzima UDG (Uracil-ADN-glicosilasa) para eliminar cualquier contaminante, estas muestras estaban enriquecidas de millones de SNPs, y a partir de ahí realizaron sus análisis consiguiendo el resultado mencionado anteriormente, Olalde *et al.*, (2019).

La genética nos permite conocer cómo de parecidos o diferentes eran los individuos de las poblaciones y cómo se desplazaron por el planeta, ya que se pensaba que muchas de estas civilizaciones permanecieron sedentarias en un mismo lugar, lo que se demostró que en realidad fueron reemplazadas por otras, las cuales asumieron gran parte de sus ritos y culturas. Al tratarse de un campo bastante reciente, hay muchos conflictos entre los arqueólogos y los genetistas, aunque también hay grupos nuevos de investigación multidisciplinarios que están dando grandes avances, además cabe destacar que hay diversos estudios que se han enfocado en la Península ibérica, Turbón, D *et al.*, (2003).

5. CONCLUSIONES

Tras la realización de este estudio, se han obtenido las siguientes conclusiones:

-Primera: El genoma Neanderthal y el genoma de humanos modernos no africanos comparten un 1-4%, probablemente producto de fenómenos de hibridación durante la expansión del hombre moderno por Oriente Medio y Europa.

-Segunda: El genoma de los denisovanos es 1,5 - 2,5 veces más pequeño que el de los humanos modernos por consecuencia de la poca eficiencia de la selección natural.

-Tercera: Aumento de la cantidad de SNPs en las poblaciones modernas en comparación con las poblaciones antiguas debido a la evolución convergente.

-Cuarta: Gracias al crecimiento de la población combinado con el flujo continuo de genes, después de la Edad del Bronce, se expandieron los componentes genómicos ancestrales.

-Quinta: Se producen al menos otros tres acontecimientos migratorios que dieron forma a la prehistoria centroeuropea.

-Sexta: La relevancia que tiene el ADN para los estudios en la Paleogenética como Arqueogenética.

6. BIBLIOGRAFÍA

ALLENTOFT, M. E., SIKORA, M., SJÖGREN, K. G., RASMUSSEN, S., RASMUSSEN, M., STENDERUP, J., ... Y WILLERSLEV, E. (2015): "Population genomics of bronze age Eurasia". *Nature*, 522(7555), 167-172. DOI: 10.1038/nature14507.

ANDREW, P (2010): "Human Molecular Genetics" (4 Edición), London: Garland Publishing.

BASTIR, M., ROSAS, A., GUNZ, P., PEÑA-MELIAN, A., MANZI, G., HARVATI, K., ... Y HUBLIN, J. J. (2011): "Evolution of the base of the brain in highly encephalized human species". *Nature Communications*, 2(1), 1-8. DOI: 10.1038/ncomms1593

BRETOS EZCURRA, M., Y PICAZO MILLÁN, J. V. (2019): "Migraciones en Prehistoria. La expansión de la cultura Yamna y el reemplazamiento poblacional en la península ibérica."

CARAMELLI, D., LALUEZA-FOX, C., VERNESI, C., LARI, M., CASOLI, A., MALLEGNI, F., ... Y BERTORELLE, G. (2003): "Evidence for a genetic

discontinuity between Neandertals and 24,000-year-old anatomically modern Europeans". *Proceedings of the National Academy of Sciences*, 100(11), 6593-6597. DOI: 10.1073/pnas.1130343100

CARDOZO, D. G. (2017): "Marcadores Genéticos y Patrones Dietarios en Poblaciones Aborígenes Pre y Posthispánicas del Paraná Inferior".

CLARK, AG, HUBISZ, MJ, BUSTAMANTE, CD, WILLIAMSON, SH Y NIELSEN, R. (2005): "Sesgo de verificación en estudios de polimorfismo en todo el genoma humano". *Investigación del genoma*, 15 (11), 1496-1502. DOI: 10.1101 / gr.4107905

CONRAD, D. F., KEEBLER, J. E., DEPRISTO, M. A., LINDSAY, S. J., ZHANG, Y., CASALS, F., ... Y 1000 GENOMES PROJECT. (2011): "Variation in genome-wide mutation rates within and between human families". *Nature genetics*, 43(7), 712.

CRESPO, C.M., J. L. LANATA, A. PÉREZ, A. HAJDUK, S. A. AVENA, F. R. CARNESE Y C. B. DEJEAN. (2012): "Primeros resultados del análisis de ADN antiguo en restos esqueléticos provenientes de colecciones de museos de Bariloche y Puerto San Julián (Patagonia Argentina)". Trabajo presentado en el XII Congreso de la Asociación Latinoamericana de Antropología Biológica, San José de Costa Rica, Costa Rica, Noviembre 2012. DOI: doi.org/10.17139/raab.19.1.21

CRICK F.H.C. AND WATSON J.D. (1954): "The complementary structure of desoxyribonucleic acid" *Proc. R. Soc. Lond. A* 1954 223, 80-96. DOI: 10.1098/rspa.1954.0101

DESTRO-BISOL G, JOBLING MA, ROCHA J, NOVEMBRE J, RICHARDS MB, MULLIGAN C, BATINI C, MANNI F. (2010): "Molecular Anthropology in

the genomic era". *Journal of Anthropological Science*, 88:93-112.

DÍEZ, S. P., Y PARDO, E. A. (2012): "Arqueogenética: una introducción a las aplicaciones de la genética sobre los restos arqueológicos". In *Nuevos paradigmas en la investigación arqueológica: actas del segundo Congreso de Arqueología de Chamartín (Ávila): Chamartín (Ávila), 5, 6 y 7 de agosto de 2011* (pp. 51-64).

FRAILE, I. R. (2007): "Utilización del ADN antiguo en estudios bioarqueológicos". *Archaeobios*, (1), 7.

FRANCISCO JAVIER, N. (2007): "Genética Humana: conceptos, mecanismos y aplicaciones de la Genética en el campo de la biomedicina" (1 Edición), Madrid: Pearson Educación, S.A.

FREDMAN, D., WHITE, SJ, POTTER, S., EICHLER, EE, DEN DUNNEN, JT Y BROOKES, AJ. (2004): "Variación de secuencia compleja relacionada con SNP en duplicaciones del genoma segmentario". *Nature genetics*, 36 (8), 861-866. DOI: 10.1038/ng1401

FU, Q., HAJDINJAK, M., MOLDOVAN, O. T., CONSTANTIN, S., MALLICK, S., SKOGLUND, P., ... Y PÄÄBO, S. (2015): "An early modern human from Romania with a recent Neanderthal ancestor". *Nature*, 524(7564), 216-219.

1000 GENOMES PROJECT CONSORTIUM. (2010): "A map of human genome variation from population scale sequencing". *Nature*, 467(7319), 1061. DOI:10.1038/nature09534

GREEN, R. E., KRAUSE, J., BRIGGS, A. W., MARICIC, T., STENZEL, U., KIRCHER, M., ... Y PÄÄBO, S. (2010): "A draft sequence of the Neandertal genome". *science*, 328(5979), 710-722. DOI: 10.1126/science.1188021

GROSS, W. L., Y JACKS, T. (2010): 5. MATERIALS AND METHODS ARE AVAILABLE AS SUPPLEMENTARY MATERIALS ON SCIENCE ONLINE. 6.

AG KNUDSON, J. *CANCER RES. CLIN. ONCOL.* 122, 135–140 (1996). 7. Y. Kim et al., “Nucleic Acids Res”. 40 (W1), W525–W530 (2012). *Arthritis Rheum*, 62, 2787-2795. DOI: 10.1126/science.1244059

HAAK, W., LAZARIDIS, I., PATTERSON, N., ROHLAND, N., MALLICK, S., LLAMAS, B., ... Y REICH, D. (2015): “Massive migration from the steppe was a source for Indo-European languages in Europe”. *Nature*, 522(7555), 207-211.

HENRÍQUEZ, O. (2001): “La Era del Genoma”. *Rev Med Hond*, 69, 2.

HOFREITER, M. (2011): “Drafting human ancestry: What does the Neanderthal genome tell us about hominid evolution? Commentary on Green et al.” (2010). *Human Biology*, 83(1), 1-11. DOI: 10.3378/027.083.0101

HOFREITER, M., JAENICKE, V., SERRE, D., HAESELER, A. V., Y PÄÄBO, S. (2001): “DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA”. *Nucleic acids research*, 29(23), 4793-4799.

LALUEZA-FOX, C. (2013): “El genoma neandertal”. *Mètode*, 38-47.

LAMOLLE, G., Y MUSTO, H. (2018): “Genoma Humano. Aspectos estructurales”. In *Anales de la Facultad de Medicina* (Vol. 5, No. 2, pp. 12-28). Universidad de la República. Facultad de Medicina. DOI: 10.25184/anfamed2018v5n2a10

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN, J., ... Y PROCTOR, M. J. (2001): “Initial sequencing and analysis of the human genome”. DOI:10.1038/35057062

LI, H., Y DURBIN, R. (2011). “Inference of human population history from individual whole-genome sequences”. *Nature*, 475(7357), 493-496.

MARÍA VICTORIA ARRUGA LAVIÑA, (2014): “Epigenética. ¿Somos únicamente la expresión de nuestro genotipo, o hay interacciones entre los genes y factores externos que modifican nuestro fenotipo?” Trabajo Fin de Grado Real Academia de Ciencias de Zaragoza.

MARTÍN, J. M. V., ORTIGOSA, F., Y PENDON, R. A. C. (2020): “Métodos de secuenciación de ácidos nucleicos: Primera generación”. Encuentros en la Biología, 13(173), 19-25.

MCKENNA, A., HANNA, M., BANKS, E., SIVACHENKO, A., CIBULSKIS, K., KERNYTSKY, A., ... Y DEPRISTO, M. A. (2010): “The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data”. Genome research, 20(9), 1297-1303. DOI: 10.1101/gr.107524.110.

MEYER, M., KIRCHER, M., GANSAUGE, M. T., LI, H., RACIMO, F., MALLICK, S., ... Y PÄÄBO, S. (2012): “A high-coverage genome sequence from an archaic Denisovan individual”. Science, 338(6104), 222-226. DOI: 10.1126/science.1224344

MYRES, N. M., ROOTSI, S., LIN, A. A., JÄRVE, M., KING, R. J., KUTUEV, I., ... Y UNDERHILL, P. A. (2011): “A major Y-chromosome haplogroup R1b Holocene era founder effect in Central and Western Europe”. European Journal of Human Genetics, 19(1), 95-101.

NATIONAL INSTITUTES OF HEALTH (US); Biological Sciences Curriculum Study. NIH Curriculum Supplement Series [Internet]. Bethesda (MD): National Institutes of Health (US); 2007. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK20364>

OLALDE, I., MALLICK, S., PATTERSON, N., ROHLAND, N., VILLALBA-MOUCO, V., SILVA, M., ... Y REICH, D. (2019): “The genomic history of the Iberian Peninsula over the past 8000 years”. Science, 363(6432), 1230-1234. DOI: 10.1126/science.aav4040

PÄÄBO, S. (1999): "Human evolution". *Trends in Genetics*, 15(12), M13-M16.
DOI: 10.1016/S0168-9525(99)01904-6

PAPAC, L., ERNÉE, M., DOBEŠ, M., LANGOVÁ, M., ROHRLACH, A. B., ARON, F., ... Y HAAK, W. (2021). "Dynamic changes in genomic and social structures in third millennium BCE central Europe". *Science Advances*, 7(35), eabi6941. DOI:10.1126/sciadv.abi6941

PENNISI, E. (2001): "The human genome". DOI:
10.1126/science.291.5507.1177

PEÑA-CASTRO, J. M., GREGORIO-RAMÍREZ, O., Y BARRERA-FIGUEROA, B. E. (2013): "Los métodos experimentales que permiten el estudio de las macromoléculas de la vida: historia, fundamentos y perspectivas". *Educación química*, 24(2), 237-246. DOI:10.1016/S0187-893X(13)72468-6

REICH, D., GREEN, R. E., KIRCHER, M., KRAUSE, J., PATTERSON, N., DURAND, E. Y., ... Y PÄÄBO, S. (2010): "Genetic history of an archaic hominin group from Denisova Cave in Siberia". *Nature*, 468(7327), 1053-1060.
DOI: doi.org/10.1038/nature09710

ROMERO, R. G., AUSIÓ, J., MÉNDEZ, J., Y LÓPEZ, J. M. E. (2011): "El papel clave de las histonas". *Investigación y Ciencia*.

ROSADO, L. A. M., Y GALARZA-BRITO, J. E. (2019): "Acerca de ciertas variaciones estructurales del genoma humano". *Revista Cubana de Genética Comunitaria*, 11(3).

RUBIO, S., PACHECO-OROZCO, R. A., MILENA GÓMEZ, A., PERDOMO, S., Y GARCÍA-ROBLES, R. (2020): "DNA Next-Generation Sequencing (NGS): Present and Future in Clinical Practice". *Universitas Medica*, 61(2), 49-63. DOI: 10.11144/Javeriana.umed61-2.sngs

SAIZ, M., ÁLVAREZ-CUBERO, M. J., GONZÁLEZ, L. J. M., ÁLVAREZ, J. C., Y LORENTE, J. A. (2012): "El ADN antiguo una herramienta para descifrar la historia". Cuadernos de Prehistoria y Arqueología de la Universidad de Granada, 22, 11-47 DOI: 10.30827/cpag.v22i0.2414

SCHMITZ, R. W., SERRE, D., BONANI, G., FEINE, S., HILLGRUBER, F., KRAINITZKI, H., ... Y SMITH, F. H. (2002): "The Neandertal type site revisited: interdisciplinary investigations of skeletal remains from the Neander Valley, Germany". Proceedings of the National Academy of Sciences, 99(20), 13342-13347. DOI: 10.1073/pnas.192464099

SMITH, C. I., CHAMBERLAIN, A. T., RILEY, M. S., STRINGER, C., Y COLLINS, M. J. (2003): "The thermal history of human fossils and the likelihood of successful DNA amplification". Journal of human evolution, 45(3), 203-217.

SIMS, D., SUDBERY, I., ILOTT, N. E., HEGER, A., Y PONTING, C. P. (2014): "Sequencing depth and coverage: key considerations in genomic analyses". Nature Reviews Genetics, 15(2), 121-132. DOI: 10.1038/nrg3642

STERN, D. L. (2013): "The genetic causes of convergent evolution". *Nature Reviews Genetics*, 14(11), 751-764. DOI: 10.1038/nrg3483

SOLARI, A. J. (2004): "Genética humana: fundamentos y aplicaciones en medicina". Ed. Médica Panamericana.

SON, M. S., Y TAYLOR, R. K. (2011): "Preparing DNA libraries for multiplexed paired-end deep sequencing for Illumina GA sequencers". Current protocols in microbiology, Unit1E-4. DOI: 10.1002/9780471729259.mc01e04s20

TIGANO, A., Y FRIESEN, V. L. (2016): "Genomics of local adaptation with gene flow". *Molecular ecology*, 25(10), 2144-2164. DOI: 10.1111/mec.13606.

TUTAR, Y. (2012): "Pseudogenes". *Comparative and functional genomics*,

2012. DOI: 10.1155/2012/424526

TURBÓN, D., DOMÍNGUEZ, E. F., PÉREZ-PÉREZ, A., Y PARDO, E. A. (2003): "ADN antiguo: química y aplicaciones". In *Anales de la Real Sociedad Española de Química* (No. 3, pp. 27-37). Real Sociedad Española de Química.

VAN DIJK, E. L., AUGER, H., JASZCZYSZYN, Y., Y THERMES, C. (2014): "Ten years of next-generation sequencing technology". *Trends in genetics*, 30(9), 418-426. DOI: 10.1016/j.tig.2014.07.001

VENTER, J. C., ADAMS, M. D., MYERS, E. W., LI, P. W., MURAL, R. J., SUTTON, G. G., ... Y KALUSH, F. (2001): "The sequence of the human genome". *science*, 291(5507), 1304-1351. DOI: 10.1126/science.1058040

WILLERSLEV, E., Y COOPER, A. (2005): "Ancient dna". *Proceedings of the Royal Society B: Biological Sciences*, 272(1558), 3-16. DOI: 10.1098/rspb.2004.2813

YOUNG, A. D., Y GILLUNG, J. P. (2020): "Phylogenomics—principles, opportunities and pitfalls of big- data phylogenetics". *Systematic Entomology*. DOI:10.1111/syen.12406

XING, J., WITHERSPOON, D. J., RAY, D. A., BATZER, M. A., Y JORDE, L. B. (2007): "Mobile DNA elements in primate and human evolution". *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, 134(S45), 2-19. DOI:10.1002/ajpa.20722