

Universidad Católica de Santa María

Facultad de Ciencias Farmacéuticas, Bioquímicas y Biotecnológicas

Escuela Profesional de Farmacia y Bioquímica



ENFERMEDAD DE ALZHEIMER Y MUTACION GENÉTICA EN FAMILIA AREQUIPEÑA

Tesis presentada por la Bachiller:

**Ortiz Manrique, Michelle
Milagros**

Para optar el Título Profesional de:

Químico Farmacéutico

Asesora:

PhD. Vera López, Karin

Arequipa – Perú

2021



DEDICATORIA

Esta tesis está dedicada a:

A mis padres, por su trabajo y sacrificio, para poder brindarme una educación superior.
Es un orgullo ser su hija.

A mis hermanas por su apoyo infinito y confianza para mi persona, saben que estaré para apoyarlas.

Al conjunto de doctores y profesores del proyecto Biomarcadores 024-2019, que me brindaron su apoyo y conocimiento para completar esta meta.

A mis amigas, hermanas del alma, por su apoyo incondicional.

A todos aquellos que me apoyaron en la escritura y culminación de la tesis.

AGRADECIMIENTOS

Un trabajo de investigación es fruto de ideas, proyectos y esfuerzos previos que corresponden a otras personas. En este caso mi más sincero agradecimiento al Dr. Christian Lino Cárdenas por su amabilidad para facilitarme su tiempo e ideas. Por su orientación y atención a mis consultas, mi agradecimiento a la Dra. Karin Vera López. Mi agradecimiento al Dr. Miguel Chavez Fumagalli y al Dr. Jorge Aguilar Pineda, por el material facilitado y las sugerencias recibidas. Finalmente, gracias a la Dra. Maria Elena Guillen Nuñez, Dra. Jenny Lopez Valencia y Dra Roxana Gutierrez Aranibar, por la revisión cuidadosa que han realizado de este texto y sus valiosas sugerencias en momentos de duda. Este proyecto también es fruto de horas de estudio y apoyo expresado por aquellos que me tienen estima y me brindaron la fuerza y energía para continuar y crecer como persona y profesional.

También deseo expresar mi gratitud a Dios, por nunca abandonarme y por siempre estar allí cuando las cosas se tornan grises, sé que en ti puedo confiar y guiarás mi camino. El camino no fue sencillo, pero se cumplió con la meta.

AGRADECIMIENTOS

A Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC) – Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica (FONDECYT), por el financiamiento brindado para el desarrollo de la investigación titulada: “Enfermedad de Alzheimer y mutación genética en familia arequipeña”, mediante el proyecto: “Diseño y elaboración de una prueba de diagnóstico basada en nuevos biomarcadores para el diagnóstico y monitoreo de la Enfermedad de Alzheimer”



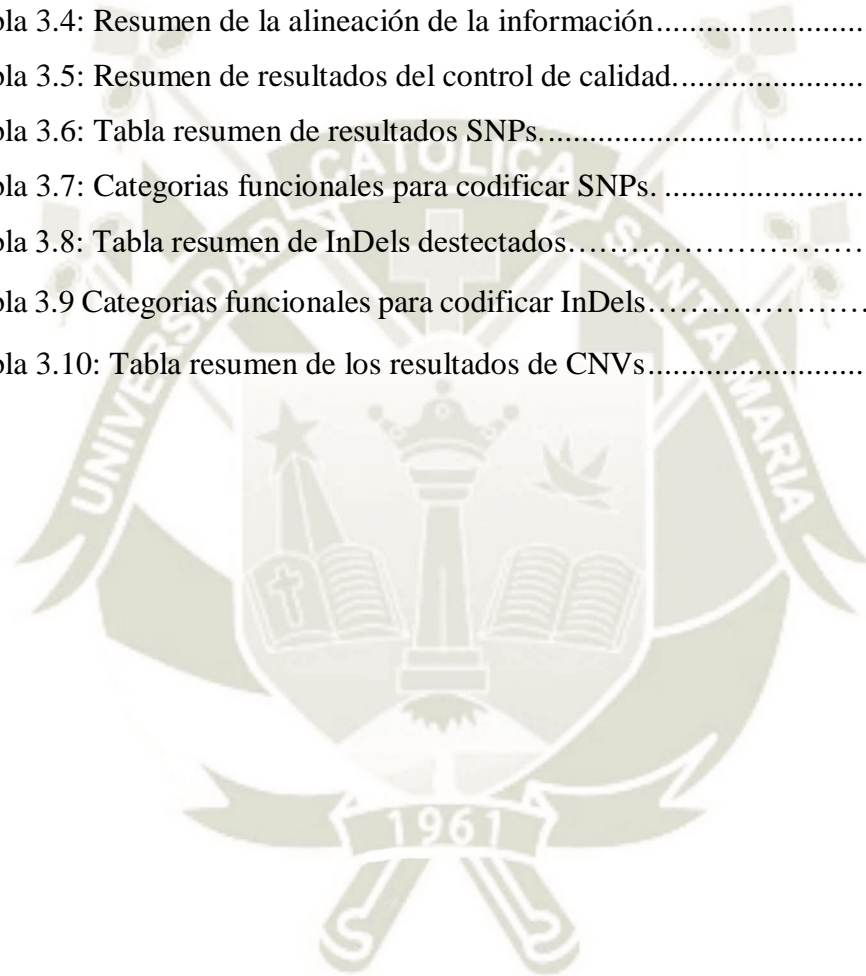
ÍNDICE DE CONTENIDO

DEDICATORIA	ii
AGRADECIMIENTOS.....	iii
ÍNDICE DE CONTENIDO	v
INDICE DE TABLAS.....	vii
INDICE DE FIGURAS	viii
RESUMEN	ix
ABSTRACT.....	x
INTRODUCCION	xi
HIPOTESIS	xiii
OBJETIVOS	xiii
Objetivo General.....	xiii
Objetivos Específicos.....	xiii
CAPITULO I	1
MARCO TEÓRICO	1
1. ENFERMEDAD DE ALZHEIMER (EA).....	1
1.1. CLASIFICACION	2
1.2. FISIOPATOLOGIA DE LA EA	4
A. EL ROL DEL β AMILOIDE ($A\beta$) EN LA EA	4
B. EL ROL DE LA PROTEINA TAU EN LA EA	6
C. EL ROL DE LA APOLIPOPROTEINA E (Apo E) EN LA EA	7
1.3. SIGNOS Y SINTOMAS	8
1.4. DIAGNÓSTICO	9
1.5. TRATAMIENTO SINTOMATICO	9
1.6. TRATAMIENTO MODIFICADOR DE LA ENFERMEDAD	10
2. BANCO DE TEJIDOS	10
3. BIOBANCO	11
3.1. DEFINICIONES.....	11
3.2. CLASIFICACION	12
3.3. BIOETICA Y SEGURIDAD EN BIOBANCO	13
4. GENETICA.....	14

4.1. MUTACIÓN.....	15
5. ESTUDIO GENETICO.....	17
5.1. TECNICAS PARA LA IDENTIFICACIÓN DE GENES.....	18
CAPITULO II.....	20
MATERIALES Y METODOS.....	20
1. LUGAR DE EJECUCION.....	20
2. MATERIALES.....	20
2.1. Muestras.....	20
2.2. Materiales.....	20
2.3. Insumos y Reactivos Químicos.....	21
2.4. Equipos e Instrumentos.....	22
2.5. Softwares.....	22
3. METODOS.....	23
3.1. SUJETOS.....	23
4. IMPLEMENTACIÓN DEL BIOBANCO:.....	23
5. RECOLECCION DE MUESTRAS.....	24
5.1. CODIFICACION DEL TUBO DE MUESTRA.....	25
5.2. ALMACENAMIENTO.....	25
6. PURIFICACION DEL DNA:.....	26
7. AMPLIFICACION DEL DNA.....	26
8. SECUENCIACION DEL DNA AUTOMATICA.....	27
9. VALIDACION DE LA MUTACION.....	27
CAPITULO III.....	31
RESULTADOS Y DISCUSIÓN.....	31
1. SUJETOS.....	31
2. EXTRACCIÓN Y SECUENCIAMIENTO DEL DNA.....	34
3. GEN: COMPLEJO DEL TRÁFICO DE PROTEINAS 12 (<i>Trafficking Protein Particle Complex 12</i> , TRAPPC 12).....	41
4. VALIDACIÓN DE LA MUTACIÓN.....	45
CONCLUSIONES.....	57
SUGERENCIAS.....	58
REFERENCIAS BIBLIOGRAFICAS.....	59
ANEXOS.....	62

INDICE DE TABLAS

Tabla 1.1: Cuadro comparativo de dos mutaciones en el gen PSEN 1	3
Tabla 1.2: Clasificación de mutaciones.....	16
Tabla 3.3: Resumen de información obtenida tras la secuenciación.....	34
Tabla 3.4: Resumen de la alineación de la información.....	35
Tabla 3.5: Resumen de resultados del control de calidad.....	37
Tabla 3.6: Tabla resumen de resultados SNPs.....	37
Tabla 3.7: Categorías funcionales para codificar SNPs.	38
Tabla 3.8: Tabla resumen de InDels detectados.....	39
Tabla 3.9 Categorías funcionales para codificar InDels.....	39
Tabla 3.10: Tabla resumen de los resultados de CNVs.....	40



INDICE DE FIGURAS

Figura 1. 1: Esquema De Un Heredograma.....	18
Figura 1.2: Heredograma del caso clínico.....	32
Figura 1.3 Tubo colector de saliva con código asignado.....	33
Figura 3.3: Distribución de la profundidad de secuenciación de los pares de bases del genoma entero.....	36
Figura 3. 4: Distribución acumulada del genoma completo.....	37
Figura 3.5: Distribución de longitudes de variantes codificantes de InDels.....	38
Figura 3.6: Ubicación del gen dentro de la estructura del cromosoma.	42
Figura 3.7: Expresion del gen TRAPPC 12 en órganos del cuerpo humano.....	42
Figura 3.8: Representacion de la estructura del gen	43
Figura 3.9: Impresión del gel de agarosa al 2% luego de la corrida electroforética.. ...	47
Figura 3.10: Fragmento amplificado de la mutación..	47
Figura 3.11: Gráfico obtenido a partir del paquete estadístico de R, SEURAT.	49
Figura 3.12: Estructuras iniciales de las proteínas TRAPCC12, P697 y V697.....	51
Figura 3.13: Resultado de las simulaciones de DM con las proteínas TRAPPC12.	53
Figura 3.14: Análisis estructural y vibracional de los sistemas en estudio.	55

RESUMEN

La enfermedad de Alzheimer es considerada la causa principal de demencia en personas mayores de 65 años. Este proyecto tuvo como objetivo principal determinar la relación entre la Enfermedad de Alzheimer (EA), y una mutación genética que podría ser propia en nuestra población. Se colectaron muestras de saliva con el tubo Orangene DNA a 9 miembros de una familia, los cuales padecen de EA y enfermedades cardiovasculares. Se extrajo el DNA de estas muestras con el kit PrepIT-L2P y para determinar las variaciones o mutaciones genéticas se realizó un análisis completo del genoma con la plataforma BGISEQ-500.

Luego de comparar los resultados con diversas bases de datos genéticos, se encontró una mutación de tipo Missense en el gen TRAPPC 12, el cual participa en el transporte y comunicación de partículas entre el aparato Golgi y el Retículo Endoplasmático. Su manifestación está relacionada con un tipo de encefalopatía que ataca a infantes provocando retraso en su desarrollo, pérdida de la audición y microcefalia. El análisis genético mostró que en la posición 2089 del genoma, en lugar de una Citosina se encuentra una Guanina, lo que provoca que en la proteína resultante en la posición 697 se encuentre el aminoácido Valina y no una Leucina.

Adicionalmente, se desarrolló un biobanco que permitió realizar el estudio genético, se identificó el gen y su tipo de mutación, la cual fue validada con la técnica de la reacción en cadena de la polimerasa (PCR) y electroforesis en gel de agarosa al 2%. También se realizó un análisis que emplea un conjunto de tecnologías, definido como secuenciación individual de una célula (scRNA-seq), en el cual se comparó la expresión del gen TRAPPC 12 en células de un cerebro sano y células de un cerebro con EA. Finalmente se realizó la dinámica molecular, para identificar la variación que presenta la proteína ante la mutación. El diseño del estudio es descriptivo, observacional, prospectivo y transversal.

Se concluyó que la mutación hallada es expresada en poca proporción, según las bases de datos genéticas, y podría ser propia de una población en específico. También se implementó un biobanco que permite almacenar muestras e información bajo una codificación y organización que certifica la confidencialidad de los datos.

Palabras clave: Enfermedad de Alzheimer, biobanco, TRAPPC 12, análisis genético.

ABSTRACT

Alzheimer's disease is considered the leading cause of dementia in people over 65 years of age. The main objective of this project was to determine the relationship between Alzheimer's disease (AD) and a genetic mutation that could be unique to our population. Saliva samples were collected with the Orangen DNA tube from 9 members of a family whose members suffer from AD and cardiovascular diseases, the DNA of these samples was extracted with the PrepIT-L2P kit. To discover the genetic variations in this project we performed whole genome sequencing with the BGISEQ-500 platform.

After comparing the results with various genetic databases, a Missense-type mutation was found in the TRAPPC 12 gene, which participates in the transport and communication of particles between the Golgi apparatus and the Endoplasmic Reticulum. Its manifestation is related to a type of encephalopathy that attacks infants causing delayed development, hearing loss and microcephaly. Genetic analysis showed that in position 2089 of the genome, instead of a Cytosine, a Guanine is found, which causes the amino acid Valine to be found in the resulting protein at position 697 and not a Leucine.

Additionally, a biobank was developed that allowed the genetic study to be carried out, the gene and its type of mutation were identified, which was validated with the polymerase chain reaction (PCR) technique and 2% agarose gel electrophoresis. An analysis was also performed using a set of technologies defined as single cell sequencing (scRNA-seq), the expression of the TRAPPC 12 gene in cells from a healthy brain was compared to cells from a brain with AD. Molecular dynamics was also carried out to identify the variation that the protein presents before the mutation. The study design is descriptive, observational, prospective, and cross-sectional.

It was concluded that the mutation found is rare and could be typical of a specific population. A biobank was also implemented that allows samples and information to be stored under a code and organization that certifies the confidentiality of the data.

Keywords: Alzheimer's Disease, biobanks, TRAPPC 12, whole genome analysis.

INTRODUCCION

Por el año 2015 cerca de 47 millones de personas fueron diagnosticadas con demencia. Se espera que para el año 2050 existan cerca de 130 millones de personas afectadas por este mal. En Latinoamérica tiene una prevalencia del 8.5% y en Perú se dispone de poca información sobre la cantidad de ciudadanos afectados. Pero la demencia puede tener diferentes causas, la principal es la Enfermedad de Alzheimer, su principal síntoma es la pérdida de memoria acompañada con episodios de desorientación espacial, depresión y rabia hasta provocar una degeneración completa que imposibilita al paciente realizar tareas cotidianas. Según su etiología, aún no definida, se conocen tres genes involucrados: el gen de la proteína precursora de beta amiloide (APP), las presenilinas (PSEN1, PSEN2) y el gen la apolipoproteína E (APO E).

El objetivo principal el presente trabajo fue determinar la relación entre la Enfermedad de Alzheimer y una mutación genética, también busco identificar algún gen que podría ser propio de una población específica y desarrollar un biobanco que permita realizar el estudio genético. A nivel mundial se están desarrollando biobancos, instituciones que albergan muestras asociadas información como datos clínicos, para investigar diversas enfermedades cuyas causas o tratamientos aún no están definidos. Estas pueden ser empleadas como herramientas en el avance de la investigación sobre enfermedades neurodegenerativas como la Enfermedad de Alzheimer, identificar factores genéticos que predisponen a una población a manifestarla y quizá encontrar genes propios de una población que brinden características únicas a la enfermedad.

En el capítulo I se encuentra el marco teórico el cual aborda conceptos sobre la Enfermedad de Alzheimer, Biobanco, genética y estudio genético. Se describe posibles causas, tratamiento y diagnóstico de la enfermedad. También se plasma diversos conceptos sobre el biobanco y sus tipos. Se habla sobre conceptos básicos de genética y los tipos de mutaciones que uno encuentra en la naturaleza.

En el capítulo II, se describe la metodología del proyecto. Forma en la cual se recolectaron las muestras de saliva, la extracción de DNA con el kit PrepIT-L2P, seguido por el análisis genético realizado en la plataforma BGISEQ-500 y la validación de la mutación hallada empleando la técnica de reacción en cadena de la polimerasa, electroforesis en gel de agarosa al 2%.

En capítulo III se establecen los resultados obtenidos tras la secuenciación completa del genoma, la comparación de la mutación hallada con bases genéticas y también las técnicas empleadas para validar la mutación como la single cell RNA sequencing y la dinámica molecular de proteínas. Además, establece la discusión e interpretación de los resultados y se presentan las conclusiones y recomendaciones de esta tesis.



HIPOTESIS

Dado que existe relación de la Enfermedad de Alzheimer y la presencia de una mutación genética es probable que en una familia de Arequipa se exprese esta enfermedad por alteración genética.

OBJETIVOS

Objetivo General

- Determinar la relación que existe entre la enfermedad de Alzheimer y la mutación genética en una familia de Arequipa.

Objetivos Específicos

- Implementar un banco de tejidos para estudio de la Enfermedad de Alzheimer.
- Identificar el gen y su tipo de mutación relacionado con la Enfermedad de Alzheimer.
- Validar el tipo de mutación encontrada en el estudio.

CAPITULO I

MARCO TEÓRICO

1. ENFERMEDAD DE ALZHEIMER (EA)

En 1907 Alois Alzheimer, psiquiatra y neuropatólogo alemán, publicó un trabajo sobre los síntomas de Auguste D. Una mujer de 50 años ingresada al Hospital Psiquiátrico de Frankfurt que sufría de trastornos de memoria, paranoia y confusión progresiva. Luego, describió la historia clínica de Johann F, un paciente olvidadizo, incapaz de orientarse y realizar tareas cotidianas. Después de su muerte y autopsia Alzheimer describió alteraciones histológicas en su cerebro, conocidas ahora como las placas amiloideas y ovillos neurofibrilares. La persona que acotó el nombre para esta enfermedad fue Emil Kraepelin, en honor a su mentor Alzheimer. (1)

La EA, considerada como un tipo de demencia, es una enfermedad neurodegenerativa en la cual el paciente presenta síntomas como episodios de agresividad y desorientación espacial, pero el principal es la pérdida de memoria. Se manifiesta, comúnmente, en personas mayores de 65 años. Aunque, existen casos en los cuales se desarrolla antes de los 65 años. El daño empieza en el hipocampo y corteza entorrinal en el lóbulo temporal, zonas encargadas del aprendizaje y memoria, afectando también la neocorteza cerebral.(2)

Esta enfermedad presenta un perfil neuropatológico específico, definido por la:

- Deposición extracelular de β amiloide en forma placas difusas y neuríticas.
- Presencia de ovillos neurofibrilares intraneuronales.
- Presencia de hilos de neuropilos dentro de neuritas distróficas compuestos por agregados de proteína TAU hiperfosforilada.(2)

La EA no tiene una causa definida, se cree que una combinación de factores ambientales y genéticos podrían desencadenar la manifestación de los síntomas. El factor de riesgo más importante es la edad. Existen factores de riesgo ambientales como enfermedades cardiacas, hipertensión, suplementos de estrógenos, fumar, accidente cerebro vascular (ACV), traumatismo en la cabeza, depresión, artritis y diabetes. Es posible que ciertos estilos de vida disminuyan el riesgo de padecer EA

como el practicar ejercicio, la estimulación intelectual y llevar una dieta Mediterránea que incluya pescado. (3)

1.1. CLASIFICACION

Se pueden establecer divisiones bajo dos criterios. El primero considera la edad de manifestación de los primeros síntomas. El segundo se enfoca en la historia familiar, así mismo ambos criterios establecen dos clasificaciones más. Bajo el primer criterio, podemos identificar dos grupos:

- EA temprana o familiar: Representa aproximadamente el 30 % de los casos. Se manifiesta a partir de los 45 años, presenta una herencia tipo autosómica dominante cuyos genes responsables podrían ser tres. El gen que codifica la Proteína Precursora Amiloide (APP) ubicada en el cromosoma 21, la Presenilina 1 (PSEN1) ubicada en el cromosoma 14 y la Presenilina 2 (PSEN2) ubicada en el cromosoma 1. (4)
- EA tardía o esporádica: Representa el 70% de los casos. Se desarrolla a partir de los 65 años. Está relacionada con el estilo de vida, proceso de envejecimiento y variabilidad genética. Los estudios asociados al genoma (GWAS), identificaron variantes en los loci de genes involucrados en la manifestación de esta, como el receptor de complemento 1 (CR1), la apolipoproteína J (CLU), la proteína de ensamblaje de clatrina unida al fosfatidilinositol (PICALM) y el integrador de uniones 1 (BIN1), antígeno de superficie de células mieloides (CD33) y proteína asociada a CD2 (CD2AP).(3) También el alelo $\epsilon 4$ de la apolipoproteína E (apo E) se ha identificado como un factor de riesgo. Este participa en la unión y transporte de lipoproteínas a su diana.(3)

Al conocer el fuerte carácter genético de esta enfermedad, se han identificado diversas mutaciones que brindan características únicas a la EA en poblaciones específicas. Se describen mutaciones en el gen APP, localizado en la región 21 del brazo largo del cromosoma 21 y está constituido por 17 exones. Los dos últimos codifican la porción que constituye el fragmento β amiloide, después de la proteólisis. Se han descrito mutaciones en el codón 717 en el cual el aminoácido Valina es sustituida por una Fenilalanina o por una Glicina.(5)

También se dan mutaciones en el gen PSEN1, localizado en la región 24.3 del brazo largo del cromosoma 14 y el gen PSEN 2, situado en la región 31-42 del brazo largo del cromosoma 1. Constituidos por 12 exones, presentan una zona reguladora la cual corresponde a los exones 0-2 y una zona codificante que incluye exones del 3 al 12. Estos presentan zonas calientes en las cuales encuentran un gran número de mutaciones, estas son la zona TM2, ubicada en el exón 5; la zona TM3, ubicada en el exón 6; la zona TM6, ubicada en la parte 3' del exón 7 y 5' del exón 8 y el loop hidrofílico, codificado por el exón 8. Un hecho particular es que, si la mutación se da en alguna de estas zonas, caracteriza la edad específica de debut de la enfermedad.(5) En la tabla 1 se describe dos ejemplos de mutaciones en el gen PSEN 1

Tabla 1.1: Comparación entre dos mutaciones en el gen PSEN 1

Diferencias	Mutación “Paisa” o E280A	Mutación “Jalisco” o A431E
Estudio realizado por	Francisco Lopera	Petra Yescas et al.
Ciudad de origen	Antioquia, Colombia.	Jalisco, México.
Síntomas	A los los 30 años pérdida de memoria y dificultad verbal. A los 45 años, disminución cognitiva. A los 50 años, un diagnóstico definitivo de EA.	Manifestación de los primeros síntomas de EA en una edad promedio de 40 años.
Esperanza de vida	10 años.	10 años.
Mutación	Se produce la transversión de Adenina por Citosina en el codón 280, lo que provoca el cambio del aminoácido Ácido Glutámico por Alanina	Mutación en el exón 12, en el cual se da una sustitución cuyo resultado es el cambio del aminoácido Alanina por el Ácido Glutámico.

En Perú se da una mutación en la Presenilina 1(L153V), la cual es responsable que, miembros de una familia en el norte manifiesten EA a partir de los 45 años.(6)

La identificación de esta mutación, nos indica que existen sectores de la población que manifiestan EA con características específicas, lo que induce a realizar mayores estudios para poder identificar el origen, brindar el tratamiento adecuado y mejorar la calidad de vida de estos pacientes.

1.2.FISIOPATOLOGIA DE LA EA

En la EA observamos el acúmulo de placas extracelulares de beta amiloide ($A\beta$) y ovillos neurofibrilares de proteína tau hiperfosforilada. Se piensa que el acumulo de $A\beta$ es debido a que los mecanismos de defensa del sistema nervioso y su depuración fallan. Activando así una cascada neurotóxica que conlleva a la hiperfosforilación de proteína TAU y apoptosis neuronal. (4) También se observa que la Apolipoproteína E (ApoE) juega un rol importante en la manifestación de la enfermedad.

A. EL ROL DEL β AMILOIDE ($A\beta$) EN LA EA

Es una proteína de 4 kD de peso molecular y el principal constituyente de las placas extracelulares. Esta se origina a partir de la proteólisis de la Proteína Precursora Amiloide (APP), por enzimas conocidas como secretasas. Es un producto normal del metabolismo de APP y es generado, en altos niveles, en las neuronas durante la vida de una persona en forma de un grupo heterogéneo de péptidos de diferente longitud. La relación con la EA se observa en estudios realizados en pacientes con EA temprana, al manifestarse una sobreproducción de $A\beta$, principalmente el de 42 aminoácidos.(2)

La APP es una proteína transmembrana que consta con tres dominios; un dominio extracelular grande, un dominio transmembrana hidrofóbico y un dominio citoplasmático corto. El dominio extracelular es complejo. Este contiene una región globular rica en cisteínas y sitios de unión para heparina, cobre y zinc; un dominio ácido rico en glutaminas, aspartatos y treoninas y dominios glicosilados que incluyen una región promotora de crecimiento. Es sintetizada en el retículo endoplasmático, modificada en el aparato de Golgi y

transportada a la superficie por la vía secretora. Endocitada y metabolizada por la vía lisosomal. (7)

Esta sigue dos vías metabólicas, en la que participan las enzimas:

- α -secretasa: Participa en la vía no amiloidogénica. Pertenece a la familia A desintegrasa y metaloproteinasa (ADAM). Produce el A β soluble.(7)
 - β -secretasa: También conocida como enzima 1 cortadora del sitio β del APP (BACE1). Es una proteína transmembrana que pertenece a la familia de la pepsina de las aspartilproteasas. Participa en la vía amiloidogénica. Responsable de la producción el A β insoluble.(7)
 - γ -secretasa: Participa en ambas vías, compuesta por un complejo enzimático compuesto por 4 subunidades. (7)
 - Presenilina (PSEN): PSEN 1
 - Presenilina potenciador (PEN 2)
 - APH 1: APH A, APH B, APH C
 - Nicastrina
- 1) Vía no amiloidogénica: La α -secretasa corta al APP en el aminoácido 83 del extremo C-terminal, generando un fragmento C terminal (CTF) de 83 residuos retenido en la membrana y un fragmento N-terminal (sAPP) soluble de gran tamaño liberado en el espacio extracelular. Posteriormente la γ -secretasa divide el CTF, esto produce un fragmento corto no tóxico llamado p3. Esta enzima corta al APP en la secuencia del A β , previene la formación de este. (4,7)
 - 2) Vía amiloidogénica: La β -secretasa corta al APP en el extremo amino terminal de la secuencia de A β , liberando sAPP β y fragmento C99 anclado a membrana. Luego la γ -secretasa se une al extremo amino terminal del fragmento (C99) e intramembrana corta el ϵ -site liberando CTF y A β 48. El complejo procesa el A β C- Terminal en péptidos cortos hasta que A β es liberado del complejo generalmente después de producir péptidos de 38, 40, 42 aminoácidos de largo. Estos tienden a agregarse en oligómeros, protofibrillas y fibrillas. (4,7)

A β 42, es el más propenso a agregarse, contiene dos aminoácidos de más, elevando su hidrofobicidad y su capacidad de oligomerizar y polimerizar más que las demás isoformas. El complejo γ -secretasa presenta su actividad catalítica en la presenilina, es por ello que las mutaciones de PSEN1 y PSEN2 del complejo γ -secretasa favorecen la formación de residuos (A β 42).(4,7)

B. EL ROL DE LA PROTEINA TAU EN LA EA

Las taupatias, son enfermedades neurodegenerativas, cuya manifestación patológica común es la presencia de agregados de proteína TAU en el cerebro. Estudios sugieren que formas hiperfosforiladas de TAU y sus agregados pueden dirigir hacia la neurodegeneración en EA. Se observa que el estado de la taupatia esta correlacionada con la progresión del deterioro cognitivo. Es decir que agregados de TAU, con el paso de tiempo, se acumulan en la corteza entorrinal y lóbulos temporales medianos sin que la persona presente algún signo de declive cognitivo, esta etapa es conocido como taupatia relacionada con la edad primaria (PART), el momento en el que estos agregados se esparcen de la corteza entorrinal a la neocorteza, se manifiesta el deterioro cognitivo. (4,8)

La proteína TAU cumple funciones estructurales como la estabilización de microtúbulos y axones neuronales, también la regulación de transporte de microtúbulos. En nuestro organismo esta proteína es soluble de estructura desordenada, bajo condiciones correctas forma oligómeros y fibrillas. Ciertos cambios en sus isoformas y su hiperfosforilación provoca que esta se torne insoluble, enrollándose hasta desprenderse de los microtúbulos. Empezando así a acumularse y creando filamentos helicoidales pares junto con fibrillas de amiloide, los cuales encontramos formando parte de los ovillos neurofibrilares. (4,9)

La proteína TAU esta codificada por el gen microtúbulo asociado a la proteína tau (MAPT) en el cromosoma 17 y es expresado en neuronas del cerebro. Esta proteína puede ser cortada alternativamente a nivel del N- terminal, creado un

dominio (N) y un dominio repetido para la unión de microtúbulos (R). Existen 6 isoformas diferentes:(4)

1. 0N3R
2. 0N4R
3. 1N3R
4. 1N4R
5. 2N3R
6. 2N4R

Tau puede sufrir modificaciones como la fosforilación que disminuye la afinidad de unión con el microtúbulo. Actúan 3 grupos de cinasas en el proceso de fosforilación:(4)

- Cinasa dirigida a prolina o motivos prolin-serina/treonina.
- Cinasa no dirigida a prolina.
- Tirosina cinasa.

Ejemplos GSK, CdK5, MAPK, CaMKII

Los neurofilamentos contienen isoformas con dominio 3R y 4R de la proteína TAU y existen otras isoformas que son sobre producidas en otras taupatias.(4)

C. EL ROL DE LA APOLIPOPROTEINA E (Apo E) EN LA EA

Identificado como el principal factor de riesgo genético para desarrollar EA, es una glicoproteína de 299 aminoácidos cuyo peso molecular es de 34 200 daltons. Constituida por una cadena peptídica que posee ocho secuencias repetidas de 22 aminoácidos con naturaleza anfipática, se reconocen dos dominios estructurales definidos por la región amino terminal en la cual se encuentra el sitio activo y la región carboxiterminal, en donde se encuentra el sitio de unión a lípidos. Se expresa en el hígado y principalmente en el cerebro por astrocitos y microglías. (10)

Apo E participa en la distribución del colesterol, entre las células de los órganos donde se produce y se usa para diferentes procesos de biosíntesis, síntesis de membranas o la producción de hormonas esteroideas. Entre otras funciones que cumple que encuentra la inmunorregulación, regeneración de nervios, y la activación de enzimas lipolíticas (lipasa hepática lipoproteica y

lecitin-colesterol acetiltransferasa).(10) Pero su función principal es el reconocimiento de receptores específicos, como el receptor de lipoproteína de baja densidad (LDLR) que media la internalización de partículas de lipoproteínas. (11,12)

El gen apo E tiene 3 alelos polimórficos diferentes que corresponden a variaciones en las combinaciones de aminoácidos en residuos 112 y 158 de la proteína. (11)

- Apo E2: Cuya frecuencia de expresión es 8%. Sus aminoácidos característicos son cys112/cys158. Esta disminuye el desarrollo de EA.
- Apo E3: Sus aminoácidos son cys 112/ arg 158. Común en población caucásica en un 78%.
- Apo E4: Principal factor de la Enfermedad de Alzheimer con una frecuencia del 14%. Sus aminoácidos característicos son arg 112/arg 158.(4)

Las diferentes sustituciones de los aminoácidos afectan la carga y estructura de la APOE, alterando el lugar de unión a receptores y a lipoproteínas, cambiando su estabilidad, producción y depuración.(2)

1.3.SIGNOS Y SINTOMAS

Esta enfermedad sigue un curso silente y progresivo, puede ser dividida en tres fases:

1. Fase inicial: Predemencia, el individuo afectado manifiesta episodios de pérdida de memoria, que pasan inadvertidos. Junto con una dificultad para recordar hechos recientes y adquirir nuevos conocimientos. También, trastornos en la memoria semántica y ciertas funciones ejecutivas, apatía.(13)
2. Fase intermedia: El paciente presenta periodos de pérdida de memoria esporádica e inusual hasta episodios persistentes. Muestra dificultad para el lenguaje, ejecutar movimientos y en el reconocimiento de personas. (13)
3. Fase avanzada: En el paciente disminuye la coordinación de movimientos, capacidades como la lectura y escritura empeoran progresivamente. Se produce el deterioro de la masa muscular perdiéndose la movilidad,

incontinencia urinaria y posible muerte por causas externas. El lenguaje se torna severamente desorganizado llegándose a perder completamente. A pesar de ello, se conserva la capacidad de recibir y enviar señales emocionales. Los pacientes no realizan las tareas más sencillas por sí mismos y requieren constante supervisión. (25)

Como síntomas de la enfermedad podemos observar:

- Alteraciones en el estado de ánimo y conducta: Episodios de depresión, frustración y arrebatos de violencia.(13)
- Pérdida de memoria: El paciente no reconoce a personas cercanas. (13)
- Alteraciones visoespaciales: Pérdida del sentido de orientación. (13)
- Problemas de lenguaje y alteración cognitiva: Afasia, apraxia, agnosia.(13)

1.4.DIAGNÓSTICO

En pacientes con sospecha de EA el médico realiza los siguientes estudios:

- A. Análisis minucioso de la historia clínica, exámenes de sangre y orina como el dosaje de vitamina B12 y perfil tiroideo para descartar demencias de origen metabólico.
- B. Examen psicológico y neurológico: Aplicación de tests psicológicos como la prueba del reloj.(13)
- C. Análisis de biomarcadores en fluidos cerebro espinales (CSF), sangre y plasma. (4)
 - Beta amiloide 42.
 - Tau total.
 - Tau hiperfosforilada.
- D. Imagenología: Electroencefalograma, tomografía axial computarizada y resonancia magnética.(14)

1.5.TRATAMIENTO SINTOMATICO

- INHIBIDORES COLINESTERASA

Se ha observado que durante el desarrollo de esta enfermedad las neuronas en el núcleo basal de Meynert y núcleo septal mueren lo que provoca un déficit colinérgico y lleva una disfunción temprana de atención y memoria. Utilizados en EA leve a moderada.(4)

- Donepezil 10 mg.
- Rivastigmina 3mg en tableta y en parche.
- Galantamina 8 mg.

- ANTAGONISTA DEL RECEPTOR NMDA

Inhiben la neurotoxicidad mediada por el glutamato que las neuronas desarrollan al morir.(4)

- Memantina 10 mg.

1.6. TRATAMIENTO MODIFICADOR DE LA ENFERMEDAD

Se considera que un tratamiento es efectivo cuando interrumpe la enfermedad tempranamente evitando desencadenar una cascada de eventos degenerativos. Existen medicamentos dirigidos al β amiloide buscan disminuir el nivel de A β parenquimal y depósitos amiloides. Otros inhiben la actividad de las enzimas secretasas, limitando la formación de nuevas placas. Otros dirigidos a la proteína Tau y Apo E. Estos se encuentran en fases de ensayos clínicos pero la mayoría fracasa al no cumplir con el objetivo de mejorar la cognición. (4)

2. BANCO DE TEJIDOS

A lo largo de los años, la comunidad científica, ha dedicado su esfuerzo y tiempo a la recolección de muestras enfocado al desarrollo y creación de información. Durante este proceso las muestras, materia de estudio, siguen una metodología que varía de acuerdo a las características del ensayo clínico, nivel de investigación o a las características intrínsecas y extrínsecas de la muestra. El descubrimiento del genoma humano y su consecuente e imparable desarrollo de aristas en el tratamiento de enfermedades genéticas como la Enfermedad de Alzheimer (15). Crea la necesidad de

desarrollar un establecimiento que certifique, asegure y ordene la información obtenida, además un lugar donde las muestras sean almacenadas bajo un sistema organizado que codifique y preserve correctamente a estas. Existen dos establecimientos que se pueden emplear con este fin, los cuales exigen ser definidos y diferenciados:

- a. Biorrepositorio: Es la infraestructura dentro de la cual se identifican, recopilan, almacenan y distribuyen muestras biológicas. Estos fueron los primeros biobancos, ya que recolectaban muestras e información al azar. (16)
- b. Biobanco: Es una colección de material biológico asociado a información ordenada en un sistema y preservada bajo directrices éticas y de seguridad con fines científicos. (16)

3. BIOBANCO

El biobanco más antiguo se desarrolló para el Estudio Framingham del Corazón (FHS) en 1948, financiado por el Instituto Nacional de Salud - Instituto Nacional del Corazón, Pulmones y Sangre (NIH-NHLBI), en Framingham, Massachusetts. Este recolectaba las muestras de sangre e integraba la información generada a una base de datos con las características del donante, para identificar factores de riesgo relacionados con enfermedades cardiovasculares (ECV) como: el incremento de la presión arterial, aumento del colesterol, tabaquismo, obesidad, edad, entre otros. (17) Pero recién en 1996, el término biobanco fue empleado en literatura científica. En el año 2006 se fundó UK Biobank, con el objeto de mejorar la prevención, diagnóstico y tratamiento de enfermedades que amenazan la vida como cáncer, enfermedades cardíacas, ACV, diabetes, artritis, desordenes oculares, depresión y formas de demencia. Reclutó a 500 000 personas entre 40-69 años desde el 2006 hasta el 2010. Obtuvieron sangre, orina y saliva, para futuros análisis, e información detallada. Esta se convertirá en una herramienta para ayudar a descubrir porque algunas personas desarrollan enfermedades y otras no. (4) Como este biobanco existen otros a gran escala en países como Corea, Estados Unidos, Japón y otros países europeos.

3.1. DEFINICIONES

Diversas instituciones, sociedades y autores han aportado diversos significados al término biobanco. Sociedades internacionales como la Organización para la

Cooperación y el Desarrollo Económico (OCDE), lo define como una colección de material biológico e información asociada y almacenada en un sistema organizado, para la población o un gran sector de esta.(19) Con el texto titulado como “Recomendaciones para Biobancos Humanos y Bases Informáticas en Investigación Genética” se establece que este puede ser empleado para investigaciones de carácter genético.(20) La sociedad de repositorios biológicos y ambientales (ISBER), indica que es la entidad que recibe, almacena, procesa y/o disemina especímenes. Este incluye la locación física, así como el rango completo de actividades relacionadas en su operación.(21)

La institución europea, con mayor participación en el campo relacionado a biobancos, Biobanking and Biomolecular Resources Research Infrastructure (BBMRI). Señala que un biobanco contiene muestras biológicas e información asociada y son materiales esenciales para el avance de la biotecnología, salud humana e investigación y desarrollo en las ciencias de la vida.(22) El Consejo de Gobierno y Ética del biobanco del Reino Unido lo define como colecciones de datos, además de muestras biológicas, desarrolladas específicamente como recursos para la investigación.(23)

Generalmente, la definición de biobanco engloba tres grupos de información:

- Material biológico humano.
- Información adjunta o conectada.
- Temas legales como el consentimiento informado, la seguridad y protección de la información obtenida. (15)

Existen también, colecciones de muestras de especímenes de origen no humano, como plantas y animales, un ejemplo es VetBioBank. La investigación se enfocará en biobancos humanos.

3.2. CLASIFICACION

Se han desarrollado un gran número de biobancos a nivel mundial. El incremento de estos coincide con el éxito obtenido en la secuenciación del genoma humano del año 2003, junto con el desarrollo de nuevas tecnologías y herramientas bioinformáticas.

Su clasificación se realiza bajo diversos criterios como su estructura organizativa, el fin al que son destinadas las colecciones de las muestras o tipos de muestras que serán recolectadas. Por ejemplo, según su estructura organizativa podemos diferenciar una red de biobancos, un biobanco en red o biobanco nacional con fines de investigación.(23)

Al trabajar con biobancos humanos relacionados a la investigación, existen múltiples diseños que podrían ser establecidos de acuerdo a diferentes objetivos. Estos podrían ser divididos en tres grandes grupos:

- **Biobancos Poblacionales:** Su objetivo principal es identificar biomarcadores de susceptibilidad y características propias de la población. Almacenan muestras de ADN y sangre de un gran número de donantes sanos, representativos de una región, país o cohorte étnico. (24)
- **Biobancos Epidemiológicos:** Enfocados en el estudio de una patología, Identifican biomarcadores de exposición. Desarrollados a partir/en base a de estudios de cohorte prospectivo, retrospectivo o caso-control. Almacenan muestras de ADN y marcadores serológicos junto con una gran cantidad de datos clínicos.(24)
- **Biobancos generales para una enfermedad:** También conocidos como biobancos de tumores. Busca identificar biomarcadores de una enfermedad a través de colecciones prospectivas y retrospectivas de tumores, muestras no tumorales y derivados (DNA, RNA y proteínas), asociados a información clínica y a veces a ensayos clínicos. No son desarrollados para un proyecto de investigación específico excepto en caso de ensayos clínicos. (24)

3.3. BIOETICA Y SEGURIDAD EN BIOBANCO

El biobanco ayuda a los investigadores, proporcionando muestras e información en condiciones adecuadas de seguridad y calidad. Los Biobancos deben estar regulados bajo directrices de cada país, con el fin de vigilar que la recolección de información y la recolección de muestras sea segura, respetando valores éticos y la privacidad de cada donante. En Perú existen normas legales que indican como establecer correctamente el biobanco. Así tenemos:

- Directiva para Muestras Biológicas en Ensayos Clínicos Directiva N° sINS/OGAT-V.01.
- DECRETO SUPREMO N° 014-2005-SA Aprueban Reglamento de la Ley General de Donación y Trasplante de Órganos y/o Tejidos Humano.

Para obtener las muestras de donantes o participantes, es necesario la creación y la disposición de un consentimiento informado que deberá ser leído y firmado por los participantes. Este documento especifica con lenguaje claro y preciso, el concepto del estudio, los beneficios y riesgos de este. Así también, se aclara que la información obtenida será manejada con total confidencialidad entre el donante y el investigador y que este es libre de abandonar la investigación en el momento que desee. (25)

4. GENETICA

Es la rama de biología que estudia los genes y mecanismos que regulan la herencia. La genética humana analiza las características que el hombre hereda tanto físicas como mentales. La herencia son las características físicas y biológicas comunes entre una familia, esta determina nuestro fenotipo (características visibles). Es derivada del genotipo, definido como la colección de genes de un individuo almacenados dentro del núcleo de las células. El material genético del ser humano es el DNA, este tiene una estructura de doble hélice y se compacta en estructuras denominadas cromosomas. (26)

A lo largo de la historia diversos científicos han aportado información sobre el material genético. Así se conoce que los primeros escritos sobre genética fueron realizados por Hipócrates, los cuales describían la teoría conocida como “Pangénesis”. La cual manifiesta que el semen se formaba en todas las partes del cuerpo. Luego Aristóteles señaló que el material reproductivo, se formaba a partir de sustancias nutrientes que se desviaban hacia el sendero reproductivo, la madre brindaba la materia prima y el padre definía la forma del embrión. (26)

En 1865 el monje Juan Gregorio Mendel presentó, ante la sociedad de ciencias naturales de Brun, el resultado de sus experimentos sobre la hibridación de plantas. Este trabajo contiene postulados teóricos sobre genética. Mendel explicó su trabajo utilizando términos como factores abstractos, hoy conocidos como genes. Basado en

el cambio de color de las plantas indicó que podían existir estados alternativos de estos factores, las variaciones observadas radicaban debido a la existencia de alelos. Si estos son iguales, el individuo es “puro”, es decir, homocigoto y si son diferentes, es heterocigoto. Actualmente se conoce que las células somáticas son diploides, contienen 46 cromosomas en 23 pares. Las células sexuales son haploides, poseen solo 23 cromosomas. Una enfermedad genética puede saltar una generación (filial 1) y manifestarse en otra generación (filial 2) lo que se conoce como “segregación”. Muchos de estos conceptos fueron ignorados durante años hasta que en 1953 James Watson y Francis Crick publicaron, en la revista Nature, su propuesta sobre la estructura del DNA. Se observó que este era capaz de duplicarse durante la división celular y dirigir la síntesis de proteínas desde el núcleo de células.(27)

En 1990 ante la unión de varias instituciones americanas se inicia el “Proyecto del Genoma Humano”. El cual concluyó, en el año 2003, con la secuenciación completa del genoma e impulso el desarrollo de la bioinformática, esta permite acceder a bases de datos y disponer de la secuencia de bases de cualquier sector del DNA en todos los cromosomas humanos.(27)

4.1. MUTACIÓN

Es la variación espontánea e impredecible en la secuencia de genes que componen el DNA en un ser vivo. Originadas por errores en la replicación del DNA y por la exposición a agentes mutagénicos como radiaciones UV o agentes químicos como el benzopireno. En la naturaleza existen diversos tipos por tal razón su clasificación depende de diversos criterios como células afectadas o efecto. En la tabla 1.2 se describen los tipos de mutaciones.

Tabla 1.2: Clasificación de mutaciones

CRITERIO	CLASIFICACION	DESCRIPCION Y SUBCLASIFICACION		
Según las células afectadas	Germinales	Afecta los gametos, las células reproductivas. Son heredables.		
	Somáticas	Afecta células del cuerpo, crea un individuo mosaico, es decir, 2 líneas celulares con genotipos distintos. No se heredan.		
Según su origen	Espontáneas	Se dan naturalmente, no son dañinas.		
	Inducidas	Exposición agente mutagénico, son dañinas.		
Según su efecto	Perjudiciales	Producen enfermedades		
	Neutras	No producen un cambio		
	Beneficiosas	Proceso evolutivo, adaptación.		
Según la extensión del material genético afectado	Génicas: Afectan la secuencia de nucleótidos del DNA.	Sustitución: se cambia una de las base del DNA	Transición: Intercambio de bases del mismo tipo.	
			Transversión: Intercambio entre bases de diferente tipo.	
		Corrimiento del orden de lectura (frameshift)	Deleciones o inserciones que dan un completo cambio en la lectura de proteína.	
	Cromosómicas:	Afectan el orden de los genes en el cromosoma.	Delección: Pérdida de una porción del cromosoma.	Duplicación: Ganancia de una porción del cromosoma
			Inversión: Cambio en la posición de una porción del cromosoma	
			Translocación: Transposon, segmento de DNA que cambia por completo de lugar.	
	Genómicas: Alteran el número de cromosomas	Aneuploidia	Nulisomia: Pérdida de 2 cromosomas Monosomia: Pérdida de un cromosoma Disomia: Ganancia de un cromosoma Trisomía/tetra: Ganancia de 3 o más cromosomas	
			Euploidias: Afecta un juego cromosómico.	Monoploidia Poliploidias

5. ESTUDIO GENETICO

Es un tipo de investigación en la cual se analizan los genes de un individuo o una población, con el propósito de identificar mutaciones que expresen enfermedades, También es empleado para estudiar la acción de fármacos y la relación del efecto del ambiente sobre el individuo y enfermedad para desarrollar tratamientos dirigidos contra un defecto genético específico. Para ello se toma muestras de tejidos como sangre, este es examinado a través de diferentes instrumentos y diversas técnicas moleculares para obtener ADN. El análisis del ADN nos permite identificar lesiones específicas en genes y diagnosticar enfermedades oportunamente, para brindar terapias que controlen el daño generado, modificando la expresión de un gen.(28)

Un estudio genético puede ser desarrollado a partir de una historia clínica dirigida, una exploración clínica general o física y exámenes complementarios destinados al diagnóstico. Al describir la historia clínica, se debe contar con informes del individuo y familiares. Una forma práctica para resumir es con la construcción del árbol genético o heredograma. Definido como, un diagrama esquemático de la familia que proporciona de manera gráfica la relación entre los miembros de la familia y enfermedades que podrían tener un significado genético. Se debe emplear símbolos que son universalmente admitidos. En la Figura 1.1 se observa un ejemplo.

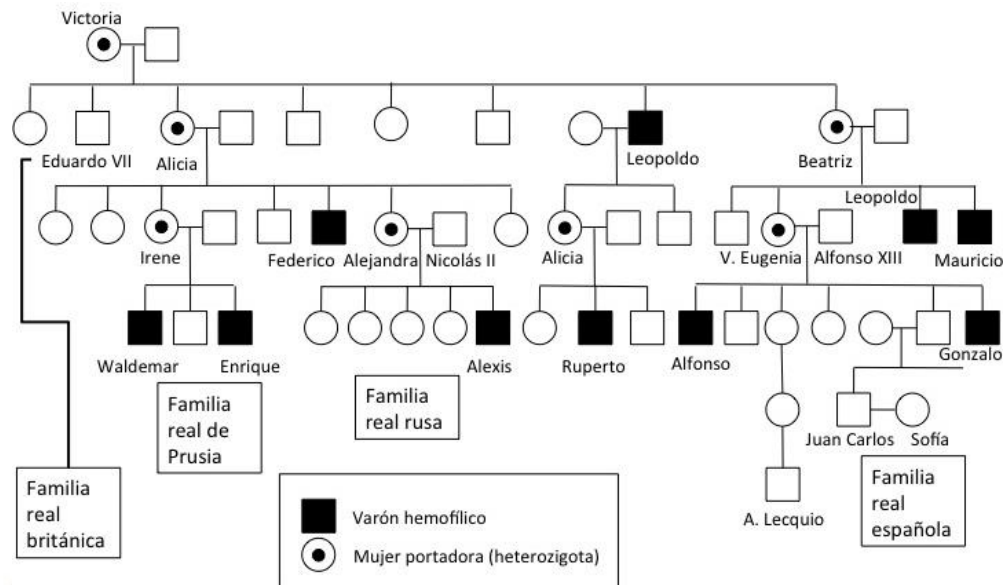


Figura 1. 1: Esquema De Un Heredograma. La imagen especifica el orden genealógico de la familia real y el tipo de herencia ligada al sexo que sigue la Hemofilia. Fuente: El blog de un profesor de genética, 2014.

Luego se procede a la exploración clínica para poder identificar rasgos dismórficos, características que no están dentro del rango normal de los individuos. La impresión clínica debe ser comprobada por una serie de medidas antropométricas, que posteriormente son valoradas según tablas de percentiles. Por último, se realizan exámenes complementarios, de acuerdo al tipo de enfermedad problema. Si la enfermedad problema es un trastorno cromosómico se hace un estudio del cariotipo. En cambio, si se trata de enfermedades monogénicas, multifactoriales o mitocondriales es necesario realizar un estudio molecular del ADN y exámenes auxiliares como análisis bioquímicos, estudios radiológicos o de neuroimagen. (28)

5.1. TECNICAS PARA LA IDENTIFICACIÓN DE GENES

Existen diversas técnicas de genotipificación que ayudan a realizar la amplificación y secuenciación *in vitro* del DNA. Estas usan como base la reacción en cadena de la polimerasa (PCR), la cual amplifica pequeños fragmentos de DNA. Existen 8 ensayos de genotipado comúnmente empleados:(29)

1. El ensayo TaqMan (Applied Biosystems)
2. Ampliación de masa homogénea (hME) e iPLEX (Sequenom)

3. SNPstream (Beckman Coulter)
4. SNPlex (Applied Biosystems)
5. Goldengate
6. Ensayo Infinium I y II (Illumina)
7. Ensayos de mapeo humano; GeneChip
8. Sondas de inversión molecular: MIP (Affymetrix)

Estos sistemas se pueden agrupar de acuerdo al número de Polimorfismo de un solo nucleótido(SNP) que pueden determinar por reacción y muestra. Existen entonces, sistemas aptos para identificar entre 1- 2 SNPs, 2-48 SNPs y 96 -500 000 SNPs.

Se define SNP como la variación de una base del DNA y pueden estar o no relacionados con una enfermedad, esto depende de la zona del genoma en la cual estén ubicados. Para poder estudiarlas se utilizan también biochips o microarrays de DNA o RNA. Estos son superficies sólidas con celdillas en las que se depositan nucleótidos de manera ordenada, permiten genotipar los marcadores genéticos tradicionales y permiten evaluar la expresión diferencial de miles de genes simultáneamente.(30)

La revolución en el estudio de los seres vivos y las interacciones entre éstos viene de la mano del desarrollo de las plataformas de secuenciación de alto rendimiento, permiten la secuenciación de genomas y transcriptomas o el genotipado masivo de individuos a un menor costo. La secuenciación masiva permite obtener marcadores genéticos, la determinación de patrones de expresión y regulación génica o la identificación de especies de microorganismos presentes en muestras ambientales.(30)

CAPITULO II

MATERIALES Y METODOS

1. LUGAR DE EJECUCION

Laboratorio de Genómica y Enfermedades Neurovasculares de la Universidad Católica de Santa María F - 401

2. MATERIALES

2.1. Muestras

Se utilizaron muestras de saliva

2.2. Materiales

- Tubo colector de saliva Oragene-DNA.
- Tubos eppendorf de 1.5 ml.
- Tubos microcentrífuga de 1.5 ml.
- Micropipetas de 20ul, 50 ul, 100 ul y 1000 ul.
- Tubos de criopreservación 2 ml.
- Parafilm.
- Rollo de papel aluminio.
- Rotulador negro indeleble.
- Papel filtro.
- Gradilla para tubos microcentrífuga.
- Punteras amarillas y azules.
- Caja para tubos eppendorf.

2.3. Insumos y Reactivos Químicos

Reactivo	Marca	Uso
Kit de DNA prepIT-L2P.	DNAGenotek	Extracción de DNA
Etanol 95 %	Sigma-Aldrich, Inc.	Presipitación del DNA
Etanol 75%	Sigma-Aldrich, Inc.	Deshidratación del DNA
Solución TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0)	Merck®	Solución para almacenar DNA
Hielo		Proceso de extracción de DNA
Taq polimerasa native	Merck®	Realizar la reacción en cadena de la polimerasa
AccuPrime Pfx Supermix	Merck®	Mezcla que brinda dNTPs y partículas de magnesio para favorecer el PCR.
TAE Buffer (Tris-acetato – EDTA) 50X	Merck®	
Primer forward y reverse TRAPPC 12	Merck®	Segmentos empleados para amplificar el fragmento deseado.
Gel de agarosa al 2 %	Bio-Rad Laboratories, Inc	Medio por el cual se separan los amplicones.
Tampón de carga (6X) (40% Sacarosa, 0.25% Azul de Bromofenol y, 0.1M EDTA)	Bio-Rad Laboratories, Inc	Permiten que la muestra se torne densa para quede en el fondo de la poza y marca la altura de la corrida.
Tampón TAE (Tris-acetato – EDTA) 50X	Bio-Rad Laboratories, Inc	Permite la conducción de la electricidad
Bromuro de Etidio (0.5ug/ml)	Merck®	Colorante que tiñe las fragmentos separados.
Marcador de pesos moleculares.	Bio-Rad Laboratories, Inc	Escala que permite identificar el peso del fragmento.

2.4. Equipos e Instrumentos

Equipo	Marca	Uso
Plataforma DNBseq™.	BIG	Secuenciación automatizada del DNA
Microcentrifuga.	Boeco	Extracción de DNA
Vortex.	Thermo Scientific	Extracción de DNA
Incubadora de agua.	Eurotech	Extracción de DNA
Refrigeradora a -20 °C.	Biobase	Almacenar las muestras de DNA.
Termociclador	Thermo Scientific	Realizar la reacción en cadena de la polimerasa
Cubeta de electroforesis.	Bio-Rad Laboratories, Inc	Electroforesis en gel de agarosa.
Fuente de poder para electroforesis	Bio-Rad Laboratories, Inc	Electroforesis en gel de agarosa.
Cámara reveladora UV.	Thermo Scientific	Identificar los fragmentos de DNA

2.5. Softwares

Software	Marca	Uso
BGISEQ-500 Software	BGI	Secuenciación automatizada del DNA
XMGRACE, Version 5.1.	XMGRACE	Dinámica molecular
UCSF Chimera	Chimera	Dinámica molecular
VMD	VMD	Dinámica molecular
I-TASSER	I-TASSER	Dinámica molecular
Gromacs v.2020.4	Gromacs	Dinámica molecular
Paquete estadístico y gráfico computacional R	R-project.org	Secuenciación de células individuales
SEURAT	R-project.org	Secuenciación de células individuales

3. METODOS

ETAPA 1:

3.1. SUJETOS

Para determinar las variaciones genéticas en este proyecto, se realizó el análisis completo del genoma a 9 muestras de saliva. Estas pertenecen a 9 personas, miembros de una familia relacionados con la EA y enfermedades cardiovasculares. Se formó 3 grupos de estudio:

- GRUPO 1: Conformado por 3 hermanos diagnosticados con EA.
- GRUPO 2: Conformado por 3 hijos de los pacientes diagnosticados con EA.
- GRUPO 3: Grupo control conformado por 1 hermana sin deterioro cognitivo, su hijo sano y la hija de un paciente diagnosticado con EA que no presenta signo alguno de deterioro cognitivo.

CRITERIOS DE INCLUSION

- Personas diagnosticadas con EA.
- Personas con antecedentes de enfermedades cardíacas.
- Personas que padecen desordenes mentales.
- Personas que formen parte de una familia.

CRITERIOS DE EXCLUSION

- Personas que no compartan lazos familiares.
- Personas sin antecedentes cardíacos o desordenes mentales.

4. IMPLEMENTACIÓN DEL BIOBANCO:

La Directiva para Muestras Biológicas en Ensayos Clínicos, Directiva N° INS/OGAT-V.01 (ANEXO1) en la legislación del Estado peruano establece que el biobanco debe ser apoyado por una institución, el cual será el Laboratorio de neurociencias y tecnología farmacéutica de la Universidad Católica de Santa María. Este es

responsable de su custodia, es decir, velar por la seguridad de las muestras almacenadas, salvaguardar los intereses de los donantes, el control del uso y disposición de muestras y resultados. También indica que este requiere tener una organización y reglamento interno escrito que incluya:

- a) Los responsables por la guarda y autorización del uso del material biológico. Un director científico y responsable del archivo de muestras.
- b) Mecanismos que garanticen la privacidad y el respeto a la confidencialidad. Aplicar una correcta codificación de las muestras.
- c) Mecanismos que aseguren la posibilidad de contacto con los donadores para suministro de información de sus intereses o para la obtención de consentimiento específico para uso de muestras en un proyecto de investigación.

También acogerse al asesoramiento de un Comité de Ética y un Comité Científico, que garantice el cumplimiento de los principios éticos aplicables a la investigación biomédica y asesore al responsable del biobanco sobre la dirección y objetivos científicos de este.

Por último, desarrollar un sistema de información y protocolo de seguridad que garantice que los datos personales de los sujetos fuente, los datos de salud y los derivados de los resultados de la investigación estén protegidos de acuerdo con la normativa vigente nacional e internacional, garantizando en todo momento la confidencialidad de los datos.

5. RECOLECCION DE MUESTRAS

Se recolectó muestras de saliva utilizando el tubo colector de saliva Oragene-DNA. Este kit ayuda a obtener muestras de ADN de alta calidad y cantidad, evita degradación del ADN y no es un método invasivo.

Para obtener una muestra de alta calidad de ADN se recomendó a los participantes que no fumen, coman o tomen liquido alguno unos 30 minutos antes de tomar la muestra.

El sujeto de estudio previamente informado de la metodología sobre la recolección de muestra, rezumó saliva en un tubo colector Oragene-DNA hasta la señalética de color

negro presente en el tubo, además al sujeto se le indicó que mantuviera el tubo en forma vertical con el fin de favorecer el sellado con la tapa tipo embudo y permitir la mezcla del reactivo con la muestra. Finalmente se cerró herméticamente con la tapa tipo rosca y agitó 5 segundos, para asegurar la mezcla completa de la saliva y el reactivo. En el ANEXO 2 se encuentre el inserto del tubo empleado.

Las muestras fueron transportadas hacia el Laboratorio de Genómica y Enfermedades Neurovasculares bajo condiciones óptimas de temperatura para evitar la degradación de la calidad de DNA en un cooler y protegerlas de la luz.

Para cumplir con el protocolo de Bioseguridad para la toma y manejo de muestras de origen biológico, se usó los implementos designados como bata blanca, guantes esteriles, barbijo y gorro descartable para garantizar la seguridad y disminuir el riesgo de contacto con la muestra.

5.1. CODIFICACION DEL TUBO DE MUESTRA

El tubo que contiene la muestra se rotuló con la codificación asignada para el proyecto, respetando orden y confidencialidad de datos del donante.

El código general para las muestras fue EA, el código para pacientes control C, los hijos de los participantes fueron enumerados e identificados con un punto.

5.2. ALMACENAMIENTO

Muestras: Fueron almacenadas en un refrigerador a una temperatura de -20°C , de manera ordenada y respetando el espacio asignado para evitar su contaminación.

Información: Se realizaron cuestionarios a los participantes para obtener mayor información, la cual se plasmó en el heredograma. Esta almacenada en un archivo Microsoft Word bajo contraseña.

Este proceso se realizó respetando los criterios éticos y la confidencialidad de la información brindada. El consentimiento informado se encontra en el ANEXO 3.

ETAPA 2:

Ya con la muestra codificada y almacenada correctamente, se purificó el DNA y se realizó su secuenciación.

6. PURIFICACION DEL DNA:

Para el proceso se empleó el kit DNA PrepIT-L2P. Este kit de simple manejo, asegura la recuperación máxima de DNA de alta calidad, a partir de los tubos colectores de saliva Oragene.

Como primer paso se homogenizó la muestra de saliva contenida en el tubo Oragene-DNA por inversión, saliva durante 5 segundos. Para liberar el DNA e inactivar las nucleasas, se colocó el tubo en Baño María durante una hora a 50 °C. Luego se colocó 500 ul de la muestra en un tubo eppendorf de 1.5 ml y adicionó 20 ul del reactivo PT-L2P, se agitó y mezcló la muestra empleando un vortex durante 5 segundos. Al término de la agitación se colocó el tubo en baño de hielo por 10 minutos para después centrifugar la muestra a temperatura ambiente a 15 000 rpm durante 5 minutos.

En el tubo se observó un pellet rodeado de sobrenadante. Se transfirió todo el sobrenadante a un tubo eppendorf nuevo y limpio con una pipeta y el pellet fue eliminado.

Por último, se agregó 600 ul de etanol al 95 % al sobrenadante y mezcló 10 veces. Se formó un precipitado compuesto por DNA, se descartó el sobrenadante y adicionó 250 ul de etanol al 70% y dejó reposar por un minuto a temperatura ambiente. Se decantó el etanol y agregó 100 ul de Solución TE, se agitó en vortex durante 5 segundos e incubó durante toda la noche a temperatura ambiente. En ANEXO 4 se encuentra el protocolo del kit empleado.

7. AMPLIFICACION DEL DNA

La muestra de DNA fue fragmentada al azar por Tecnología Covaris en fragmentos de 350bp. La reparación terminal de los fragmentos de DNA fue realizada y una base A se agregó al terminal 3' de cada hebra de DNA. Los adaptadores se ligaron en ambos terminales y al terminal reparado/dA y amplificados por ligación mediada PCR (LM-PCR), separadas por una sola hebra para ser, luego, cicladas.

La amplificación de círculos rodantes (RCA) fue realizada para producir DNA nano bolas (DNBs). Las DNBs de calidad fueron cargadas en nanoarrays y leídas por BGISEQ-500 Software, definidos como datos no analizados y guardados en formato FASTQ. Se empleó la plataforma DNBseq™. En el ANEXO 5 se describen las características de la plataforma empleada.

8. SECUENCIACION DEL DNA AUTOMATICA

El análisis bioinformático empezó con los datos secuenciados. Primero las lecturas limpias, fueron producidas por filtración para detectar y corregir errores en los datos brutos. El software Burrows-Wheeler Aligner (BWA), se usó para alinear los datos. Para asegurar la precisión en las medidas de las variaciones, seguimos Buenas Prácticas para análisis de variaciones con el kit de herramientas para análisis del genoma (GATK). Las lecturas duplicadas se removieron por Picard Tools. La profundidad de la secuenciación fue calculada en base a los alineamientos. Los SNPs en indels se detectaron por The State Of Heart Software como HaplotypeCaller de GATK. El número de copias de variantes (CNVs) fueron nombradas por el CBVnator y las variaciones estructurales (SV), fueron detectadas por Breakdancer. La herramienta SnpEff fue aplicada para realizar una serie de anotaciones de variantes.

9. VALIDACION DE LA MUTACION

Se identificó el gen y empleó la reacción en cadena de la polimerasa (PCR) para certificar la variación hallada.

Reacción en cadena de la polimerasa (PCR)

Previamente, en un tubo eppendorf se colocó AccuPrime Pfx Supermix, junto con los Primers Forward y Reverse del gen TRAPPC 12. Luego se agregó la Taq Polimerasa native y por último la muestra en estudio. Se mezcló y centrifugó la mezcla brevemente. El tubo eppendorf se introdujo en el termociclador bajo el programa PCR para completar el siguiente ciclo:

- a) Se calentó la Taq polimerasa 94°C durante 2 minutos, para separarla del anticuerpo que evita que actúe a temperaturas bajas.

- b) Se desnaturalizó la muestra de ADN llevando la reacción a 94°C durante 30 segundos.
- c) Alineación: los primers se unieron con sus zonas específicas mediante enlaces de Hidrógeno, este proceso se realizó a una temperatura de 55°C durante 30 segundos.
- d) Se aumentó la temperatura a 72°C. Durante 30 segundos, la Taq polimerasa tomó los dNTPs del medio para complementar la hebra molde. Se realizaron 30 ciclos.

Proceso de electroforesis

Se empleó la técnica de electroforesis en gel de agarosa al 2%, cuyo fundamento es la migración de moléculas de acuerdo a su carga eléctrica. Entonces, se conoce que el DNA está cargado negativamente, si lo colocamos en un campo eléctrico, este viajará del polo negativo al positivo.

Se agregó tampón de carga (6X) (40% Sacarosa, 0.25% Azul de Bromofenol y, 0.1M EDTA) a la muestra, para aumentar la densidad de la muestra y se quede en la poza del gel de agarosa. El colorante azul de bromofenol, indicó la altura de la corrida en el gel. Esta mezcla se agregó a las pozas formadas en el gel de agarosa. En una poza del gel se agregó el marcador de pesos moleculares, el cual es un patrón que permite identificar el tamaño de fragmento que se amplificó. El gel se sumergió completamente, en tampón TAE (Tris- acetato – EDTA) 50X, sus sales permitieron la transmisión de la corriente. Una vez que las muestras se cargaron se estableció la corriente en la cubeta de electroforesis. Completada la separación se utilizó el colorante Bromuro de Etidio para visualizar las bandas de DNA, el gel se sumergió en el colorante durante 15 minutos. Se iluminó el gel con luz ultravioleta (UV) para detectar los fragmentos separados. En el ANEXO 6 se encuentra el protocolo seleccionado.

Secuenciación de células individuales (Single-cell RNA sequencing)

Todos los análisis en esta etapa fueron realizados en R junto con su paquete SEURAT. Esta técnica se empleó para analizar la heterogeneidad de la célula ubicadas en el cerebro, ya que este órgano es el más afectado por la Enfermedad de Alzheimer.

R es un software gratis para estadística computacional y gráficas que puede ser descargado de la página oficial de proyecto R. (<https://www.r-project.org/>). SEURAT es un paquete de R diseñado para el control de calidad, integración de información, análisis, exploración e interpretación biológica de la data obtenida del scRNA-seq. Está disponible en CRAN (<https://cran.r-project.org/>). Para procesar la data disponible se realizó la configuración del objetivo de SEURAT. En primer lugar, se analizó el conjunto de datos de células del cerebro disponible, de personas sanas y de pacientes con la enfermedad de Alzheimer.

Luego creamos el objeto de SEURAT, el cual sirve como un espacio donde se procesó de la información y se realizó el análisis. Para este proceso se realizó un control de calidad y selección de células, Seurat nos permite explorar fácilmente parámetros de control de calidad y filtrar células basándose en el criterio del usuario. Los parámetros más usados son el número de genes únicos detectados en cada célula y el porcentaje de leídas que mapean el genoma mitocondrial. A continuación, se normalizaron los datos, luego de remover células no deseadas.

Como siguiente paso el programa identificó características altamente variables, mediante un cálculo del subconjunto de características que exhiben una variación alta entre células, genes que se expresan altamente en unas células y pobremente en otras. Al enfocarse en estos genes en los análisis “*downstream*”, como PCA ayuda a resaltar la señal biológica de un conjunto de datos. El programa realizó la alineación de los datos y una reducción linear dimensional, el programa provee muchas formas para visualizar las células y sus características. Para interpretar los resultados se determinó la dimensionalidad de estos y las células fueron agrupadas. Luego se corrió una reducción dimensional no lineal (UMAP/tSNE) y se encontró las características diferenciales expresadas. Como último paso, se asignó la identidad del tipo de células a los grupos, para este paso se emplearon marcadores comunes.

Dinámica molecular (DM)

En primer lugar, se estableció una configuración inicial del sistema que puede provenir de datos experimentales (cristalografía de rayos X, NMR), de modelos teóricos o de una combinación de ambos, siempre se debe considerar que la calidad de los resultados (output), depende de la información inicial o datos de entrada (input).

Una vez optimizada la estructura, se asignó velocidades iniciales para cada uno de los átomos. Esto se realizó de manera aleatoria a partir de una distribución de Maxwell-Boltzmann a una determinada temperatura. Finalmente, es preciso mantener fijas algunas de las condiciones de simulación: número de partículas (N), volumen (V), temperatura (T), presión (P) o energía total del sistema (E). Al combinarlas se obtienen simulaciones microcanónicas (NVE), isotérmico-isobáricas (NPT) y canónicas (NVT) siendo NPT y NVT las más empleadas en DM de proteínas y ácidos nucleicos.

Las DM se componen de dos etapas: la primera una fase de equilibrado y la segunda una fase de producción. Durante la primera fase se llevó el sistema a un estado de equilibrio a partir de una configuración inicial. En esta fase parámetros como la energía potencial, la temperatura y la densidad son monitoreadas hasta que se estabilizan. En la segunda fase comenzó la producción en la cual se permitió la evolución del sistema obteniendo una trayectoria que será analizada. Mientras mayor sea el tiempo de la simulación, se puede analizar más espacios conformacionales y por lo tanto los resultados serán más precisos.(31)

Herramientas de análisis en una dinámica molecular

Una de las propiedades fundamentales que se debe analizar, es la estabilidad de la proteína. Para este análisis se midió el desplazamiento cuadrático medio (RMSD), de todos los átomos pesados, la RMSD cuantifica la distancia entre cada residuo de una alineación de dos estructuras, además de hacer una comparación a lo largo de toda la trayectoria. Las vibraciones alrededor del equilibrio no son aleatorias, sino que dependen de la flexibilidad de la estructura local. La fluctuación cuadrática media (RMSF), de cada residuo es fácil de calcular a lo largo de la trayectoria, pero lo más importante es que se pueden convertir a factores de temperatura, datos presentes para cada átomo de un archivo PDB. Hay dos herramientas básicas de análisis útiles para analizar: el tamaño de la proteína, si es flexible o rígida definido por el radio de giro (R_g) y el número de enlaces de hidrógeno.(32,33)

Finalmente se visualizaron los resultados con el programa XmGrace que lee archivos en formato .xvg y para visualizar las estructuras obtenidas se utilizó el UCSF Chimera y el VMD.

CAPITULO III

RESULTADOS Y DISCUSIÓN

1. SUJETOS

En el estudio fueron incluidos 9 personas miembros de una familia cuyo cuadro clínico es característico. La familia es de la ciudad de Arequipa, está compuesta por 7 hermanos. Tres hermanos fallecieron debido a problemas cardiacos, 3 hermanas fueron diagnosticadas con EA y una hermana se encuentra sana, lúcida en tiempo y espacio.

Las 3 personas diagnosticadas con EA están dentro rango de edad entre 65-80 años y son de sexo femenino, el diagnóstico clínico se efectuó de acuerdo a protocolos establecidos.

Los participantes señalaron que el padre murió a la edad de 65 años debido a problemas cardiacos y la madre falleció a los 75 años debido a un infarto con una sospecha de demencia. La hermana mayor falleció por infarto sin antecedente alguno a la edad de 87 años y no tuvo hijos. La segunda hermana no presenta signo alguno de deterioro cognitivo, tuvo cuatro hijos todos sanos 2 viven en Arequipa y 2 viven en Estados Unidos. La tercera hermana padece de EA y tuvo 2 hijos, una hija padece una enfermedad mental y su hijo es saludable. La cuarta hermana nació con problemas cardiacos, para los cuales tomaba medicación y falleció de infarto a los 68 años, adoptó a un hijo. El quinto hijo era hombre y falleció debido a un paro cardiaco, tuvo 3 hijos, uno de los cuales padece hipotiroidismo. La sexta hermana vive en el norte, tiene 76 años y está diagnosticada con EA Tiene 4 hijos, dos de los cuales padecen esquizofrenia, uno falleció y el último esta sano. Y la séptima hermana vive en EEUU también padece EA y recibe tratamiento, tiene 3 hijos, sanos. Estos datos se emplearon para construir un heredograma, que se observa en la Figura 1.2.

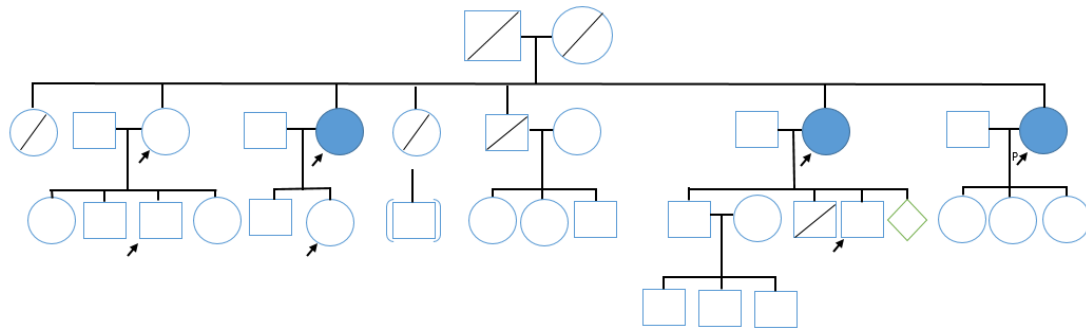



Figura 1.2: Heredograma del caso clínico. El gráfico permite organizar la información recabada tras realizar entrevistas con los miembros de la familia, se conoce que la familia está conformada por 7 hermanos. Cada figura geométrica representa una característica:

□ Sexo masculino, ○ Sexo femenino,   Muertos,   Enfermos con EA,   Participantes del estudio.

Ante un cuadro clínico único se procedió a tomar muestras de saliva a dos generaciones de la familia, madres e hijos. Tres hermanas diagnosticadas con EA forman parte del Grupo 1. Sus hijos forman parte del Grupo 2, en este se incluye una hija con leve retraso mental, el hijo que padece esquizofrenia y la hija de última hermana. Por último, formaron parte del Grupo 3 o control; la hermana sana, su hijo y la hija de la última hermana.

Una vez colectadas las muestras, se creó un código para proteger la información e identidad de los participantes, respetando las normas éticas que deben ser aplicadas en toda investigación. Cada tubo de muestra fue codificado y almacenado a -20°C en Laboratorio de Genómica y Enfermedades Neurovasculares de la Universidad Católica de Santa María.

Se trata de una familia amplia cuya primera generación fue afectada por la EA y enfermedades cardiovasculares, su segunda generación manifiesta enfermedades de tipo mental como el retardo y la esquizofrenia, señalando que existe un gen cuya manifestación ataca el Sistema Nervioso y puede expresarse en una población específica. También, existen familiares completamente sanos que, a pesar de la edad, se encuentran lúcidos como es el caso de una hermana y su hijo. Estas personas permiten la creación del grupo control, el cual permite comparar la presencia del gen cuya expresión provoca el cuadro clínico descrito.

Se empleó como sistema de codificación las letras EA y se enumeraron de acuerdo al orden de toma de muestra de saliva. Para identificar si la segunda generación se colocó un punto a lado del número y el grupo control se diferenció con una letra C. A continuación, se describe el tipo de código empleado para identificar las muestras de saliva durante el proyecto, ya que por cuestiones éticas los nombres de los participantes no pueden ser revelados. En la Figura 1.3 se observa el tubo de muestra junto con la codificación asignada.

1. EA 001
2. EA 001.1
3. EA 002
4. EA 002.2
5. EA 003
6. EA 003.1
7. EA C 001
8. EA C 001.1
9. EA C 002

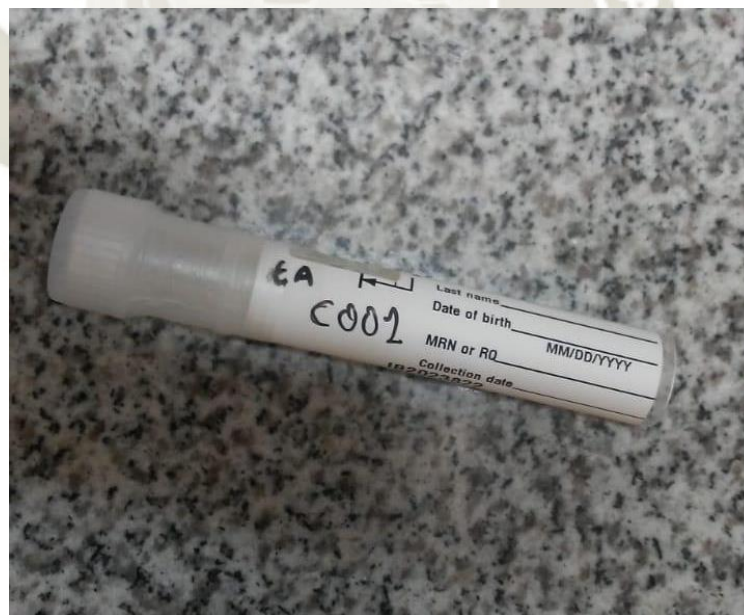


Figura 1.3 Tubo colector de saliva con código asignado.

Con las muestras de saliva se implementó el biobanco, para disponer de muestras asociadas a información y realizar futuros estudios genéticos, sobre enfermedades neurodegenerativas como la Enfermedad de Alzheimer. Para cumplir con la normativa dictada por el Estado Peruano, el laboratorio de Genómica y Enfermedades

Neurovasculares, facilitó sus instalaciones. El área esta delimitada para almacenar correctamente las muestras, el sistema de codificación para salvaguarda la identidad del participante. Cabe resaltar, que solo tiene acceso a la información obtenida los investigadores que forman parte del proyecto de investigación Biomarcadores contrato 024-2019- FONDECYT-BM-INC.INV. Este biobanco fue creado para poder ampliar el entendimiento sobre la EA y que futuros investigadores tengan a disposición muestras e información para realizar mayores estudios.

Este proyecto busca fomentar la investigación sobre enfermedades neurodegenerativas basándose en el perfil genético que manifiestan. Los biobancos facilitan la disponibilidad de muestras e información para poder investigar sobre causas, factores de riesgo y genéticos que predisponen a unos individuos a manifestar este tipo de enfermedades y así poder brindar un diagnóstico temprano y tratamiento adecuado, facilitando sus vidas.

2. EXTRACCIÓN Y SECUENCIAMIENTO DEL DNA

El objetivo principal de este proyecto fue identificar un gen que relacione la EA con una población específica. Para poder identificarlo se realizó la purificación y extracción del DNA de las muestras de saliva con el kit PrepIT-2LP, el cual permitió obtener DNA de alta calidad.

Para este secuenciar el DNA, las muestras fueron enviadas al laboratorio BGI en Hong Kong, China. Allí, los fragmentos de DNA fueron procesados por la plataforma DNBseq™ y la información resultante fue analizada por el software BGISEQ-500. Luego de diversas comparaciones con bases de datos genéticos, se identificó un gen cuya variación estaría relacionada con la EA. Tras la secuenciación de las muestras se obtuvo un promedio de 113,895.38 Mb de bases. Luego de remover los resultados de baja calidad se observó una media de 106,676.25 Mb de lecturas limpias. Las lecturas limpias de cada muestra tuvieron una alta Q20 y Q30, lo que señala una alta calidad para secuenciar. El promedio del contenido GC fue 41.44%. Toda la información obtenida tras la secuenciación está resumida en tabla 3.3.

Tabla 3.3: Resumen de información obtenida tras la secuenciación

Muestra	Leidas brutas	Base brutas (Mb)	Leidas limpias	Bases limpias (Mb)	Tasa de datos limpios (%)	Datos Limpios Q20 (%)	Datos Limpios Q30 (%)	Contenido GC (%)
EA 001.1	1,143,210,644	114,321.06	1,048,743,078	104,874.31	91.74	96.46	87.88	40.61
EA 001	1,113,650,932	111,365.09	1,055,147,476	105,514.75	94.75	97.59	91.94	41.30
EA 002.1	1,159,471,266	115,947.13	1,093,365,902	109,336.59	94.30	97.48	91.31	40.66
EA 002	1,086,897,602	108,689.76	1,020,779,680	102,077.97	93.92	97.53	91.56	43.23
EA 003	1,102,120,922	110,212.09	1,040,034,316	104,003.43	94.37	98.28	93.25	41.77
EA 003.1	1,217,842,106	121,784.21	1,151,093,592	115,109.36	94.52	98.15	92.69	41.70
EA C001.1	1,107,650,370	110,765.04	1,013,050,116	101,305.01	91.46	97.98	91.88	41.01
EA C001	1,128,269,024	112,826.90	1,071,759,618	107,175.96	94.99	97.40	91.35	41.70
EA C002	1,191,471,584	119,147.16	1,106,888,568	110,688.86	92.90	98.15	92.46	41.00
Promedio	1,138,953,827	113,895.38	1,066,762,482	106,676.25	93.66	97.67	91.59	41.44

Resumen estadístico de la alineación de la información.

Toda la información limpia se alinea según la referencia genómica humana (GRCh38/HG38), utilizando la herramienta Burrows-Wheeler Aligner (BWA). Se mapeo exitosamente el 88.10% de la información y 93.23 % de los datos son únicos. Se removieron los datos duplicados en el rango de 2.48% y 30.72 veces en la medida de la profundidad media de la secuenciación en el GWAS excluyendo regiones GAP. Un 99.34% de los datos secuenciados, excluyendo regiones GAP, fueron cubiertas por al menos 1x, 98.78% por 4X de cubrimiento y 97.42% por 10X de cubrimiento, estos resultados se encuentran resumidos en la Tabla 3.4. Adicionalmente, la distribución de profundidad de secuenciación por par de bases y la profundidad de secuenciación acumulada se observan en la Figura 3.3 y Figura 3.4

Tabla 3.4: Resumen de la alineación de la información

Muestra	Lecturas limpias	Bases Limpias (Mb)	Rango de mapeo (%)	Rango Único (%)	Rango duplicado (%)	Rango de discordancia (%)	Promedio de profundidad de secuen	Cubrimiento (%)	4X (%)	10X (%)	20X(%)

							ciación (X)				
EA 001.1	1,048,743,078	104,874.31	92.84	92.91	1.51	0.62	31.40	99.77	99.23	97.46	86.60
EA 001	1,055,147,476	105,514.75	96.93	94.30	2.15	0.42	33.92	99.13	98.62	97.47	92.08
EA 002.1	1,093,365,902	109,336.59	98.41	93.30	2.89	0.45	35.25	99.13	98.64	97.64	94.20
EA 002	1,020,779,680	102,077.97	64.38	92.86	1.97	0.41	20.81	99.05	98.18	95.34	54.26
EA 003	1,040,034,316	104,003.43	85.76	93.25	2.76	0.36	29.16	99.11	98.55	97.61	90.93
EA 003.1	1,151,093,592	115,109.36	89.71	92.76	3.12	0.38	33.62	99.13	98.67	97.79	94.80
EA C001.1	1,013,050,116	101,305.01	91.13	92.57	3.08	0.40	30.02	99.82	99.32	97.91	88.32
EA C001	1,071,759,618	107,175.96	81.62	93.62	2.61	0.45	28.75	99.08	98.45	97.23	87.44
EA C002	1,106,888,568	110,688.86	92.14	93.46	2.22	0.39	33.55	99.80	99.38	98.29	92.01
PROMEDIO	1,066,762,482	106,676.25	88,10	93.23	2.48	0.43	30.72	99.34	98.78	97.42	86.74

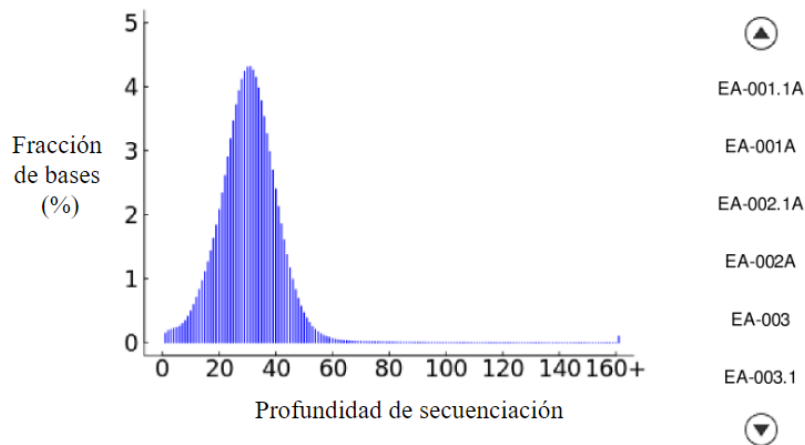


Figura 3.3: Distribución de la profundidad de secuenciación de los pares de bases del genoma entero. En el eje X se denota la profundidad de secuenciación y en el eje Y indica el porcentaje del genoma completo excluyendo la región GAP.

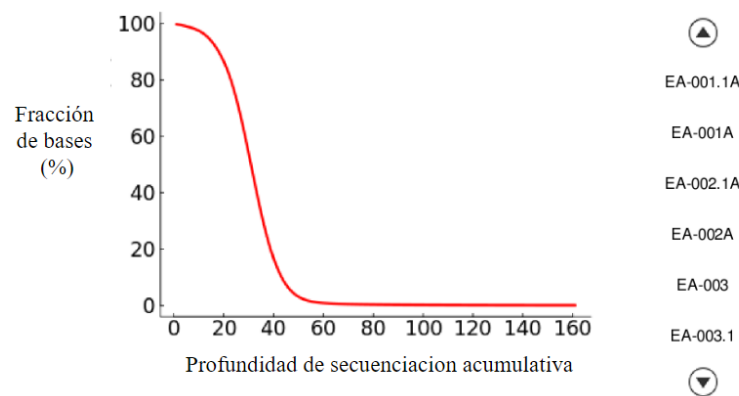


Figura 3. 4: Distribución acumulada del genoma completo. En el eje X se denota la profundidad de secuenciación y en el eje Y indica la fracción del genoma completo excluyendo regiones gap.

Control de calidad de los datos:

El estricto control de calidad fue realizado durante todo el proceso de análisis del gen. Se revisaron diversos controles los cuales se observan en la Tabla 3.5, donde Y muestra los datos que superaron el control de calidad y N muestra datos que no superaron el control de calidad.

Tabla 3.5: Resumen de resultados del control de calidad.

Muestra	Lectura limpia Q20(%)	Lectura limpia Q30 (%)	Porcentaje GC (%)	Rango de mapeo (%)	Rango duplicado (%)	Rango de discordancia (%)	Promedio de profundidad de secuenciación (X)	Cubrimiento (%)	4X (%)
EA 001.1	Y(96.46)	Y(87.88)	Y(40.61)	N(92.84)	Y(1.51)	Y(0.62)	Y(31.40)	Y(99.77)	Y(99.23)
EA 001	Y(97.59)	Y(91.31)	Y(40.66)	Y(98.41)	Y(2.89)	Y(0.45)	Y(35.25)	Y(99.13)	Y(98.64)
EA 002.1	Y(97.48)	Y(91.31)	Y(40.66)	Y(98.41)	Y(2.89)	Y(0.45)	Y(35.25)	Y(99.13)	Y(98.64)
EA 002	Y(97.53)	Y(91.56)	Y(43.23)	N(64.38)	Y(1.97)	Y(0.41)	Y(20.81)	Y(99.05)	Y(98.18)
EA 003	Y(98.28)	Y(93.25)	Y(41.77)	N(85.76)	Y(2.76)	Y(0.36)	Y(29.16)	Y(99.11)	Y(98.55)
EA 003.1	Y(98.15)	Y(92.69)	Y(41.70)	N(89.71)	Y(3.12)	Y(0.38)	Y(33.62)	Y(99.13)	Y(98.67)
EA C001.1	Y(97.98)	Y(91.88)	Y(41.01)	N(91.13)	Y(3.08)	Y(0.40)	Y(30.02)	Y(99.82)	Y(99.32)
EA C001	Y(97.40)	Y(91.35)	Y(41.70)	N(81.62)	Y(2.61)	Y(0.45)	Y(28.75)	Y(99.08)	Y(98.45)
EA C002	Y(98.15)	Y(92.46)	Y(41.00)	N(92.14)	Y(2.22)	Y(0.39)	Y(33.55)	Y(99.80)	Y(99.38)

Resultados de los polimorfismos de un solo nucleótido (SNPs)

Se identificó un promedio de 3,933,470 SNPs en todos los individuos. De estas variantes el 98.53% se encontraron en dbSNP y el 90.72% en 1000 Genomes Project database. Se analizaron 57,217 SNPs nuevas. La proporción entre transición y transversión (Ti/Tv) fue de 2.01. Entre todos SNPs, hay un promedio de 12,273 mutaciones sinónimas, 11,831 missense, 36 stop-loss, 115 stop-gain, 22 start-loss y 203 splice-site mutaciones en la región codificante. La proporción de Ti/Tv de SNPs codificante fue 2.84 (Tabla 3.7). El resumen estadístico de SNPs se observa en Tabla 3.6.

Tabla 3.6: Tabla resumen de resultados SNPs.

Muestra	Total de SNPs	Fracción de SNPse n dbSNP (%)	Fracción de SNPs en 1000genomas (%)	Nuevas SNPs	Homocigoto	Heterocigoto	Intrones	5' UTRs	3' UTRs	Upstream	Downstream
EA 001.1	3,901,265	98.05	90.73	75,535	1,469,010	2,432,255	1,530,528	4,590	24,877	57,039	55,389
EA 001	3,963,155	98.63	90.75	53,356	1,470,882	2,492,273	1,548,716	5,082	25,128	59,549	57,946
EA 002.1	3,944,514	98.61	90.70	54,169	1,491,727	2,452,787	1,542,351	4,973	24,914	59,042	56,628
EA 002	3,875,050	98.78	91.63	46,781	1,475,361	2,399,689	1,526,135	4,890	24,659	57,923	55,949
EA 003	3,953,166	98.62	90.78	53,884	1,479,894	2,473,272	1,546,203	5,046	24,739	59,120	57,122
EA 003.1	4,005,447	98.59	90.58	55,569	1,420,145	2,585,302	1,566,647	5,009	25,256	60,150	57,840
EA C001.1	3,958,952	98.52	90.41	57,840	1,487,061	2,471,891	1,554,806	4,944	25,320	59,846	57,147
EA C001	3,928,090	98.56	90.88	55,946	1,486,156	2,441,934	1,541,188	4,888	25,030	58,736	56,716
EA C002	3,871,598	98.38	90.01	61,875	1,579,405	2,292,193	1,511,027	4,918	24,841	58,135	56,295
PROMEDIO	3,933,470	98.53	90.72	57,217	1,484,404	2,449,066	1,540,844	4,926	24,973	58,837	56,781

Tabla 3.7: Categorías funcionales para codificar SNPs.

Muestra	Sinonima	Missense	Stopgain	Stoploss	Startloss	Splicing	Ti/Tv
EA 001.1	11,783	11,644	116	36	22	208	2.87
EA 001	12,269	11,965	120	35	24	202	2.88
EA 002.1	11,838	11,700	117	38	23	204	2.89
EA 002	11,992	11,603	110	34	26	199	2.86
EA 003	11,956	11,812	113	39	21	209	2.87
EA 003.1	12,057	11,905	128	38	25	205	2.91
EA C001.1	12,016	11,686	114	36	21	199	2.91
EA C001	14,740	12,432	108	37	22	209	2.49
EA C002	11,811	11,732	109	36	15	193	2.88
PROMEDIO	12,273	11,831	115	36	22	203	2.84

Resultados de InDels.

Hubo un promedio de 869,587 InDels en todas las muestras. El 85 % se describen en dbSNP y 54.29% en la base de datos 1000 Genomes Project. El número de nuevos InDels es de 130,022. Entre todos los InDels, hay un promedio de 504 frameshift, 0 stoploss, 1 startloss y 142 splice site (Tabla 3.8). El resumen estadístico de InDel se muestran en la Tabla 3.9. La longitud de la distribución de InDels en la región codificante se observa en Figura 3.5.

Tabla 3.8 Categorías funcionales para codificar InDels

Muestra	Frameshift	Insercion No-frameshift	Delección No-frameshift	Stoploss	Startloss	Splicing
EA 001.1	1,135	201	134	2	2	153
EA 001	348	103	135	0	3	145
EA 002.1	345	98	114	0	1	139
EA 002	508	132	129	0	1	135
EA 003	652	131	143	0	4	141
EA 003.1	369	104	130	0	2	141
EA C001.1	378	98	143	0	2	140
EA C001	455	102	142	0	1	147
EA C002	353	106	140	1	0	137
PROMEDIO	504	119	135	0	1	142

Tabla 3.9: Tabla resumen de InDels detectados.

Muestra	Total de InDels	Fracción de InDels en dbSNP (%)	Fracción de InDels en 1000genomes (%)	Nuevos InDel	Homocigoto	Heterocigoto	Intronico	5' UTRs	3' UTRs	Upstream	Downstream
EA 001.1	862,720	82.96	53.91	146,379	308,473	554,247	371,143	822	7,048	14,520	14,533
EA 001	859,615	85.97	55.22	119,966	315,775	543,840	368,833	872	6,778	14,982	14,809
EA 002.1	873,022	85.29	54.38	127,788	319,507	553,515	374,371	809	6,759	15,161	15,021
EA 002	803,055	87.48	57.15	99,944	320,439	482,616	345,233	806	6,501	13,937	13,854
EA 003	893,472	84.23	53.38	140,211	319,227	574,245	383,693	910	6,778	15,636	15,296
EA 003.1	907,485	84.28	53.05	141,920	302,585	604,900	389,857	878	6,897	15,881	15,564
EA C001.1	893,146	84.41	53.21	138,527	320,914	572,232	385,155	833	6,824	15,595	15,327
EA C001	851,257	86.19	55.31	116,929	319,340	531,917	365,665	807	6,785	14,675	14,602
EA C002	882,515	84.23	53.02	138,539	342,947	539,568	378,992	835	6,911	15,423	15,187
PROMEDIO	869,587	85.00	54.29	130,022	318,800	550,786	373,660	841	6,809	15,090	14,910

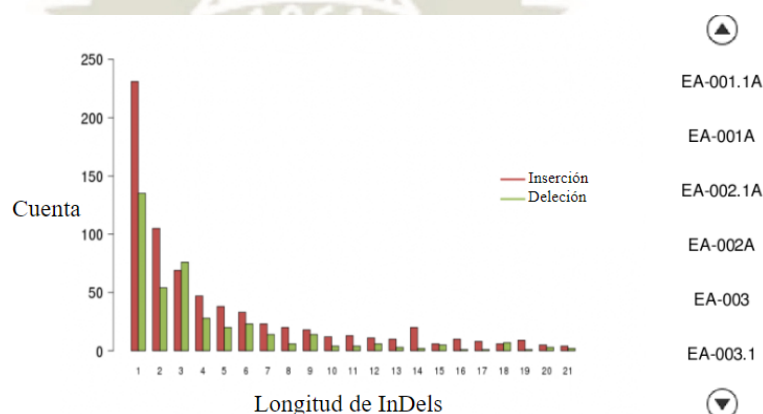


Figura 3.5: Distribución de longitudes de variantes codificantes de InDels.

Resultados del numero de copias de las variaciones (Copy Number Variations, CNVs)

Se identificaron 5,889 CNVs, con un total de 13,247,111 bp amplificadas y 75,654,955 bp eliminadas en todos los individuos. De estos, 1,774 se superponen a regiones codificantes del exón, 0 se superponen en area 5' sin traducción, 0 se superponen en área 3' sin traducción y 1,033 superponen en los intrones. Un número de 118 CNVs superponen 2-kb en el sitio de inicio de la región upstream de transcripción, y 87 superponen 2-kb en la región downstream del sitio terminal de transcripción. El resumen estadístico se muestra en la Tabla 3.10.

Tabla 3.10: Tabla resumen de los resultados de CNVs

Muestra	Total de CNVs	Exónicos	Splicing	Nc RNA	Intrónicos	5' UTRs	3' UTRs	Up stream	Downstream	Intergénico	Longitud de amplificación (bp)	Longitud de delección(bp)
EA 001.1	5,525	2,362	0	0	735	0	0	120	72	2,236	18,102,200	65,381,000
EA 001	12,831	2,423	0	0	2,705	0	0	187	187	7,332	13,680,500	104,051,8
EA 002.1	8,604	3,787	0	0	1,300	0	0	96	96	3,215	12,370,500	91,241,10
EA 002	4,142	789	0	0	867	0	0	82	82	2,326	12,360,600	43,303,500
EA 003	3,423	676	0	0	659	0	0	61	61	1,951	12,194,400	63,613,00
EA 003.1	3,642	786	0	0	705	0	0	66	66	2,015	13,394,600	73,531,30
EA C001.1	4,408	1,876	0	0	583	0	0	73	73	1,770	12,508,600	77,454,000
EA C001	6,619	2,234	0	0	1,127	0	0	91	91	3,035	13,192,900	87,164,80
EA C002	3,810	1,033	0	0	621	0	0	55	55	2,007	11,419,700	75,154,10
PROMEDIO	5,889	1,774	0	0	1,033	0	0	87	87	2,876	13,247,111	75,654,95

Con la información descrita en las tablas se realizaron comparaciones con diversas bases de datos, como 1000 genomes y dbSNPs para poder identificar la variación que se encuentre en menor proporción. El estudio reveló la presencia de una mutación tipo Missense en el gen TRAPPC 12 (NM_001321102.1:p.Leu697Val/c.2089C>G), según la notación en la posición 2089 de dio un cambio de la base Citosina por una Guanina, lo que a nivel de la proteína se traduce como el cambio del aminoácido leucina por valina.

La EA no tiene una causa definida, existen factores genéticos relacionados, como las mutaciones en la proteína precursora del beta amiloide y las presenilinas. Afecta a personas mayores de 65 pero se han observados casos en los cuales ataca a personas menores de 45 y esta peculiaridad se ha observado en personas que provienen de ciudades como Antioquia, Colombia o Jalisco, México.(34,35). Estas personas que presentan razgos únicos de la enfermedad, fomentaron realizar estudios genéticos para encontrar mutaciones o variaciones propias de una población, y es así como se descubrieron las mutaciones “PAISA” o “JALISCO”, las cuales afectan la PSEN 1.(34,35). Por esta razón es importante realizar estudios genéticos que permitan

encontrar variaciones en genes que proveen a la EA características únicas o propias de un grupo de personas así personalizar el diagnóstico y el tratamiento. En este proyecto se encontró el gen TRAPPC 12, cuya variación estaría relacionada con el desarrollo de la EA, este gen está relacionado con una enfermedad que provoca un retraso en el desarrollo de niño y microcefalia, y que brinda una característica especial a los miembros de la familia ya que padecen desordenes mentales como la esquizofrenia o enfermedades cardiovasculares.

3. GEN: COMPLEJO DEL TRÁFICO DE PROTEÍNAS 12 (*Trafficking Protein Particle Complex 12*, TRAPPC 12)

El transporte intracelular entre vesículas requiere de una serie de actividades moleculares, para asegurar el direccionamiento apropiado de la vesícula y su contenido. TRAPP es un complejo compuesto por diversas subunidades que participan en el transporte de partículas entre el Retículo Endoplasmático (RE) y el aparato Golgi.

Funciones:

- Transporte entre RE y el aparato Golgi.
- Proteólisis
- Organización del aparato Golgi.
- Oligomerización de proteínas
- Congresión de metafase.
- Regulación del ensamblaje del cinetocoro.
- Regulación positiva de la localización de proteína hacia el cinetocoro.

Localización:

El gen TRAPPC 12 se ubica en el cromosoma 2 en el brazo corto p25.3 (chr2:3,379,675-3,479,567), en la Figura 3.6 se señala con una línea roja la zona en la cual se encuentra este gen.

se brinda un esquema que representa la estructura del gen estudiado, su orientación es del extremo 5' UTR al 3' UTR, también se señala la posición en la cual se encuentra el cambio de la base y proteína.

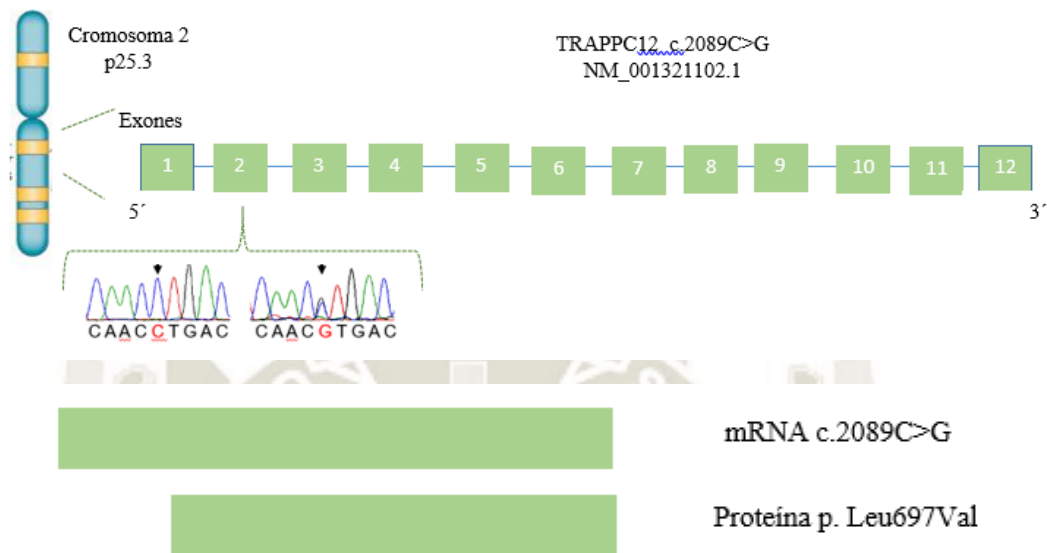


Figura 3.8: Representación de la estructura del gen. El gen está compuesto por exones e intrones, a nivel del exón se da una mutación tipo Missense.

RNA mensajero (mRNA): c.2089C>G

Está compuesto por 2483 bases, representa la zona codificante (exones). En el siguiente párrafo se observa la estructura del mRNA, el cual lleva orden para poder sintetizar la proteína en el ribosoma.

```

TTTCCGGCCTGGGGGCAGGGCCTCGGGTTCCTGCATCGGGCCTGAGCAG
AACTAGGCGCGCCGCGGCCCGGCACGCGCAGGCGTACAGAGCTCCCCGG
GGGTGCCAGTTCCTCTCCGATTCCCGGCTTCTGTACCTTCCTCACGGACC
TTGGTCACGGCCGCAGGGTCATGGAGGACGCTGGCGGGCGGCGAGGAGAC
CCCGGCCCGGAGGCCCGCACCCCCCTCAGCTCGCGCTCCGGAGGAG
CAGGGGTTGCTCTTCCAGGAGGAACCATCGATCTTGCCGGAGATGAGT
TTGGATCCGAAGAGAACGAGACCGCATCGGAAGGCTCGAGTCCTCTCGC
GGACAAGCTGAACGAACACATGATGGAGAGCGTCCTCATCTCTGACTCC
CCCAACAGCGAGGGCGACGCGGGCGACCTGGGGCCGAGTGCGGGACGAA
GCTGAGCCCCGGAGGGGAAGGCGACCCAGGCCCGGAGCCCCGCGGGCACC
CCGAGTCCCAGCGGCGAGGCCGACGGCGACTGTGCCCCCGAGGACGCGG
CACCCAGTAGCGGAGGGGCCCCGAGGCAGGACGCGGCCCGCGAGGTCCC
AGGCAGCGAAGCCGCGCGCCCGGAGCAGGAGCCTCCCGTTGCGGAGCCG
GTCCCGGTGTGCACCATCTTCAGCCAGCGCGCGCCCCAGCCTCCGGGGA
    
```

CGGCTTCGAGCCGCAGATGGTGAAGTCGCCAGCTTCGGTGGCGCCAGC
GAGGCTCGGCCAGGACACCGCCCCAGGTCGTGCAGCCCAGCCCCAGCC
TCAGCACGTTCTTCGGAGACACGGCCGCCAGCCACTCCTTGGCCTCGGAC
TTCTTCGACTCCTTTACTACCTCCGCCTTCATTTCCGTCAGCAATCCCGGC
GCGGGCTCCCCGGCCCCCGCCAGCCCGCTCCCCTCGCTGTGCCCGGGAC
CGAGGGGCGCCCCGAACCCGTGGCCATGCGAGGGGCCCCAGGCAGCTGCG
CCCCGGCGTGCAGAGCCTTTTCGCGCACATCCAGGCAGTGTTCGAGG
GAGTGACGACCCCTTTGCCACCGCCCTGAGCATGAGCGAGATGGACCGG
AGGAACGACGCCTGGCTTCCCGGCGAGGCTACGCGTGGAGTCTTGCGGG
CCGTGGCCACCCAGCAGCGCGGCCGTGTTTCGTGGACAAGGAGAACCT
CACCATGCCGGGCTCAGGTTTCGACAACATCCAGGGAGATGCAGTAAA
GACTTGATGCTTCGCTTTCTGGGTGAAAAAGCTGCAGCAAAGAGACAAG
TCCTAAATGCCGACTCAGTGGAACAATCTTTTGTGGATTGAAACAGCTA
ATCAGCTGCAGAACTGGAGGGCAGCAGTGGACCTGTGCGGACGTCTCC
TCACAGCCCACGGCCAGGGCTACGGCAAGAGCGGGCTGCTCACCAGCCA
CACGACAGATTCAGTGCAGCTCTGGTTTGTGAGGCTGGCACTACTAGTGA
AGTTGGGCCTTTTCCAGAATGCTGAGATGGAATTTGAACCCTTCGGAAAT
CTTGATCAGCCAGATCTTTATTACGAGTACTACCCGCACGTGTACCCTGG
GCGCAGGGGCTCCATGGTCCCCTTCTCGATGCGCATCTTGCACGCGGAGC
TTCAGCAGTACCTGGGGAACCCACAGGAGTCGCTGGATAGACTGCACAA
GGTGAAGACTGTCTGCAGCAAGATCCTGGCCAATTTGGAGCAAGGCTTA
GCAGAAGACGGCGGCATGAGCAGCGTACTCAGGAGGGCAGACAAGCC
TCTATCCGGCTGTGGAGGTCACGTCTGGGCGGGTGATGTACTCCATGGC
AACTGTCTGCTCCTGATGAAGGATTATGTGCTGGCCGTGGAGGCGTATC
ATTCGGTTATCAAGTATTACCCAGAGCAAGAGCCCCAGCTGCTCAGCGG
CATCGGCCGGATTTCCCTGCAGATTGGAGACATAAAAACAGCTGAAAAG
TATTTTCAAGACGTTGAGAAAGTAACACAGAAATTAGATGGACTACAGG
GTAAAATCATGGTTTTGATGAACAGCGCTTCCTTACCTCGGGCAGAAT
AACTTTGCAGAAGCCCACAGGTTCTTACAGAGATCTTAAGGATGGATCC
AAGAAACGCAGTGGCCAACAACAACGCTGCCGTGTGTCTGCTCTACCTG
GGCAAGCTCAAGGACTCCCTGCGGCAGCTGGAGGCCATGGTCCAGCAGG
ACCCAGGCACTACCTGCACGAGAGCGTGTCTTCAACCTGACCACCATG
TACGAGCTGGAGTCCTCACGGAGCATGCAGAAGAAACAGGCCCTGCTGG
AGGCTGTGCCCGCAAGGAGGGGGACAGCTTCAACACACAGTGCCTCAA
GCTGGCCTAGCTGCCTCCAACACACTACGTCAGAAGGACCCGGGTCTTTG
AACTGTGTCTTGAAGCTAATGTATTAATGTGACATGGAGGAACTCAATA
AACTCCTGCTTCACT

Proteína: Leu697Val

Se formó a partir de mRNA, está conformada por 735 aminoácidos. En el párrafo siguiente observamos los aminoácidos que conforman la proteína. En la mutación se observa que el aminoácido Leucina en la posición 697 cambio por una Valina, debido a la variación de la base Citosina por Guanina.

MEDAGGGEETPAPEAPHPQLAPPEEQGLLFQEETIDLGGDEFSGSEENETASE
GSSPLADKLNEHMMESVLISDSPNSEGDAGDLGRVRDEAEPGGEGDPGPEP
AGTPSPSGEADGDCAPEDAAPSSGGAPRQDAAREVPGSEAAARPEQEPPVAEP

VPVCTIFSQRAPPASGDGFEPQMVKSPSFGGASEASARTPPQVVQPSPSLSTFF
GDTAASHSLASDFFDSTTSAFISVSNPGAGSPAPASPPPLAVPGTEGRPEPVA
MRGPQAAAPPASPEPFAHIQAVFAGSDDPFATALSMSEMDRRNDAWLPGEA
TRGVLRAVATQQRGAVFVDKENLTMPGLRFDNIQGDVAVKDLMLRFLGEKA
AAKRQVLNADSVEQSFVGLKQLISCRNWRAAVDLCGRLLTAHGQGYGKSG
LLSHTTDSLQLWFVRLALLVKLGLFQNAEMEFEPFGNLDQPDLYEYYPH
VYPGRRGSMVPPFSMRILHAELQQYLGNPQESLDRLHKVKTVC SKILANLEQ
GLAEDGGMSSVTQEGRQASIRLWRSRLGRVMYSMANCLLLMKDYVLAVE
AYHSVIKYYPEQEPQLLSGIGRISLQIGDIKTAEKYFQDVEKVTQKLDGLQK
IMVLMNSAFLHLGQNNFAEAHRFFTEILRMDPRNAVANNNAVCLLYLGKL
KDSLRLQLEAMVQQDPRHYLHESVLFNL(V)TTMYELESSRSMQKKQALLEA
VAGKEGDSFNTQCLKLA

4. VALIDACIÓN DE LA MUTACIÓN

Como método de validación de la mutación hallada se empleó la técnica del PCR, para amplificar el fragmento en el cual se sospecha se dio la variación. Para este proceso se empleó la Taq Polimerasa, el PCR mix y se diseñaron Primers del gen TRAPPC 12 cumpliendo siguiente esquema:

DISEÑO DE PRIMERS

Como primer paso se diseñaron 2 primers: Forward(F) y Reverse(R). Aquí se utilizaron páginas web de libre acceso: GenScrip y Reverse complement. Se obtuvo la secuencia de los exones gracias la página web Gene Browser. El párrafo siguiente muestra las bases que conforman los exones del gen TRAPPC 12. Los resultados nos indican que la base Citosina cambia por una Guanina en la posición número 2089 del gen.

El gen TRAPPC12 y su mRNA c.2089C>G

ATGGAGGACGCTGGCGGCGGCGAGGAGACCCCGGCCCGGAGGCCCGG
CACCCCTCAGCTCGCGCCTCCGGAGGAGCAGGGGTTGCTCTTCCAGGA
GGAAACCATCGATCTTGGCGGAGATGAGTTTGGATCCGAAGAGAACGAG
ACCGCATCGGAAGGCTCGAGTCTCTCGCGGACAAGCTGAACGAACACA
TGATGGAGAGCGTCCTCATCTCTGACTCCCCAACAGCGAGGGGCGACGC
GGGCGACCTGGGCCGAGTGCGGGACGAAGCTGAGCCCGGAGGGGAAGG
CGACCCAGGCCCGGAGCCCGCGGGCACCCCGAGTCCCAGCGGCGAGGCC
GACGGCGACTGTGCCCCGAGGACGCGGCACCCAGTAGCGGAGGGGCC
CGAGGCAGGACGCGGCCCGGAGGTCCCAGGCAGCGAAGCCGCGCGCC
CGGAGCAGGAGCCTCCCGTTGCGGAGCCGGTCCCGGTGTGCACCATCTTC
AGCCAGCGCGCCCCAGCCTCCGGGGACGGCTTCGAGCCGCAGATGG
TGAAGTCGCCAGCTTCGGTGGCGCCAGCGAGGCCTCGGCCAGGACACC
GCCCCAGGTCGTGCAGCCCAGCCCCAGCCTCAGCACGTTCTTCGGAGAC

ACGGCCGCCAGCCACTCCTTGGCCTCGGACTTCTTCGACTCCTTTACTAC
CTCCGCCTTCATTTCCGTCAGCAATCCCGGCGCGGGCTCCCCGGCCCCCG
CCAGCCCGCCTCCCCTCGCTGTGCCCGGGACCGAGGGGCGCCCCGAACC
CGTGGCCATGCGAGGGCCCCAGGCAGCTGCGCCCCCGGCGTCGCCAGAG
CCTTTCGCGCACATCCAGGCAGTGTTCGAGGGAGTGACGACCCCTTTC
CACCGCCCTGAGCATGAGCGAGATGGACCGGAGGAACGACGCCTGGCTT
CCCGGCGAGGCTACGCGTGGAGTCTTGCGGGCCGTGGCCACCAGCAGC
GCGGCGCCGTGTTTCGTGGACAAGGAGAACCTCACCATGCCGGGCCTCAG
GTTTCGACAACATCCAGGGAGATGCAGTTAAAGACTTGATGCTTCGCTTTC
TGGGTGAAAAAGCTGCAGCAAAGAGACAAGTCCTAAATGCCGACTCAGT
GGAACAATCTTTTGTGGATTGAAACAGCTAATCAGCTGCAGAACTGG
AGGGCAGCAGTGGACCTGTGCGGACGTCTCCTCACAGCCCACGGCCAGG
GCTACGGCAAGAGCGGGCTGCTCACCAGCCACACGACAGATTCACTGCA
GCTCTGGTTTGTGTCAGGCTGGCACTACTAGTGAAGTTGGGCCTTTTCCAGA
ATGCTGAGATGGAATTTGAACCCTTCGGAAATCTTGATCAGCCAGATCTT
TATTACGAGTACTACCCGCACGTGTACCCTGGGCGCAGGGGCTCCATGGT
CCCCTTCTCGATGCGCATCTTGCACGCGGAGCTTCAGCAGTACCTGGGGA
ACCCACAGGAGTCGCTGGATAGACTGCACAAGGTGAAGACTGTCTGCAG
CAAGATCCTGGCCAATTTGGAGCAAGGCTTAGCAGAAGACGGCCGGCATG
AGCAGCGTGACTCAGGAGGGCAGACAAGCCTCTATCCGGCTGTGGAGGT
CACGTCTGGGCCGGGTGATGTACTCCATGGCAAACCTGTCTGCTCCTGATG
AAGGATTATGTGCTGGCCGTGGAGGCGTATCATTTCGGTTATCAAGTATTA
CCCAGAGCAAGAGCCCCAGCTGCTCAGCGGCATCGGCCGGATTTCCTG
CAGATTGGAGACATAAAAACAGCTGAAAAGTATTTTCAAGACGTTGAGA
AAGTAACACAGAAATTAGATGGACTACAGGGTAAAATCATGGTTTTGAT
GAACAGCGCGTTCCTTCACCTCGGGCAGAATAACTTTGCAGAAGCCAC
AGGTTCTTCACAGAGATCTTAAGGATGGATCCAAGAAACGCAGTGGCCA
ACAACAACGCTGCCGTGTGTCTGCTCTACCTGGGCAAGCTCAAGGACTCC
CTGCGGCAGCTGGAGGCCATGGTCCAGCAGGACCCAGGCACTACCTGC
ACGAGAGCGTGCTCTTCAAC(C/G)TGACCACCATGTACGAGCTGGAGTCC
TCACGGAGCATGCAGAAGAAACAGGCCCTGCTGGAGGCTGTCGCCGGCA
AGGAGGGGGACAGCTTCAACACACAGTGCCTCAAGCTGGCCTAG

En el siguiente texto se observa la porción del DNA compuesta por exones (representados por letras mayúsculas) e intrones (letras minúsculas)

tctctctccagtcaccagctccaggcagtggtgtggcccttaggggaagtgaccaacaggggccacctaggtctg
cccagcagctctgaccacccccaggcccctaacacggcccagcagcagcctggaatggcgggcgctgtctctgg
ttgtgtgggccaaccctggcggtgtgctctcccctagGCCAACAAACGCTGCCGTGTGTCT
GCTCTACCTGGGCAAGCTCAAGGACTCCCTGCGGCAGCTGGAGGCCATG
GTCCAGCAGGACCCAGGCACTACCTGCACGAGAGCGTGCTCTTCAACC
TGACCACCATGTACGAGCTGGAGTCTCACGGAGCATGCAGAAGAAACA
GGCCCTGCTGGAGGCTGTCGCCGGCAAGGAGGGGGACAGCTTCAACAC
ACAGTGCCTCAAGCTGGCCTAGCTGCCTCCAACACACTACGTCAGAAGG
ACCCGGGTCTTTGAAACTGTGTCTTGAAGCTAATGTATTAATGTGACATG
GAGGAACTCAATAAACTCCTGCTTCACTggtgtctgctgcgtgtcttctgtgccaagccac
ggcccagcccagcactccccgagttggtcgcggttcagccacgcagtcctgcagctgggtcactgttca

El primer tiene una longitud de 231 bp, esta compuesto por:

- TRAPPC12 Leu697Val_Forward tagGCCAACAACAACGCTG
- TRAPPC12 Leu697Val_Reverse GCACTGTGTGTTGAAGCTGT

Luego de realizar el proceso de amplificación por PCR, los fragmentos se encuentran juntos en la muestra, para el fin del estudio es importante identificar aquel fragmento que presenta la variación. Para separar los fragmentos e identificar aquel que esta mutado empleamos la técnica de electroforesis en gel de agarosa al 2% y tras el revelado con luz UV se observó que el fragmento en estudio se encuentra a la altura de 300 bp. El resultado de la revelación se observa en la Figura 3.9 y Figura 3.10.

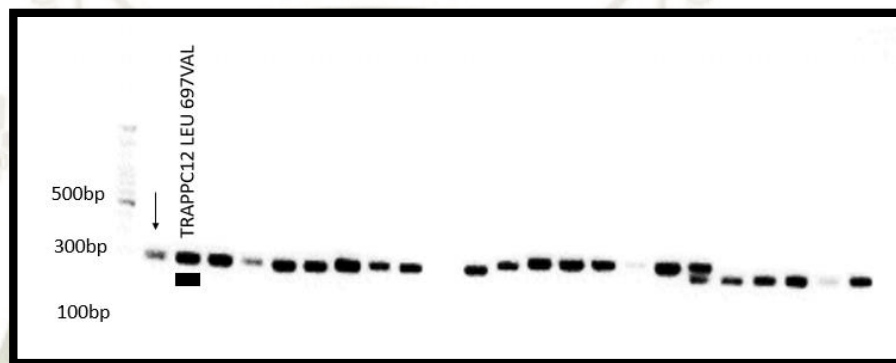


Figura 3.9: Impresión del gel de agarosa al 2% luego de la corrida electroforética. En la primera poza se colocó la escala de pesos moleculares, a continuación, se colocaron las muestras. Luego de la corrida se observó la presencia del gen mutado TRAPPC 12 LEU697VAL a la altura de 300bp.

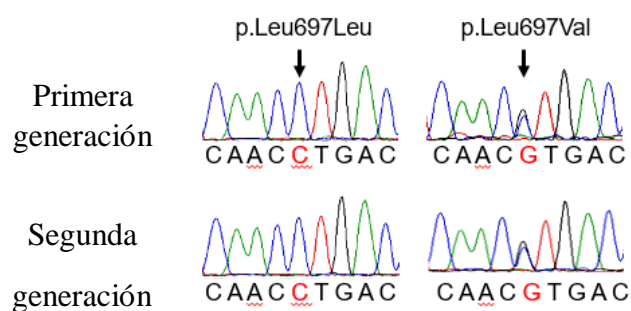


Figura 3.10: Fragmento amplificado de la mutación. Se compara la secuencia de bases presente en la muestra de esputo de la madre (primera generación) e hijo (segunda generación). En la prima columna se observa la base nitrogenada Citosina y en la segunda columna, en el lugar de la Citosina se encuentra una Guanina en la posición número 2089 del genoma.

Este tipo de mutación se encuentra en baja proporción en la base de datos pública de variaciones genéticas (dnSNP), la cual incluye información de más de 1000 muestras de todas las poblaciones. Todos estos resultados son compatibles con la hipótesis de que la mutación en esta familia es rara y esta relacionada con el desarrollo de la enfermedad de Alzheimer. Existe un reporte sobre de una familia peruana con EA temprana asociada con una mutación de tipo Missense de la PSEN 1, esta es la cuarta familia reportada a nivel mundial con esta variación(6), se entiende que este el primer y único estudio realizado en Perú y el resultado es importante ya que señala en nuestro país también se encuentran variantes genéticas únicas, responsables de la EA. La familia que forma parte de este proyecto expresa un cuadro clínico particular que podría ser reportado como una mutación propia en la población de Arequipa.

Secuenciación de células individuales (scRNA-seq)

Este tipo de secuenciación incluye un conjunto de tecnologías que permiten identificar la expresión del gen en una célula individual. Aunque todas las células en nuestro cuerpo comparten genotipos casi idénticos, se conoce que los diferentes tipos de células de nuestro cuerpo expresan un único transcriptoma, lo cual explicaría la variabilidad de respuesta de cada célula.(36)

Para identificar la expresión del gen TRAPPC 12 en el órgano del cerebro se utilizó una tecnología llamada scRNA-seq, esta permite estudiar el material genético propio de cada célula, lo que permite individualizar y diferenciarlas entre sanas y dañadas o enfermas. Mide el nivel de distribución de cada gen a través de una población de células y permite estudiar nuevas preguntas biológicas como la identificación del tipo de célula o la heterogeneidad en la respuesta celular. En la actualidad se cuenta con softwares de libre acceso que permiten emplear bibliotecas de scRNA-seq y realizar el análisis con la información disponible. Se trabajó con R y su paquete SEURAT, está diseñado para realizar el control de calidad, análisis y exploración en la información de scRNA-seq.

Se programó el objetivo de Seurat para identificar las células que forman parte del cerebro sano y de un cerebro con la EA. Se identificó gracias a marcadores comunes del programa la expresión del gen TRAPPC 12 en este tipo de células. Se

formó un grupo conformado por células del cerebro sano y otro por células del cerebro con la EA, para diferenciarlos se asignó un color a cada grupo; el color celeste para células sanas y el color naranja para células con EA. Al analizar las dos figuras se observó que la expresión del gen TRAPPC 12 en células de un cerebro sano es baja y en células con EA la expresión de este gen se multiplica. Lo cual indicaría una posible relación entre el gen TRAPPC 12 y la EA, porque en un cerebro sano este gen casi no se manifiesta, pero en un cerebro con EA se observa su alta prevalencia. (36)

Además, este gen está relacionado con encefalopatías que no permiten el desarrollo correcto de niños, pérdida de la audición y microcefalia, lo que señala su relación directa con el sistema nervioso. (36) El Figura 3.11 se observa los resultados luego del agrupamiento celular y análisis de la expresión del gen.

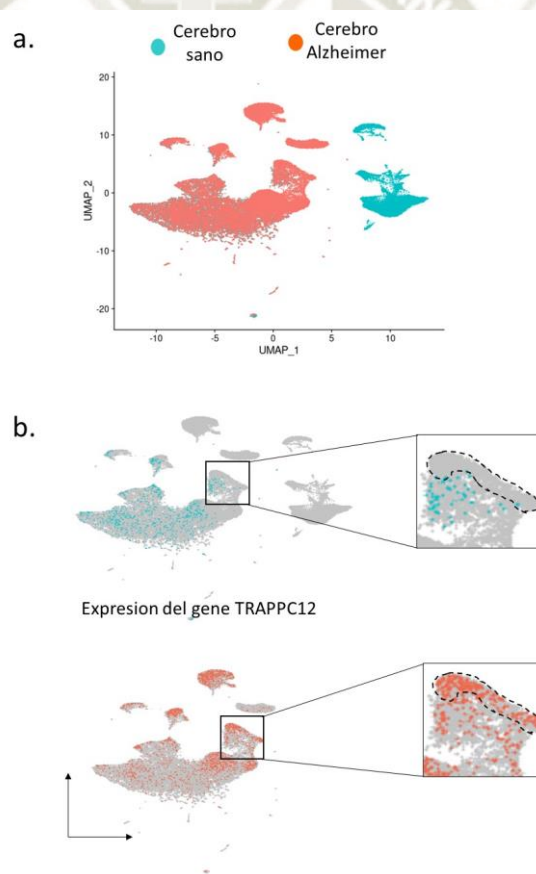


Figura 3.11: Gráfico obtenido a partir del paquete estadístico de R, SEURAT. a. Luego de realizar el análisis, las células del cerebro fueron divididas en 2 grupos: células sanas (color celeste) y células enfermas con EA (color naranja). B. Comparación entre la expresión de gen TRAPPC 12 en células de cerebro sano y células del cerebro enfermo. La manifestación del gen TRAPPC 12 en células del cerebro sano es pobre,

pero al comparar se observa que el gen TRAPPC 12 tiene una alta manifestación en células de cerebro enfermo.

La mutación de tipo Missense hallada en el gen TRAPPC 12 es muy rara. Ya que, al realizar la comparación con diversas bases de datos genéticos, se encuentra en muy baja proporción a nivel mundial. Para validar este resultado e indicar que se trataría de una mutación propia de una población específica se emplearon técnicas como PCR y electroforesis, las cuales señalaron la presencia de la mutación en el fragmento del gen. Entonces, esta podría ser señalada como una mutación propia de una población específica, ya que según el análisis scRNA-sequencing se observa su amplia manifestación en células de cerebros que manifiestan la EA.

DINAMICA MOLECULAR

En el campo de la simulación computacional, una de las técnicas de amplio uso en el estudio de sistemas proteicos es la Dinámica Molecular (DM), ya que permite analizar el comportamiento o evolución de un sistema (físico, químico o biológico) a través de tiempo, calculando las fuerzas entre los átomos que lo conforman, mediante la resolución de las ecuaciones de movimiento de Newton. Por medio de DM, se puede calcular diferentes propiedades fisicoquímicas de un sistema como la energía libre (G o H), entropía, solubilidad, viscosidad, presión, y en sistemas biológicos, permite hacer predicciones sobre propiedades como la dinámica de enlace o plegado de una proteína que luego se puede confirmar en el laboratorio. También permite medir la fuerza de interacción entre posibles fármacos y sus dianas biomoleculares o receptores, en general este tipo de estudios han demostrado buena correlación con resultados experimentales.(37)

Estructuras Iniciales

En el presente trabajo, se analizaron mediante simulaciones de dinámica molecular a la proteína TRAPPC12 nativa y su mutación L697V (Figura 3.12a). Las estructuras fueron obtenidas a partir de su secuencia lineal y modeladas usando el servidor I-TASSER. (38)

Para realizar la mutación en la posición 697, se utilizó el programa UCSF Chimera permitiendo la relajación de los residuos vecinos, mediante una minimización local con el fin de impedir choques estéricos, debidos al cambio en volumen del residuo mutado. El cambio estructural debido a la mutación puede observarse en el alineamiento de ambas estructuras (Figura 3.12b), obteniéndose una diferencia en el RMSD de 0.393 Å.

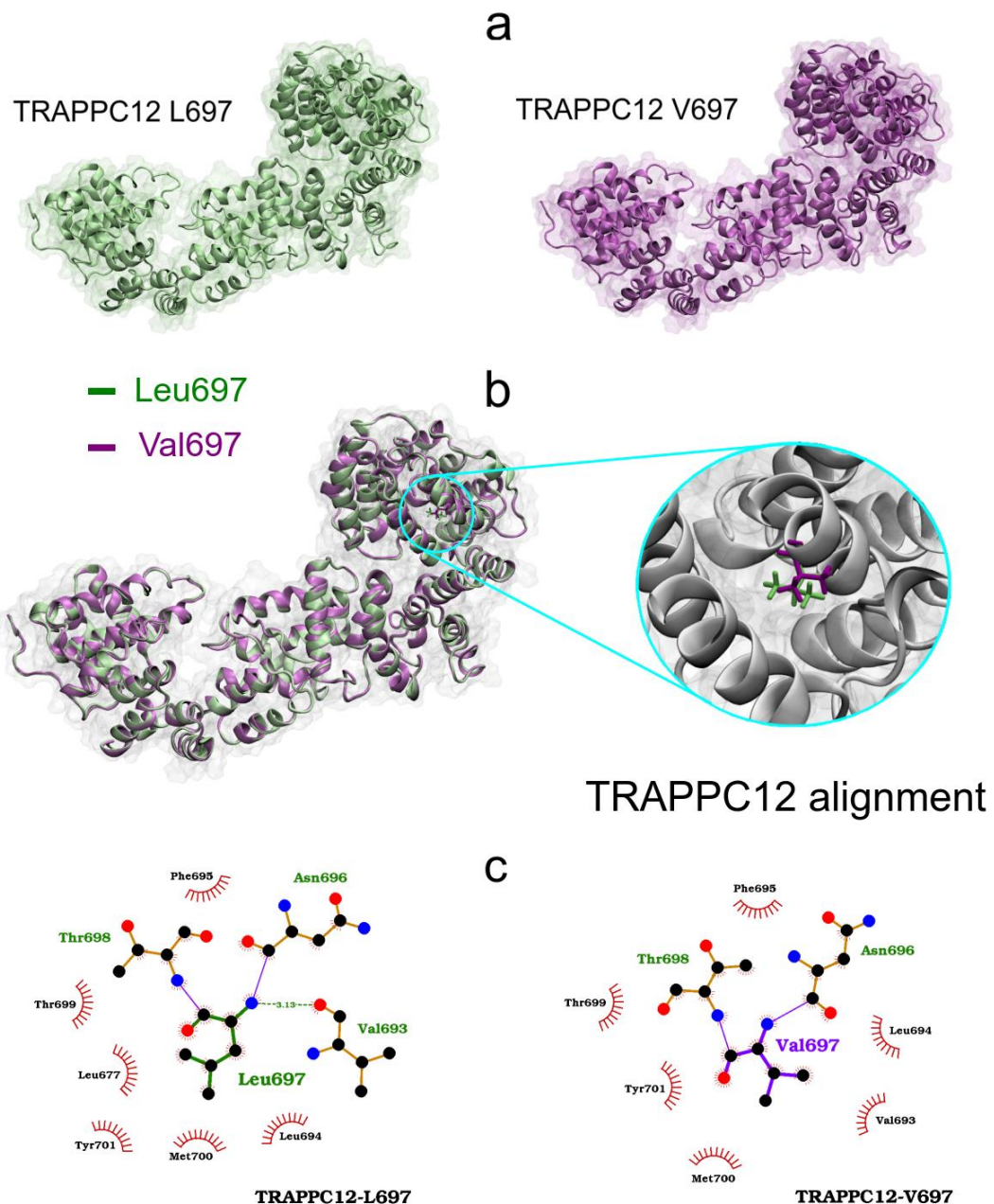


Figura 3.12: Estructuras iniciales de las proteínas TRAPCC12, P697 y V697. a) En ambas estructuras se realizó una minimización local en la zona donde se localiza el residuo 697. b) Alineamiento estructural de las proteínas donde se observa el cambio

sufrido por la estructura debido a la mutación L697V. c) Residuos vecinos a la posición 697 y que estén en contacto con éste. Las interacciones hidrofóbicas se muestran en semicírculos en rojo y las interacciones de carácter electrostáticas, como líneas discontinuas en color verde. Los residuos Asn696 y Thr698 se muestran unidos mediante líneas continuas de color púrpura.

La Figura 3.12c muestra las interacciones de los residuos 697, pudiéndose observar que, en el caso de la Leu697, ésta tiene una interacción de carácter electrostático con la Val693, el cual permite una mayor estabilización estructural, además de una interacción extra de carácter hidrofóbico con la Leu677. Ambas estructuras fueron optimizadas antes de las simulaciones DM haciendo uso del servidor ModRefiner.(39)

La metodología usada para obtener estructuras estables está basada en el tutorial realizado por Lemkul (40), y que comprende tres etapas: minimización, equilibrio y producción. En las tres etapas, se utilizó la suite de Gromacs v.2020.4.(41) y el campo de fuerza OPLS-AA. Para las etapas de minimización y equilibrio se usaron restricciones en las posiciones de los átomos pesados y sirvieron para estabilizar la estructura proteica en la caja de simulación bajo la presencia del solvente. Para la etapa de producción, el tiempo de simulación fueron 100 ns usando el ensamble NPT a 309.65K, 1 bar de presión y pH7 y como solvente se usó el modelo de agua TIP4P a una concentración salina de 0.154M. Los resultados pueden verse en las Figuras 3.13 y 3.14.

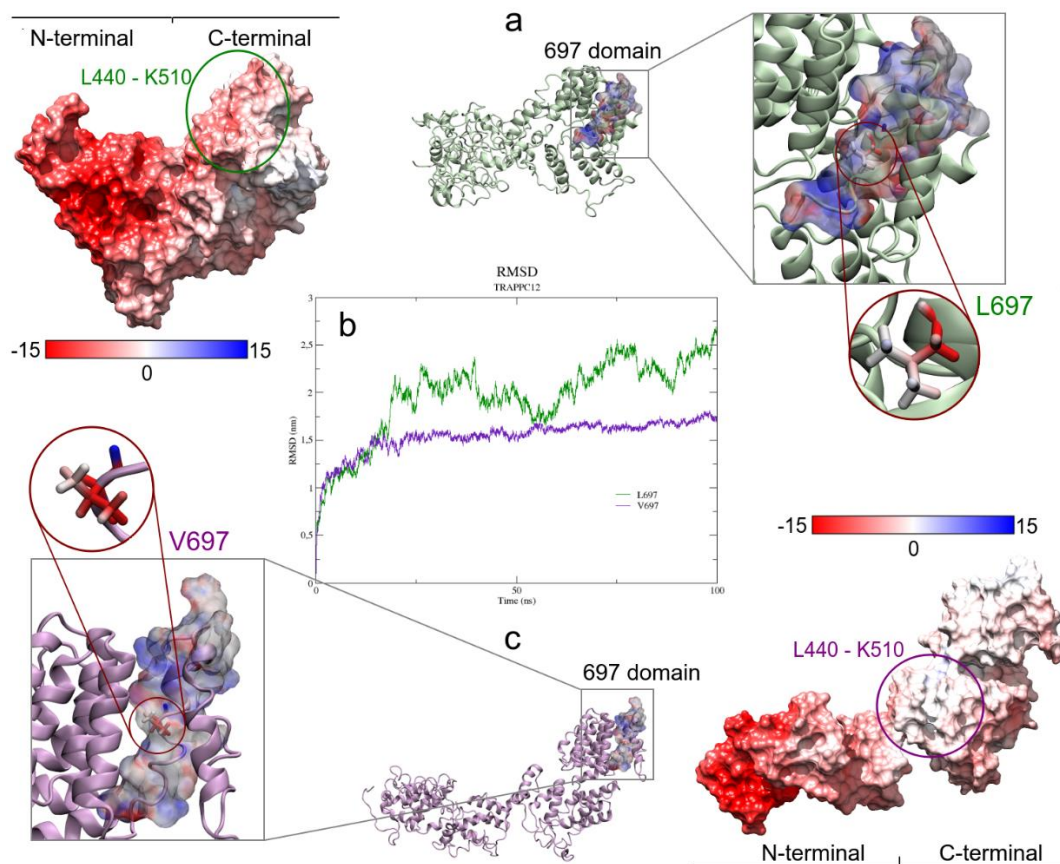


Figura 3. 13: Resultado de las simulaciones de DM con las proteínas TRAPPC12 a 100ns. a) Superficie electrostática de la TRAPPC12 L697, el color rojo indica zonas susceptibles a ataques electrofílicos; el color blanco, zonas hidrofóbicas y el color azul, zonas susceptibles a ataques nucleofílicos. Las unidades están en kBT/e , donde kB es la constante de Boltzmann; T , la temperatura del sistema y e , es la carga del electrón. b) Gráfico del RMSD de ambas proteínas, el color verde es la estructura L697 y el púrpura, la estructura V697. c) Superficie electrostática de la TRAPPC2 V697.

El gráfico RMSD (Figura 3.13b) nos indica el cambio conformacional que sufre la proteína a través de la simulación, comparándola con la estructura inicial y es un excelente indicador de la estabilidad proteica. En este caso, los resultados muestran que la estructura de la TRAPPC12 V697 (línea púrpura), estabiliza a partir de los 20 ns, teniendo un RMSD de 1.55 ± 0.18 nm en comparación con la estructura nativa (L697, línea verde), cuyo RMSD es de 2.01 ± 0.44 , lo cual se observa en los saltos bruscos en las líneas del gráfico. Esto puede ser ocasionado por la diferencia en el plegamiento de las estructuras, mientras que la L697 mantiene una conformación plegada, la estructura V697 abre, manteniéndose este desplegamiento a lo largo de la simulación.

Al analizar las superficies del potencial electrostático (Figura 3.13 a y 3.13b), se observan dos zonas claramente diferenciadas por su capacidad de interactuar con otros sistemas. Mientras que la zona N-terminal es una zona polar, con predominio a las interacciones con moléculas o sustancias electrofílicas, indicado por el color rojo intenso; el dominio C-terminal es ligeramente menos electrostático y con más propensión a las interacciones hidrofóbicas (zonas blancas). Entre ambas estructuras se presenta también una clara diferencia en cuanto a la capacidad de la proteína de interactuar con su entorno en la zona C-terminal.

En la estructura L697 se presenta una zona entre los residuos Leu440 y Lys510 de color rojo intenso por lo que se favorecen las interacciones polares, mientras que en la estructura V697, esta zona es apolar. Haciendo un acercamiento a la zona de la mutación, puede observarse la diferencia electrostática entre la leucina y la valina, la primera tiene carácter hidrófobo, mientras que la valina tiene carácter hidrofílico. Sin embargo, al observar la superficie del potencial electrostático por átomo entre los residuos His690 – Gln710, puede observarse que el aporte electrostático de la valina es la de neutralizar el carácter electrostático de la zona, dándole el carácter hidrófobo en ese dominio.

Por otro lado, si analizamos a los residuos de forma individual y su efecto en la energía y estructura de las proteínas podemos encontrar diferencias importantes entre los dos sistemas en estudio. Las cuales observamos en la Figura 3.14.

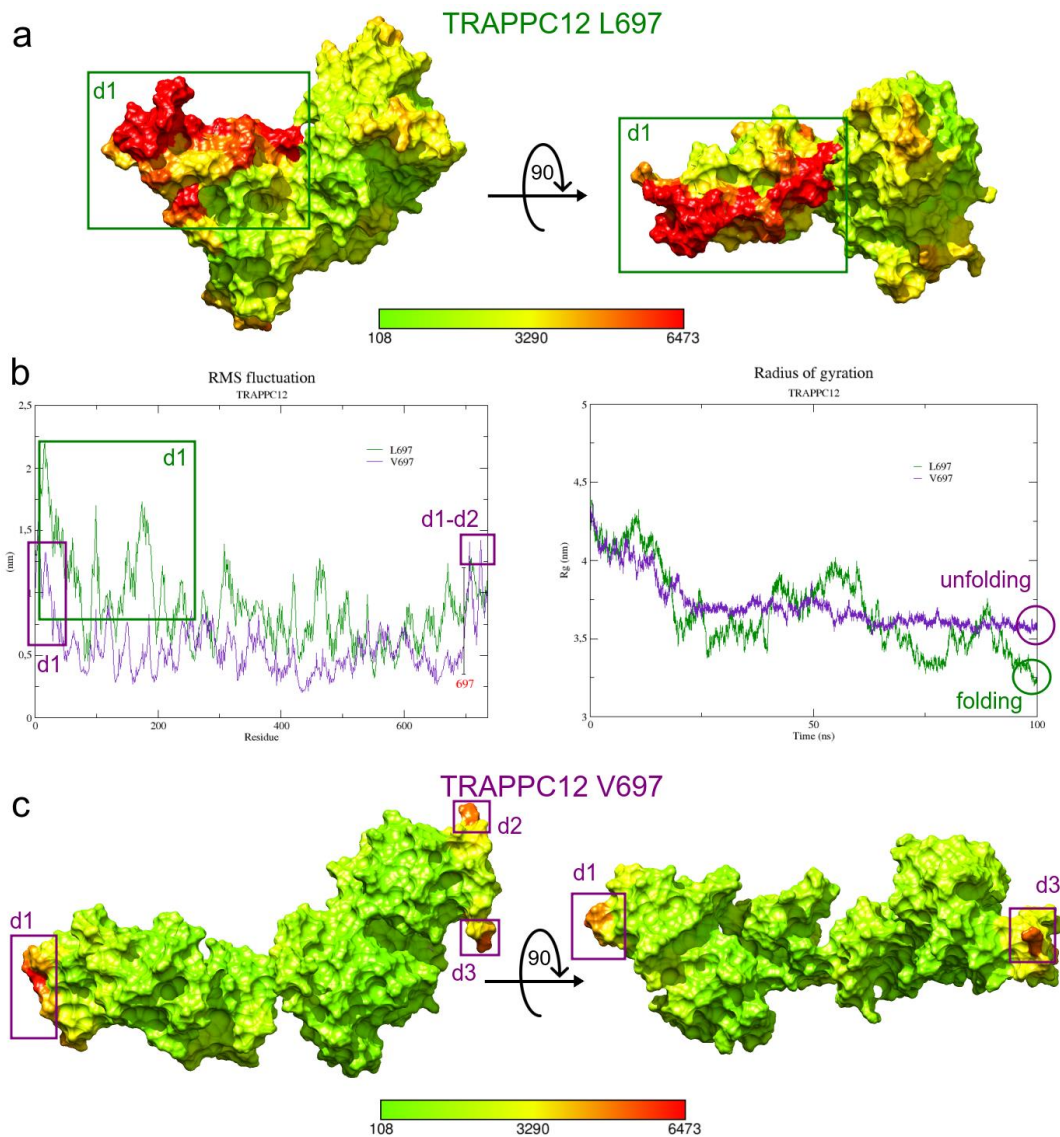


Figura 3.14: Análisis estructural y vibracional de los sistemas en estudio. a) Mapeo energético del factor-B sobre la superficie isodensa de la TRAPPC12 L697. Los colores representan el nivel energético (alta vibración) de los residuos; en color verde se muestran los residuos de baja energía, mientras que los de color rojo son los residuos de alta energía; el color amarillo indica vibraciones medias. En la estructura se observa un área de alta energía en la zona inicial del dominio N-terminal identificada como d1. b) Gráfico RMSF y del radio de giro de las proteínas. En el primer gráfico puede observarse la cantidad de residuos involucrados en el dominio d1 para la estructura L697 y los tres dominios d1, d2 y d3 de la estructura V697. En el segundo gráfico se observa el grado de compactación de cada sistema. Se observa que la estructura nativa se compacta y la V697 sufre un proceso de desplegamiento. c) Mapeo energético de la estructura TRAPPC12 V697, en la superficie se observan tres dominios de alta energía, identificados como d1, d2 y d3.

El gráfico de RMSF (Figura 3.14b), mide el promedio vibracional de cada residuo y nos sirve para dilucidar en qué zonas de la proteína se desestabiliza la estructura o, en casos especiales, que sitios pueden ser susceptibles a interacciones o ataques debido a la alta frecuencia vibracional. En el caso de la estructura L697, se observa una alta fluctuación (0.89 ± 0.32 nm), sobre todo en el dominio comprendido entre los residuos Met1 y Gly211 (d1), la cual, aunada al estudio del potencial electrostático, nos muestra que esta zona es activa y que interaccionará a sustancias electrofílicas. Por otra parte, en la estructura V697, la fluctuación es menor (0.55 ± 0.22 nm) y aunque se identifican tres dominios de alta fluctuación, d1 (Met1-Ala22), d2 (Ser705-Lys712) y d3 (Gly721-Thr729), éstos comprenden pocos residuos. Un punto interesante a recalcar en la estructura V697 es que los dominios d2 y d3 son cercanos a la mutación L697V y que, a pesar de ser una zona electrostáticamente neutra, esta alta fluctuación indica que en esos dominios se encuentran zonas de alta interacción con sustancias hidrofóbicas. Las Figuras 3.14b y 3.14c se muestran los mapeos del factor-B sobre la superficie isodensa de ambas proteínas y los dominios de alta energía están identificados por recuadros.

Un punto importante que se observó después de los 100 ns de simulación, es la diferencia en el plegamiento entre las dos estructuras. Esto es posible medirlo a partir del cálculo del radio de giro, el cual, mide el grado de compactación estructural de las proteínas. Este dato nos indica si una proteína está plegándose, como el caso de la estructura L697, cuyo promedio en el radio de giro fue de 3.64 ± 0.28 nm; o si la estructura está sufriendo un desplegamiento estructural, como se observa que es el caso de la estructura V697 (3.71 ± 0.16 nm en el radio de giro).

Gracias a este cálculo se identifica que la proteína mutada (V697) presenta un cambio en su estructura en comparación con la proteína nativa (L697). La valina encontrada en esa posición ocasiona que la proteína se desdoble y amplie su estructura, generando cambios en cuanto interacciones y funciones.

CONCLUSIONES

Primera, se encontró el gen TRAPPC 12 mutado en una familia de Arequipa, esta mutación está relacionada con la Enfermedad de Alzheimer, por su alta expresión en células de cerebro con EA y relación con enfermedades neurodegenerativas.

Segunda, con las muestras de saliva y la información recolectadas se implementó un Biobanco en el Laboratorio de Genómica y Enfermedades Neurovascular bajo un sistema de codificación y organización que garantiza su seguridad.

Tercera, los miembros de la familia estudiada manifiestan una mutación tipo Missense. En la posición número 2089 del RNA se da el cambio de la base Citosina por la base Guanina generando un cambio en la proteína. Luego de contrastar la información obtenida con diversas bases de datos genéticos se concluyó que la expresión de este tipo de mutación se da en bajas proporciones a nivel mundial.

Cuarta, se emplearon las técnicas scRNA y dinámica molecular para validar la mutación hallada, estas son herramientas apropiadas para el tipo de estudio realizado ya que emplean softwares de libre acceso que facilitan el libre manejo de información.

SUGERENCIAS

1. A lo largo de la realización de este proyecto se detectaron ciertos puntos que podrían mejorar y facilitar el estudio genético de grupo de individuos. En la actualidad existen diversas técnicas que se emplean para realizar un análisis del genoma, estas abaratan costos y reducen tiempo, pero el principal objetivo es obtener información limpia, sin errores o duplicaciones. Una recomendación sería emplear la técnica que analiza polimorfismos de longitud de fragmentos de restricción (RLFP), esta permite identificar fragmentos de diferentes longitudes empleando las enzimas de restricción. Las enzimas de restricción cortan el DNA en secuencias específicas de nucleótidos, permitiendo así realizar un análisis de variantes. Es decir, ayuda a identificar la mutación o variación que se puede dar dentro del genoma estudiado y disminuye el uso de reactivos y exposición del investigador a reactivos tóxicos ya que se dispone de páginas web que brindan este tipo de información gratuitamente.
2. Otra sugerencia es fomentar la creación de biobancos a nivel nacional, resaltar su función e importancia para aumentar el conocimiento sobre diversos tipos de enfermedades, no solo sobre el cáncer. La población padece diversos males como diabetes o la hipertensión arterial, pero existen enfermedades degenerativas que también deben ser estudiadas ya que diezman la calidad de vida de las personas, familiares y cuidadores. Es importante crear conciencia y educar a la población para que conozcan factores de riesgo o características que nos predisponen a manifestarlas.
3. Se recomienda también, fomentar la investigación sobre enfermedades neurodegenerativas y realizar estudios genéticos que permitan identificar posibles causas, tratamientos específicos y diagnósticos tempranos para desacelerar el avance de la enfermedad.

REFERENCIAS BIBLIOGRAFICAS

1. Research Arc. Neurodegenerative Disease [Internet]. Volumen 2; Disponible en: www.cellpress.com
2. Masters CL, Bateman R, Blennow K, Rowe CC, Sperling RA, Cummings JL. Alzheimer's disease. *Nat Rev Dis Primer*. 15 de octubre de 2015;1(1):1-18.
3. Ridge PG, Mukherjee S, Crane PK, Kauwe JS, Consortium ADG. Alzheimer's disease: analyzing the missing heritability. *PloS One*. 2013;8(11).
4. Long JM, Holtzman DM. Alzheimer Disease: An Update on Pathobiology and Treatment Strategies. *Cell*. octubre de 2019;179(2):312-39.
5. Lanoiselée H-M, Nicolas G, Wallon D, Rovelet-Lecrux A, Lacour M, Rousseau S, et al. APP, PSEN1, and PSEN2 mutations in early-onset Alzheimer disease: A genetic screening study of familial and sporadic cases. *PLoS Med* [Internet]. 28 de marzo de 2017 [citado 3 de septiembre de 2020];14(3). Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5370101/>
6. Cornejo-Olivas MR, Yu C-E, Mazzetti P, Mata IF, Meza M, Lindo-Samanamud S, et al. Clinical and molecular studies reveal a PSEN1 mutation (L153V) in a Peruvian family with early-onset Alzheimer's disease. *Neurosci Lett*. 20 de marzo de 2014;563:140-3.
7. González Sureda D. Escritura de una revisión sobre las Bases Moleculares de la Enfermedad de Alzheimer. 2015;
8. Zhang X, Gao F, Wang D, Li C, Fu Y, He W, et al. Tau pathology in Parkinson's disease. *Front Neurol*. 2018;9:809.
9. Gra Menéndez S, Padrón Pérez N, Llibre Rodríguez J de J. Péptido beta amiloide, proteína Tau y enfermedad de Alzheimer. *Rev Cuba Investig Bioméd*. 2002;21(4):253-61.
10. Huang Y, Mahley RW. Apolipoprotein E: Structure and function in lipid metabolism, neurobiology, and Alzheimer's diseases. *Neurobiol Dis*. diciembre de 2014;72:3-12.
11. Kraft L, Serpell LC, Atack JR. A Biophysical Approach to the Identification of Novel ApoE Chemical Probes. *Biomolecules*. febrero de 2019;9(2):48.
12. Utermann G. Alzheimer's Disease: The apolipoprotein E connection. *Curr Biol*. abril de 1994;4(4):362-5.
13. Romano M, Nissen MD, Del Huerto N, Parquet C. Enfermedad de alzheimer. *Rev Posgrado Vía Cátedra Med*. 2007;75:9-12.

14. Fuentes P, Slachevsky Ch A. Enfermedad de Alzheimer: Actualización en terapia farmacológica. *Rev Médica Chile*. 2005;133(2):224-30.
15. Kinkorová J. Biobanks in the era of personalized medicine: objectives, challenges, and innovation. *EPMA J* [Internet]. 22 de febrero de 2016 [citado 29 de junio de 2020];7(1). Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4762166/>
16. Mora E, Robb JA, Stefanoff G, Mellado RH, Coppola D, Muñoz-Antonia T, et al. Proceedings of the 1st Puerto Rico Biobanking Workshop. *Rev Recent Clin Trials*. 2014;9(4):233-44.
17. Dawber TR KWB. The Framingham Study An Epidemiological Approach to Coronary Heart Disease. octubre de 1966;XXXIV(4):553-5.
18. About UK Biobank | UK Biobank [Internet]. [citado 10 de marzo de 2020]. Disponible en: <https://www.ukbiobank.ac.uk/about-biobank-uk/>
19. OECD Glossary of Statistical Terms - Biobank Definition [Internet]. [citado 11 de julio de 2020]. Disponible en: <https://stats.oecd.org/glossary/detail.asp?ID=7220>
20. Co-Operation O for E, Development. OECD guidelines on human biobanks and genetic research databases. OECD Paris; 2009.
21. 2012 Best Practices for Repositories *Collection, Storage, Retrieval, and Distribution of Biological Materials for Research* International Society for Biological and Environmental Repositories. *Biopreservation Biobanking*. abril de 2012;10(2):79-161.
22. Asslaber M, Zatloukal K. Biobanks: transnational, European and global networks. *Brief Funct Genomic Proteomic*. 20 de agosto de 2007;6(3):193-201.
23. Serrano-Díaz N, Páez-Leal MC, Luna-González ML, Guío-Mahecha E. Biobanco: Herramienta fundamental para la investigación biomédica actual. *Rev Univ Ind Santander Salud*. 48(1):97-117.
24. Riegman PHJ, Morente MM, Betsou F, de Blasio P, Geary P. Biobanking for better healthcare. *Mol Oncol*. octubre de 2008;2(3):213-22.
25. Kang B, Park J, Cho S, Lee M, Kim N, Min H, et al. Current status, challenges, policies, and bioethics of biobanks. *Genomics Inform*. 2013;11(4):211.
26. Solari AJ. *Genética humana: fundamentos y aplicaciones en medicina*. Ed. Médica Panamericana; 2004. 572 p.
27. Yourkowitzky RL, Dehesa AZ, González PG. *Introducción a la genética humana*. Editorial El Manual Moderno; 2013. 332 p.
28. Gómez EG. INDICACIONES DEL ESTUDIO GENÉTICO. :6.
29. Ragoussis J. Genotyping technologies for all. *Drug Discov Today Technol*. junio de 2006;3(2):115-22.

30. López de Heredia U. Las técnicas de secuenciación masiva en el estudio de la diversidad biológica. *Munibe Cienc Nat* [Internet]. 2016 [citado 26 de julio de 2020];64. Disponible en: <http://www.aranzadi.eus/fileadmin/docs/Munibe/mcn.2016.64.07.pdf>
31. Sebastián Yagüe Á, Pascual-García A, Abascal F, Aguirre J, Bajic D, Baú D, et al. «Bioinformática con Ñ v1.0»: a collaborative project of young Spanish scientists to write a complete book about Bioinformatics. 21 de septiembre de 2014 [citado 2 de marzo de 2021]; Disponible en: <https://digital.csic.es/handle/10261/115052>
32. Kukol A. *Molecular modeling of proteins*. Vol. 443. Springer; 2008.
33. Simulación de Propiedades en Disolución de Polímeros. (Dr.), Universidad de Murcia.
34. Lopera F, Arcos M, Madrigal L, Kosik K, Cornejo W, Ossa J. Demencia tipo Alzheimer con agregación familiar en Antioquia, Colombia. *Acta Neurol Colomb*. 4 de octubre de 1994;10.
35. Rojas AG. Qué es la mutación Jalisco y por qué es importante estudiarla para saber más sobre el Alzheimer. *BBC News Mundo* [Internet]. 16 de enero de 2018 [citado 25 de marzo de 2020]; Disponible en: <https://www.bbc.com/mundo/noticias-42657547>
36. Chen G, Ning B, Shi T. Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. *Front Genet* [Internet]. 2019 [citado 21 de febrero de 2021];10. Disponible en: <https://www.frontiersin.org/articles/10.3389/fgene.2019.00317/full>
37. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005;26(16):1781-802.
38. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc*. 2010;5(4):725-38.
39. Xu D, Zhang Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J*. 2011;101(10):2525-34.
40. Lemkul J. From proteins to perturbed Hamiltonians: A suite of tutorials for the GROMACS-2018 molecular simulation package [article v1. 0]. *Living J Comput Mol Sci*. 2019;1(1):5068.
41. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. *J Comput Chem*. 2005;26(16):1701-18.

ANEXOS

Anexo 1: Directiva para Muestras Biológicas en Ensayos Clínicos Directiva N°
INS/OGAT-V.01

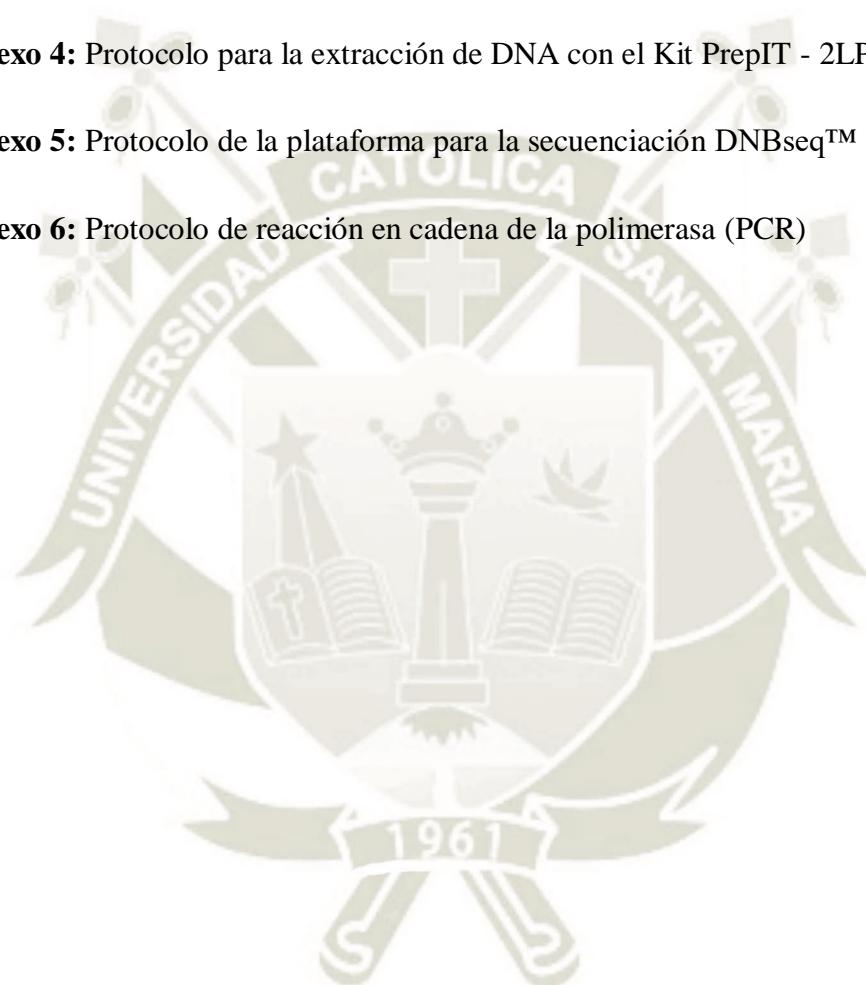
Anexo 2: Inserto tubo colector de saliva Orangene DNA

Anexo 3: Consentimiento informado para el estudio

Anexo 4: Protocolo para la extracción de DNA con el Kit PrepIT - 2LP

Anexo 5: Protocolo de la plataforma para la secuenciación DNBseq™

Anexo 6: Protocolo de reacción en cadena de la polimerasa (PCR)



**Anexo 1: Directiva para Muestras Biológicas en Ensayos Clínicos Directiva N°
INS/OGAT-V.01**

Directiva para Muestras Biológicas en Ensayos Clínicos
Directiva N° INS/OGAT-V.01

Directiva para Muestras Biológicas en Ensayos Clínicos

Finalidad.

Proteger los derechos del sujeto en investigación que participa en los Ensayos Clínicos con relación al uso de muestras biológicas.

1. Objetivos

1.1. Objetivo general:

Normar la utilización de muestras biológicas en el contexto de los ensayos clínicos

1.2. Objetivos específicos:

- 1.2.1. Normar la obtención, traslado, uso, almacenamiento y disposición de las muestras biológicas.
- 1.2.2. Normar el procedimiento que garantice la protección de datos generados a partir de las muestras biológicas.

2. Ámbito de Aplicación.

La presente Directiva es de aplicación a todas las personas naturales o jurídicas, públicas o privadas, nacionales o extranjeras, que realicen o estén vinculadas con los ensayos clínicos en seres humanos en el territorio nacional, incluyéndose a los biobancos en lo que corresponda.

3. Base Legal.

1. Ley General de Salud. Ley N° 26842
2. Ley del Ministerio de Salud Ley N° 27657
3. Ley que regula el Transporte Terrestre de Materiales y Residuos Peligrosos. LEY N° 28256. 2004
4. Decreto Supremo 017-2006 que aprueba el Reglamento de Ensayos clínicos en el Perú.
5. Resolución Jefatura N° 310-J-OPE/INS que aprueba la Directiva No 001-INS/OGAT-V.02: "Directiva para la Elaboración, Revisión, Aprobación, Difusión, Actualización y Control de los Documentos Normativos del Instituto Nacional de Salud".
6. Resolución Jefatural N° 478-2005-J-OPD/INS que aprueba el Manual Bioseguridad en laboratorios de ensayo, biomédicos y clínicos. Serie de Normas Técnicas N° 18. Ministerio de Salud - Instituto Nacional de Salud.

4. Disposiciones Generales.

- 4.1. Toda muestra biológica en un ensayo clínico sólo podrá ser obtenida previo consentimiento informado.
- 4.2. La obtención de muestras biológicas de poblaciones vulnerables en el desarrollo de un Ensayo Clínico debe seguir lo establecido, en lo que corresponda, a lo dispuesto en el Título III, Capítulos I: De la protección de los sujetos en investigación y II: Del consentimiento informado del Reglamento de Ensayos Clínicos en el Perú.
- 4.3. La obtención de la muestra biológica no puede representar un riesgo significativo para el sujeto en investigación, ni para su comunidad.

Directiva para Muestras Biológicas en Ensayos Clínicos
Directiva N° INS/OGAT-V.01

- 4.4. Se debe asegurar el cumplimiento de las normas de bioseguridad y calidad del procedimiento en la obtención, conservación, transporte, uso y disposición de las muestras biológicas. Así como el cumplimiento de las Buenas Prácticas Clínicas del personal implicado en el ensayo clínico.
- 4.5. Se debe cautelar el tratamiento de la información generada de los resultados de las muestras biológicas.
- 4.6. El uso de muestras biológicas dentro de estudios de farmacogenética y farmacogenómica, se regirán por las disposiciones contempladas en este documento.
- 4.7. **Definiciones Operativas.**
- a. **Anonimización:** Proceso por el cual deja de ser posible establecer por medios razonables el nexo entre una muestra biológica y el sujeto al que se refiere. Es aplicable también a un dato.
 - b. **Biobanco:** Establecimiento público o privado, que acoge una colección de muestras biológicas concebida con fines diagnósticos o de investigación biomédica y organizada como una unidad técnica con criterios de calidad, orden y destino.
 - c. **Cegamiento:** Procedimiento en el cual una o más partes del estudio desconocen las asignaciones al tratamiento. El cegamiento simple generalmente se refiere a que los sujetos en investigación desconocen la asignación; el cegamiento doble se refiere a que los sujetos en investigación e investigadores desconocen la asignación al tratamiento, y el cegamiento triple se refiere a que los sujetos en investigación, los investigadores y el que analiza los resultados desconocen la asignación al tratamiento
 - d. **Consentimiento Informado:** Es el acto del individuo en el que expresa voluntariamente su aceptación de participar en un estudio, siendo el resultado de un proceso de información y explicación detallada sobre todos los aspectos de la investigación que permitan su toma de decisión, y que puede suspenderse en cualquier momento a solicitud del propio interesado.
 - e. **Dato anónimo:** Dato registrado sin un nexo con una persona identificada o identificable.
 - f. **Dato codificado o reversiblemente disociado:** Dato no asociado a una persona identificada o identificable por haberse sustituido o desligado la información que identifica a esa persona utilizando un código que permita la operación inversa.
 - g. **Ensayo Clínico:** Toda investigación que se efectúe en seres humanos, para determinar o confirmar los efectos clínicos, farmacológicos, y/o demás efectos farmacodinámicos; detectar las reacciones adversas; estudiar la absorción, distribución, metabolismo y eliminación de uno o varios productos en investigación, con el fin de determinar su eficacia y/o su seguridad.
 - h. **Muestra biológica:** Cualquier material biológico de origen humano susceptible de conservación y que pueda albergar información sobre la dotación genética característica de una persona; que incluye: sangre y sus derivados, tejidos u órganos y sus remanentes, obtenido dentro de un ensayo clínico.
 - i. **Muestra biológica anonimizada o irreversiblemente disociada:** Muestra que no puede asociarse a una persona identificada o identificable por haberse destruido el nexo con toda información que identifique al sujeto, o porque dicha asociación exige un esfuerzo no razonable.
 - j. **Muestra biológica codificada o reversiblemente disociada:** Muestra no asociada a una persona identificada o identificable por haberse sustituido o desligado la información que identifica a esa persona utilizando un código que permita la operación inversa.
 - k. **Patrocinador:** Persona individual, grupo de personas, empresa, institución u organización, incluidas las académicas, con representatividad legal en el país, que asume la responsabilidad de la iniciación, mantenimiento y/o financiación de un ensayo clínico. Cuando un Investigador independiente

Directiva para Muestras Biológicas en Ensayos Clínicos
Directiva N° INS/OGAT-V.01

inicia y toma toda la responsabilidad de un ensayo clínico, entonces el Investigador asume el papel de patrocinador.

- l. **Población vulnerable:** Aquellas personas absoluta o relativamente incapaces de proteger sus propios intereses. Específicamente, pueden tener insuficiente poder, inteligencia, educación, recursos, fuerza u otros atributos necesarios para proteger sus intereses. Por ejemplo grupos subordinados, pacientes con enfermedades incurables, personas en asilos, sin empleo o indigentes, pacientes en situaciones de emergencia, grupos étnicos de minoría, personas sin hogar, nómadas, refugiados, menores y aquellos que no pueden dar su consentimiento.
- m. **Procedimiento invasivo:** Toda intervención realizada con fines de investigación que implique un riesgo físico para el sujeto en investigación.
- n. **Producto en Investigación:** A efecto del presente Reglamento, el producto en investigación es la forma farmacéutica de una sustancia activa (medicamento), dispositivo médico, producto de origen biológico o afines que se investiga o se utiliza como comparador activo en un ensayo clínico, incluidos los productos con registro sanitario cuando se utilicen o combinen (en la formulación o en el envase) de forma diferente a la autorizada, o cuando se utilicen para tratar una indicación no autorizada, o para obtener más información sobre su uso autorizado.
- o. **Sujeto en investigación:** Es el individuo que participa voluntariamente en un ensayo clínico.
- p. **Sustancias infecciosa Categoría A:** Aquella sustancia que cuando ocurre su exposición, es capaz de causar incapacidad permanente, enfermedad fatal o para toda la vida en humanos y animales. Esta exposición ocurre cuando la sustancia infecciosa es liberada fuera del empaque de protección, teniendo contacto físico con las entidades anteriormente mencionadas. Si la sustancia infecciosa, causa enfermedad sólo en el humano o en humanos y animales se deberá asignar el número de naciones unidas UN 2814, pero si ésta afecta sólo a animales se colocará el número UN 2900. Ejemplos de sustancias infecciosas que afectan sólo a humanos (UN 2814): Hantavirus, cultivos de Bacillus anthracis, Coccidioides immitis, Mycobacterium tuberculosis, Brucella melitensis, Yersinia pestis, Hepatitis B. Entre las sustancias infecciosas que afectan sólo a animales (UN 2900) tenemos a cultivos del virus de fiebre clásica del cerdo, Mycoplasma mycoides, virus de la peste de pequeños rumiantes, virus del carnero. (18)
- q. **Trazabilidad:** Capacidad de asociar un material biológico determinado con información registrada referida a cada paso en la cadena de su obtención, así como a lo largo de todo el proceso de investigación.

5. Disposiciones Específicas.

5.1. De la obtención de muestras biológicas:

- 5.1.1. Se debe asegurar la calidad técnica, en capacitación y experiencia, del personal que obtendrá las muestras biológicas.
- 5.1.2. Se debe garantizar la calidad del procedimiento en la toma de muestras biológicas.
- 5.1.3. Se debe asegurar el cumplimiento de las normas de bioseguridad en la obtención de las muestras biológicas.
- 5.1.4. Se debe garantizar las condiciones de esterilidad en la extracción, manipulación y la conservación de la muestra biológica hasta el destino final.
- 5.1.5. Se debe asegurar el etiquetado y codificación de las muestras en el momento de la obtención.

Directiva para Muestras Biológicas en Ensayos Clínicos
Directiva N° INS/OGAT-V.01

5.2. Almacenamiento de muestras biológicas y sus remanentes en Ensayos Clínicos:

- 5.2.1. Se debe asegurar la conservación en estabilidad y esterilidad de las muestras biológicas en todo momento.
- 5.2.2. Las muestras biológicas deben almacenarse en cumplimiento de las normas de bioseguridad.
- 5.2.3. Se debe garantizar la calidad de la conservación de las muestras biológicas según su naturaleza.
- 5.2.4. La decisión de que las muestras biológicas sean anonimizadas o irreversiblemente disociadas debe ser justificada por el patrocinador e informada en el consentimiento informado al sujeto en investigación.
- 5.2.5. Se debe asegurar la trazabilidad de las muestras biológicas
- 5.2.6. La conservación de las muestras biológicas se realiza en biobancos. En el Perú los biobancos deben estar registrados en el Instituto Nacional de Salud. En caso que la conservación de la muestra biológica se realice en el extranjero, el patrocinador debe indicar el biobanco, especificando su ubicación y la acreditación o autorización para su funcionamiento.
- 5.2.7. Las muestras biológicas se conservarán únicamente en tanto sean necesarias para los fines que justificaron su recolección, salvo que el sujeto en investigación haya otorgado su consentimiento explícito para otros usos posteriores.
- 5.2.8. Cuando el consentimiento informado sea retirado total o parcialmente y se refiera a cualquier uso de la muestra, se procederá a su inmediata destrucción, sin perjuicio de la conservación de los datos resultantes de las investigaciones que se hubiesen realizado con carácter previo.
- 5.2.9. El tiempo de almacenamiento de la muestra biológica debe estar especificado y justificado en el Consentimiento Informado.

5.3. Biobancos

5.3.1. Organización de un Biobanco:

- 5.3.1.1. La institución que acoge el biobanco es la responsable de su custodia y debe dotarlo de una estructura, una organización y un reglamento interno escrito que determine su funcionamiento y en el que se definan las responsabilidades, la política de la calidad y los objetivos asistenciales y/o científicos.
- 5.3.1.2. El Biobanco debe contar como mínimo con un director científico y un responsable del archivo de muestras
- 5.3.1.3. El biobanco debe acogerse al asesoramiento de un Comité de Ética que garantice el cumplimiento de los principios éticos aplicables a la investigación biomédica de los proyectos que incorporen muestras de origen humano al biobanco, así como del uso que se haga de las mismas.
- 5.3.1.4. El biobanco debe acogerse a un Comité Científico que asesore al responsable del biobanco sobre la dirección y objetivos científicos del biobanco y desarrolle los estándares de funcionamiento del mismo. Igualmente, aconsejará sobre las actividades de investigación a desarrollar, que sean de importancia estratégica en la explotación óptima del biobanco, aprobará cualquier transferencia de muestras a terceras partes y asesorará en la priorización de la cesión de las muestras.
- 5.3.1.5. El biobanco debe organizar un sistema de información y el correspondiente protocolo de seguridad que garantice que los datos personales de los sujetos fuente, los datos de salud y los derivados de los resultados de la investigación estén protegidos de acuerdo con la normativa vigente nacional e internacional, garantizando en todo momento la confidencialidad de los datos.

Directiva para Muestras Biológicas en Ensayos Clínicos
Directiva N° INS/OGAT-V.01

5.3.1.6. Funciones del director del Biobanco:

- a) Velar por el cumplimiento de la normatividad vigente
- b) Mantener un registro de actividades del biobanco
- c) Garantizar la calidad, la seguridad y la trazabilidad de los datos y muestras biológicas almacenadas.
- d) Elaborar un informe anual de actividades, el cual deberá ser enviado al INS.
- e) Atender las consultas o reclamos que puedan dirigirse al biobanco
- f) Elaborar el documento de buenas prácticas del biobanco.
- g) Elaborar la memoria descriptiva que recoja las características de las muestras, los criterios de inclusión, los propósitos para los cuales fueron obtenidas las muestras y el uso de las mismas hasta el momento.

5.3.1.7. El responsable del archivo de muestras biológicas, atenderá las solicitudes de ejercicio de los derechos de acceso, rectificación, cancelación u otras formuladas por los sujetos donantes en conformidad con la normatividad vigente.

5.3.1.8. Todo el personal relacionado directamente a las muestras biológicas, deberá estar capacitado en Buenas Prácticas de bioseguridad, Buenas Prácticas de conservación, Buenas Prácticas de almacenamiento y transporte de las mismas, según corresponda a sus funciones y responsabilidades y todos sin excepción en Buena Prácticas clínicas

5.3.2. Registro Nacional de Biobancos:

Todos los Biobancos deben registrarse en el Instituto Nacional de Salud, para lo cual deben presentar los siguientes documentos:

- a) Solicitud de registro de Biobanco según formato del INS.
- b) Autorización de la Institución donde funcionará el Biobanco, indicando el nombre del Director y del responsable del archivo de muestras.
- c) Informe indicando las características de las muestras que serán almacenadas, la forma de obtención de las mismas, y los objetivos del almacenamiento.
- d) Documento de Buenas Prácticas de Biobanco
- e) Carta de adscripción del Comité de Ética de investigación
- f) Carta de adscripción del Comité Científico.
- g) Declaración jurada firmada por el director de la institución donde funcionará el Biobanco y el director del mismo, sobre el acondicionamiento de las instalaciones, según el anexo respectivo.
- h) Declaración jurada garantizando la calidad de la conservación, la seguridad en el almacenamiento y trazabilidad de las muestras biológicas, firmada por el director de la Institución donde funcionará el Biobanco y el Director del mismo.
- i) Declaración jurada de resguardo de la confidencialidad de la información relacionada a las muestras biológicas y sus donantes, firmada por cada una de las personas que labora en el Biobanco.
- j) Currículo vitae no documentado del Director, del responsable del archivo de muestras y de todo el personal en contacto con las muestras biológicas.

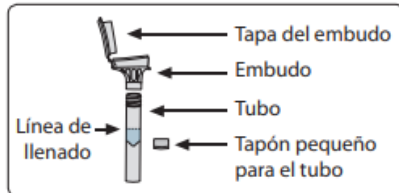
5.3.3. **Inspecciones y medidas de control al Biobanco:** El Instituto Nacional de Salud, llevará a cabo inspecciones periódicas para garantizar que los biobancos cumplan con las condiciones de instalación, organización y funcionamiento con las que fue autorizado su registro. La notificación de las inspecciones se efectuará con una anticipación no menor de dos (02) ni mayor de cinco (05) días útiles.

Directiva para Muestras Biológicas en Ensayos Clínicos
Directiva N° INS/OGAT-V.01

- 5.3.4. **Clausura o cierre del biobanco:** El Instituto Nacional de Salud podrá decidir, por decisión propia o a solicitud del Biobanco, la clausura o el cierre del biobanco en los casos en los que no se cumplan los requisitos sobre su creación, organización y funcionamiento establecidos en este documento, o cuando su titular manifieste la voluntad de no continuar con su actividad. En dicha Para lo cual se el INS emitirá una Resolución Directoral, en la cual se indicará el destino de las muestras almacenadas en el biobanco que va a ser clausurado o cerrado, las cuales deberán ser especificadas por el Director del Biobanco previamente al cierre.
- 5.4. **Del transporte de las muestras biológicas**
- 5.4.1. Si la muestra biológica será enviada al extranjero, se debe presentar un convenio en el que se establezcan los deberes y beneficios tanto del país donante como del país receptor en relación a la muestra y los resultados obtenidos a partir de ella. Dicho convenio debe garantizar la protección de nuestra diversidad genética y cultural. En conformidad con el artículo 18, 19 y 20 de la **Declaración Internacional sobre los Datos Genéticos Humanos de la UNESCO** sobre la Circulación y Cooperación Internacional y los artículos 17, 18 y 19 de la **Declaración Universal sobre el Genoma Humano y los Derechos Humanos de UNESCO** sobre Solidaridad y Cooperación Internacional.
- 5.4.2. Se debe aplicar el Manual de Bioseguridad en laboratorios de ensayo, biomédicos y clínicos, según lo estipulado en el Manual de Bioseguridad en laboratorios de ensayo, biomédicos y clínicos de la Serie de Normas Técnicas N° 18 Instituto Nacional de Salud.
- 5.4.3. El personal que manipulará y transportará las muestras, debe estar capacitado en normas de Bioseguridad, en Buenas Prácticas Clínicas y además tener experiencia en la manipulación y transporte de las mismas.
- 5.4.4. Se debe asegurar la identificación y garantía de la trazabilidad de las muestras biológicas para garantizar sus características originales.
- 5.4.5. Se debe asegurar el acondicionamiento de las muestras biológicas.
- 5.4.6. Se debe usar mensajerías especializadas en el transporte de muestras biológicas para los envíos.
- 5.4.7. Se debe asegurar el cumplimiento de las normas nacionales en el embalaje para transporte internacional.
- 5.4.8. Se debe asegurar el cumplimiento de las normas nacionales en el transporte y el cumplimiento de la normatividad internacional para la importación y exportación de muestras biológicas.
- 5.4.9. Las sustancias infecciosas de categoría A (sólo pueden ser transportadas en cajas que cuenten con las especificaciones de las Naciones Unidas clase 6.2 P620 (Figura 1)).



Anexo 2: Inserto tubo colector de saliva Orangene DNA



Precauciones para la recolección:

NO ingiera alimentos ni bebidas, no fume ni mastique goma de mascar 30 minutos antes de obtener su muestra de saliva.

NO quite la película plástica de la tapa del embudo.

Uso previsto: Para la recolección de ADN humano de muestras de saliva.

Contenido: El kit contiene líquido estabilizador.

Advertencias y precauciones: Enjuague con agua si el líquido estabilizador entra en contacto con los ojos o la piel. No lo ingiera. Consulte la hoja de datos de seguridad del material (MSDS) en www.dnagenotek.com.

Tapón pequeño, peligro de asfixia.

Almacenamiento: 15°C a 30°C

Resumen y explicación del kit:

Oragene-DNA es un kit de autorrecolección que contiene los materiales y las instrucciones para recolectar y estabilizar muestras de saliva.

Leyenda de la etiqueta:

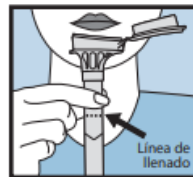
	Consulte el prospecto del paquete
	Recoja la saliva antes del (fecha de caducidad)
	Dispositivo médico para diagnóstico in vitro
	Número de catálogo
	Marca CE
	Precaución: consulte las instrucciones de uso
	Instrucciones de almacenamiento
	Representante autorizado
	Fabricante
	Número de lote

INSTRUCCIONES PARA EL USUARIO

Lea todas las instrucciones antes de la recolección

Procedimiento:

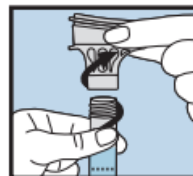
A la mayoría de las personas les lleva entre 2 y 5 minutos proporcionar una muestra de saliva siguiendo los pasos 1 a 5.



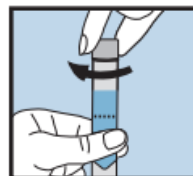
- 1 Escupa hasta que la cantidad de saliva líquida (no de burbujas) alcance la línea de llenado que se muestra en la figura n.º 1.



- 2 Sostenga el tubo en posición vertical con una mano. Cierre la tapa con la otra mano (como se muestra), empujándola con firmeza hasta que escuche un chasquido fuerte. El líquido de la tapa pasará al tubo para mezclarse con la saliva. Asegúrese de que la tapa esté correctamente cerrada.



- 3 Sostenga el tubo en posición vertical. Desensrosque el embudo del tubo.



- 4 Utilice el tapón pequeño para cerrar el tubo herméticamente.



- 5 Agite el tubo tapado durante 5 segundos. Deseche o recicle el embudo.

Anexo 3

CONSENTIMIENTO INFORMADO PARA EL ESTUDIO:

“ESTUDIO GENÉTICO DE LA ENFERMEDAD DE ALZHEIMER EN POBLACIÓN AREQUIPEÑA”

Datos de contacto:

Investigador: Milagros Ortiz Manrique, correo-e: michellemilagrosom@gmail.com,
teléfono móvil: 987189223

He sido invitado a participar de la investigación cuyo objetivo es *determinar la relación existente entre la Enfermedad de Alzheimer y una mutación genética que la genere.*

El procedimiento que se realizará es: **1) toma de muestras de saliva.**

La información de los datos proporcionados será guardada en forma confidencial bajo codificación manteniendo el anonimato. Sólo los investigadores tendrán acceso a la información que deriven de las muestras, obtenidas y no serán compartidos con ninguna persona o institución ajena al presente proyecto. Si los resultados de este estudio son publicados en revistas científicas, no se mostrará información que identifique a las personas que participaron en este estudio.

Habiendo leído el formato de consentimiento que se me ha entregado, he podido hacer las preguntas que considere necesarias sobre el estudio, las que han sido contestadas en un lenguaje comprensible y han sido mi entera satisfacción. Asimismo, entiendo que mi participación es VOLUNTARIA.

Al firmar este consentimiento no estoy renunciando a ninguno de mis derechos legales, según leyes locales.

ACEPTO VOLUNTARIAMENTE PARTICIPAR EN ESTE ESTUDIO.

Nombre del participante
DNI:

____/____/____

Fecha

Firma

Anexo 4: Protocolo para la extracción de DNA con el Kit PrepIT - 2LP

prepIT[®]•L2P

Laboratory protocol for manual purification of DNA from whole sample

Ethanol precipitation protocol and prepIT[®]•L2P reagent for the purification of genomic DNA from Oragene[®] products and ORAcollect[®] formats OC-175, OCD-100 and OCR-100. Not for use with OCD-100A.

Visit our website at www.dnagenotek.com for any additional languages and protocols.

Note: This protocol requires the use of a centrifuge capable of generating at least 3,500 × g to obtain optimal results.

The procedure is described for purifying DNA from the entire collected sample (approximately 1 mL to 4 mL total volume). The volumes shown should be adjusted for the actual collected volume.

Reagents included

- prepIT•L2P (catalog #: PT-L2P)

Equipment and reagents

- Centrifuge that accommodates 15 mL tubes, and is capable of generating at least 3,500 × g (see Table 2)
- 15 mL conical polypropylene tubes (e.g., BD Falcon #352196)
- Microcentrifuge capable of running at 15,000 × g (optional)
- 1.5 mL microtubes (e.g., Axygen #MCT-150-C)
- Air or water incubator at 50°C
- Ethanol (95% to 100%) at room temperature
- Ethanol (70%) at room temperature
- DNA storage buffer: TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) or similar solution

Pre-purification check

Weigh the sample to estimate the amount of saliva provided by the donor (see Table 1; not required for ORAcollect products). The amount of saliva collected is directly proportional to the amount of DNA recovered. As an example, if a donor has provided less than 2 mL of saliva, you should expect to recover a lower total yield from this sample. Correspondingly, a donor providing more than 2 mL of saliva should result in higher total yield.

Weight of kit (without sample)	Table 1	
Once a sample arrives at the lab, we suggest weighing the sample to estimate if the right amount of saliva was provided by the donor. You can expect some variability across donors as determined by weight of triggered kit with sample. The average weight of an empty kit is provided (Table 1). To calculate the amount of sample collected (assuming 1 g/mL), perform the following subtraction: Weight of kit containing sample - Weight of kit without sample = Amount of sample collected	Product #	Weight of kit without sample
	OG-250/OGR-250	14.15 g
	OG-500/OGD-500/OGR-500/OG-600*/OGD-600*/OGR-600*	6.81 g
	OG-510/OGD-510/OG-610*/OGD-610*	5.83 g
	OG-520	5.41 g
	OG-575/OGD-575/OGR-575/OG-675*/OGD-675*/OGR-675*	5.66 g
	ON-500/ON-600*	6.47 g
	* 2D bottom barcoded kits are 1 g heavier than the respective non-barcoded kit.	

PD-PB-015 Issue 11/2018-10

© 2018 DNA Genotek Inc., a subsidiary of OraSure Technologies, Inc., all rights reserved.

www.dnagenotek.com • support@dnagenotek.com

DNAgenotek



Anexo 5: Protocolo de la plataforma para la secuenciación DNBseq™

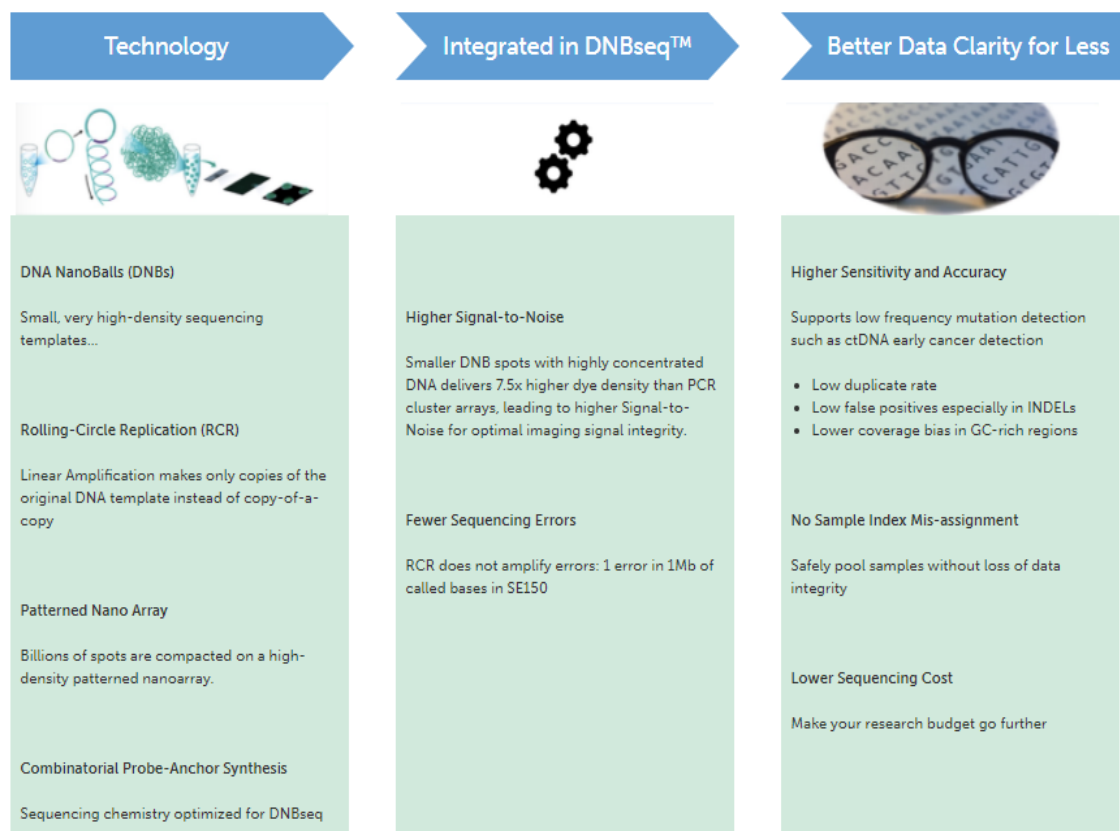


Proven DNBseq sequencing technology combines the power of DNA Nanoballs (DNB™), PCR-free Rolling Circle Replication, Patterned Nano Arrays and cPAS to deliver a new level of data clarity and affordability.

Contact us

The combination of linear amplification and DNB technology reduces the error rate while enhancing the signal, resulting in real advantages.

NGS data from DNBseq technology is well documented with over 800 peer-reviewed publications and is exclusively available from BGI.



Anexo 6: Protocolo de reacción en cadena de la polimerasa (PCR)

Taq DNA Polymerase

Cat. Nos.

18038-018 (100 units)

18038-067 (1500 units)

18038-042 (500 units)

18038-240 (5000 units)

Conc: 5 U/ μ l

Store at -20°C (non-frost-free)

Licensed for PCR

Description

Taq DNA Polymerase is isolated from *Thermus aquaticus* YT1. The enzyme consists of a single polypeptide with a molecular weight of approximately 94 kDa. *Taq* DNA polymerase is heat-stable and will synthesize DNA at elevated temperatures from single-stranded templates in the presence of a primer.

Component	Kit Size			
	100 U	500 U	1500 U	5.000 U
<i>Taq</i> DNA Polymerase	20 μ l	100 μ l	300 μ l	1000 μ l
10X PCR Buffer, Minus Mg	1.25 ml	2.5 ml	7.5 ml	20 ml
50 mM Magnesium Chloride	1 ml	1 ml	3 ml	10 ml

Storage Buffer

20 mM Tris-HCl (pH 8.0), 0.1 mM EDTA, 1 mM DTT, 50% (v/v) glycerol, Stabilizers

10X PCR Buffer

200 mM Tris-HCl (pH 8.4), 500 mM KCl

The PCR Buffer is supplied as a 10X concentrate and should be diluted for use.



Unit Definition

One unit incorporates 10 nmol of deoxyribonucleotide into acid-precipitable material in 30 minutes at 74°C. Unit assay conditions: 25 mM TAPS (pH 9.3), 50 mM KCl, 2 mM MgCl₂, 1 mM DTT, 0.5 mg/ml activated salmon sperm DNA, 0.2 mM dATP, dCTP, dGTP, dTTP

Basic PCR Protocol

The following basic protocol serves as a general guideline and a starting point for any PCR amplification. Optimal reaction conditions (incubation times and temperatures, concentration of *Taq* DNA Polymerase, primers, MgCl₂, and template DNA) vary and need to be optimized.

Critical parameters and troubleshooting information are documented in reference 1. PCR reactions should be assembled in a DNA-free environment. Use of "clean" dedicated automatic pipettors and aerosol resistant barrier tips are recommended. **Always** keep the control DNA and other templates to be amplified isolated from the other components.

1. Add the following components to a sterile 0.5-ml microcentrifuge tube sitting on ice:

<u>Components</u>	<u>Volume</u>	<u>Final Conc.</u>
10X PCR buffer minus Mg	10 µl	1X
10 mM dNTP mixture	2 µl	0.2 mM each
50 mM MgCl ₂	3 µl	1.5 mM
Primer mix (10 µM each)	5 µl	0.5 µM each
Template DNA	1–20 µl	n/a
<i>Taq</i> DNA Polymerase (5 U/µl)	0.2–0.5 µl	1.0–2.5 units
Autoclaved distilled water	to 100 µl	n/a

We recommend preparing a master mix for multiple reactions, to minimize reagent loss and enable accurate pipetting.

2. Mix contents of tube and overlay with 50 µl of mineral or silicone oil.
3. Cap tubes and centrifuge briefly to collect the contents to the bottom.

Basic PCR Protocol, continued

4. Incubate tubes in a thermal cycler at 94°C for 3 minutes to completely denature the template.
5. Perform 25–35 cycles of PCR amplification as follows:

Denature	94°C for 45 s
Anneal	55°C for 30 s
Extend	72°C for 1 min 30 s
6. Incubate for an additional 10 min at 72°C and maintain the reaction at 4°C. The samples can be stored at –20°C until use.
7. Analyze the amplification products by agarose gel electrophoresis and visualize by ethidium bromide staining. Use appropriate molecular weight standards.

"Hot-start" Protocol

In the "hot-start" method, the addition of *Taq* DNA Polymerase is withheld until the reaction temperature is at 80°C, to ensure high specificity of the products being synthesized.

1. Add all components as in the Basic PCR Protocol, except for the *Taq* DNA Polymerase.
2. Mix contents of tube and overlay with 50 µl of mineral or silicone oil.
3. Cap tubes and centrifuge briefly to collect the contents to the bottom.
4. Incubate tubes in a thermal cycler at 94°C for 3 minutes to completely denature the template.
5. After denaturation at 94°C, maintain the reaction at 80°C.
6. Add 0.2–0.5 µl of *Taq* DNA Polymerase (1.0–2.5 U) to each reaction. Be certain to add the enzyme beneath the layer of oil.
7. Continue with 25–35 cycles of denaturation, annealing and extension as in the Basic PCR Protocol.