



# Universidad Europea

TRABAJO FINAL DE MÁSTER  
MÁSTER EN BIOINFORMÁTICA

## **IDENTIFICACIÓN DE VARIANTES GENÉTICAS POTENCIALMENTE ASOCIADAS CON LA HIPERTENSIÓN INTRACRANEAL IDIOPÁTICA (HII) A TRAVÉS DEL ESTUDIO DEL GENOMA COMPLETO**

Autor: José Alberto Fernández Expósito

Tutor: Ángela Lacombe Antonelli

*CURSO ACADÉMICO 2023-2024*

## Índice de contenido

Agradecimientos.....	6
Resumen.....	7
Abstract.....	8
<b>1. Introducción.....</b>	<b>9</b>
<b>1.1. Contexto.....</b>	<b>9</b>
<b>1.2. Presentación clínica de la HII.....</b>	<b>11</b>
<b>1.3. Epidemiología.....</b>	<b>12</b>
<b>1.4. Pronóstico y tratamiento.....</b>	<b>13</b>
<b>1.5. Patogenesis de la HII.....</b>	<b>15</b>
<b>1.5.1. Homeostásis del LCF.....</b>	<b>15</b>
<b>1.5.2. Desregulación hormonal.....</b>	<b>17</b>
<b>1.5.3. Perfil metabólico.....</b>	<b>17</b>
<b>1.5.4. Mediadores pro-inflamatorios y sistema inmunitario.....</b>	<b>18</b>
<b>1.6. Estudios previos que relacionana la HII con variantes genéticas.....</b>	<b>20</b>
<b>1.7. La secuenciación de genoma completo en la práctica clínica.....</b>	<b>22</b>
<b>2. Objetivos.....</b>	<b>24</b>
<b>3. Material y métodos.....</b>	<b>24</b>
<b>3.1. Población de Estudio.....</b>	<b>24</b>
<b>3.2. Aspectos éticos y legales.....</b>	<b>25</b>
<b>3.3. Obtención de los datos.....</b>	<b>26</b>
<b>3.4. Software y bases de datos empleados.....</b>	<b>27</b>
<b>3.4.1. Software.....</b>	<b>27</b>
<b>3.4.2. Bases de datos.....</b>	<b>30</b>
<b>3.4.3. Instalación y puesta a punto.....</b>	<b>31</b>
<b>3.5. Estructura del flujo de trabajo.....</b>	<b>32</b>
<b>3.5.1. Control de calidad.....</b>	<b>32</b>
<b>3.5.2. Mapeo.....</b>	<b>32</b>
<b>3.5.3. Eliminaicón de duplicados y recalibración del valor de calidad.....</b>	<b>34</b>
<b>3.5.4. Variant calling.....</b>	<b>35</b>
<b>3.5.5. Anotación y priorización de variantes con <i>hiPhivePrioritiser</i>.....</b>	<b>35</b>

3.5.6. Anotación de variantes. Panel de genes de AQP y estudio GWAS.....	37
3.6. Procesamiento y análisis estadístico de variantes priorizadas.....	37
4. Resultados.....	41
4.1. Estudio de asociación genómica WGS.....	41
4.1.1. Variantes localizadas en zonas codificantes .....	41
4.1.2. Variantes localizadas en zonas no codificantes .....	42
4.2. Validación del estudio de acuoporinas.....	44
4.3. Validación del estudio de GWAS.....	45
5. Discusión.....	45
6. Conclusiones.....	50
Bibliografía.....	51
Anexo.....	63
Declaración del uso de herramientas de IA.....	64

## Índice de imágenes

<b>Imagen 1.</b> Anatomía del sistema ventricular y flujo del LCR en la cavidad craneal .....	9
<b>Imagen 2.</b> Mecanismos fisiopatológicos propuestos para el incremento de la PIC basados en la alteración del flujo del LCR .....	16
<b>Imagen 3.</b> Mecanismos fisiopatológicos propuestos para el incremento de la PIC causados o agravados por la obesidad .....	20
<b>Imagen 4.</b> Diagrama de flujo de una pipeline bioinformática general de análisis de WGS .....	23
<b>Imagen 5.</b> Representación del flujo de trabajo efectuado para la obtención del genoma del grupo de afectados por HII .....	26
<b>Imagen 6.</b> Parametros básicos de uno de los archivos FASTQ empleados en el análisis .....	33
<b>Imagen 7.</b> Gráficas de calidad media por base y calidad promedio por secuencia de un archivo FASTQ .....	33
<b>Imagen 8.</b> Distribución de la densidad del valor $(f_{casos} - f_{controles})^2$ antes y después de aplicar el filtrado de variantes basado en la frecuencia de aparición .....	39
<b>Imagen 9.</b> Gráfico Manhattan donde se recogen todas las variantes exónicas incluidas en el análisis .....	42
<b>Imagen 10.</b> Gráfico Manhattan donde se recogen todas las variantes no codificantes incluidas en el análisis .....	43
<b>Imagen 11.</b> Análisis de enriquecimiento de términos GO y rutas KEGG de los genes asociados a la HII .....	44
<b>Imagen 12.</b> Conteo de cada uno de las variantes estudiadas por Tellería-Orrriols et al. en nuestra población de estudio .....	44
<b>Imagen 13.</b> Conteo de las variantes relevantes en la publicación Kuehn et al. en nuestra población de estudio. ....	45
<b>Imagen 14.</b> Metabolismo de la esfingosina y su implicación en diferentes enfermedades. ..	47
<b>Imagen 15.</b> Genes asociados a los procesos biológicos más enriquecidos en el análisis de términos GO realizado sobre variantes no codificante. ....	48
<b>Imagen 16.</b> Genes asociados a los procesos biológicos más enriquecidos en el análisis de rutas KEGG realizado sobre variantes no codificante. ....	49



## Índice de fórmulas

<b>Fórmula 1.</b> Cálculo del Phred Score.....	33
<b>Fórmula 2.</b> Cálculo de la métrica de calidad QUAL.....	35
<b>Fórmula 3.</b> Cálculo del cuadrado de la diferencia de frecuencias.....	38

## Indice de tablas

<b>Tabla 1.</b> Variantes exónicas significativamente asociadas con la HII en nuestro grupo de estudio .....	41
<b>Tabla 2 .</b> Puntuaciones asignadas por las herramientas de predicción de patogenicidad a las variantes exónicas asociadas con la HII en nuestra población de estudio. ....	41
<b>Tabla 3.</b> Variantes estudiadas por Tellería-Orriols et al. 2023 anotadas y analizadas en nuestra población de estudio .....	44
<b>Tabla 4.</b> Variantes relevantes en la publicación de Kuehn et al. 2019 anotadas y analizadas en nuestra población de estudio. ...	45

## Agradecimientos

Este trabajo fue realizado en colaboración con la Unidad de Bioinformática y Medicina Personalizada de Laboratorios Larrasa, empresa en la que realicé prácticas extracurriculares durante el año 2024, como parte de mi formación en el Máster Oficial de Bioinformática ofrecido por la Universidad Europea.

Quiero expresar mi sincero agradecimiento a Javier Gómez Román, bioinformático clínico y analista de datos NGS de Laboratorios Larrasa por su acompañamiento y apoyo durante todo el proceso. Su experiencia y disposición para compartir sus conocimientos han sido fundamentales para el desarrollo de este trabajo.



## Resumen

La hipertensión intracraneal idiopática (HII) es una enfermedad rara, con una prevalencia estimada de 1 por cada 10,000 habitantes, notablemente más frecuente en mujeres con sobrepeso en edad reproductiva. Esta condición se caracteriza por un aumento crónico de la presión intracraneal, manifestado a través de síntomas como dolores de cabeza persistentes, deterioro cognitivo leve y alteraciones visuales, siendo la pérdida total de la visión una complicación potencialmente irreversible. Existen diversas hipótesis sobre el origen y la fisiopatología de esta enfermedad, que abarcan desde alteraciones en la homeostasis del líquido cefalorraquídeo hasta desregulaciones metabólicas, hormonales e inmunológicas. Estudios previos han reportado casos de HII en varios miembros de una misma familia, lo que sugiere una posible predisposición genética; sin embargo, hasta el momento son pocos los estudios genéticos que han identificado variantes causales o que predispongan al desarrollo de esta condición. Este trabajo aborda la HII desde una perspectiva integral, aplicando la secuenciación del genoma completo (WGS) en la determinación de variantes genéticas que puedan estar implicadas en su desarrollo.

### Objetivos:

El objetivo principal de este trabajo consiste en identificar, a través del WGS variantes genéticas asociadas con la HII que puedan ser utilizadas en un futuro como potenciales biomarcadores así como para esclarecer la fisiopatología de la enfermedad. Como objetivo secundario, también se buscan variantes genéticas descritas en estudios anteriores para comprobar su posible asociación con la HII.

### Material y métodos:

Se secuenció el genoma completo de 25 pacientes afectados con HII pertenecientes a la Asociación de Enfermos y Familiares con Hipertensión Intracraneal Idiopática (ADEFHIC) y de 25 controles en la plataforma NovaSeq 6000, con una cobertura promedio de 40x. Se diseñó una pipeline de análisis genómico a través de la cual se anotaron y priorizaron las variantes genéticas identificadas en base a su efecto funcional, así como, a su interactoma con la herramienta *hiviPrioritiser* con el objetivo de centrar el análisis estadístico posterior en variantes con un mayor potencial patogénico y clínico en el contexto de la HII.

### Resultados:

Se identificaron más de 90 variantes fuertemente asociadas a la HII en regiones no codificantes y un total de 3 variantes, aunque con una menor fuerza de asociación, en regiones codificantes del genoma. Destaca el gen *ASAH2* como aquel con mayor acumulación de variantes significativas, tanto en regiones codificantes como no codificantes del genoma, así como variantes fuertemente correlacionadas con la HII como rs71583510 (p-value=3.16x10<sup>-6</sup>), rs58930175 (p-value = 1.21x10<sup>-5</sup>) y rs111963051 (p-value=1.89x10<sup>-5</sup>), localizadas en genes involucrados en la regulación de la fosforilación de proteínas y la señalización intracelular.

### Conclusiones:

Este trabajo reafirma la importancia de las regiones no codificantes en el desarrollo y predisposición a diferentes enfermedades, apuntando a un posible componente inmunológico o de inflamación crónica, ocasionado posiblemente por un desequilibrio en el metabolismo lipídico de las ceramidas, como la causa de la HII.

### Palabras clave:

Hipertensión Intracraneal Idiopática. Estudio de asociación genómica. Secuenciación de genoma completo

## Abstract

Idiopathic intracranial hypertension (IIH) is a rare disease with an estimated prevalence of 1 in 10,000 individuals, and it is notably more common in overweight women of reproductive age. This condition is characterized by a chronic increase in intracranial pressure, manifesting through symptoms such as persistent headaches, mild cognitive impairment, and visual disturbances, with total vision loss as a potentially irreversible complication. There are various hypotheses regarding the origin and pathophysiology of this disease, ranging from disruptions in cerebrospinal fluid homeostasis to metabolic, hormonal, and immunological dysregulation. Previous studies have reported cases of IIH among multiple family members, suggesting a possible genetic predisposition. However, few genetic studies have identified causal or predisposing variants for this condition. This study approaches IIH from a comprehensive perspective, using whole genome sequencing (WGS) to identify genetic variants potentially involved in its development.

### Objectives

The primary objective of this study is to identify IIH-associated genetic variants through WGS that could serve as potential biomarkers in the future, as well as to clarify the disease's pathophysiology. A secondary objective is to examine previously described genetic variants to verify their possible association with IIH.

### Methods

The whole genome of 25 IIH patients from the Association of Patients and Families with Idiopathic Intracranial Hypertension (ADEFHIC) and 25 controls was sequenced on the NovaSeq 6000 platform, with an average coverage of 40x. A genomic analysis pipeline was designed to annotate and prioritize identified genetic variants based on their functional impact and interactome with the *hiPhivePrioritiser* tool, focusing subsequent statistical analysis on variants with higher pathogenic and clinical potential in the context of IIH.

### Results

Over 90 variants strongly associated with IIH were identified in non-coding regions, and a total of 3 variants, albeit with lower association strength, in coding regions of the genome. The *ASAH2* gene emerged as having the highest accumulation of significant variants, both in coding and non-coding regions, along with strongly IIH-correlated variants such as rs71583510 ( $p\text{-value}=3.16\times 10^{-6}$ ), rs58930175 ( $p\text{-value}=1.21\times 10^{-5}$ ), and rs111963051 ( $p\text{-value}=1.89\times 10^{-5}$ ), located in genes involved in protein phosphorylation regulation and intracellular signaling.

### Conclusions

This study reaffirms the importance of non-coding regions in the development and predisposition to various diseases. It also suggests a possible underlying immunological or chronic inflammatory component, primarily due to an imbalance in ceramide lipid metabolism as a potential cause of IIH.

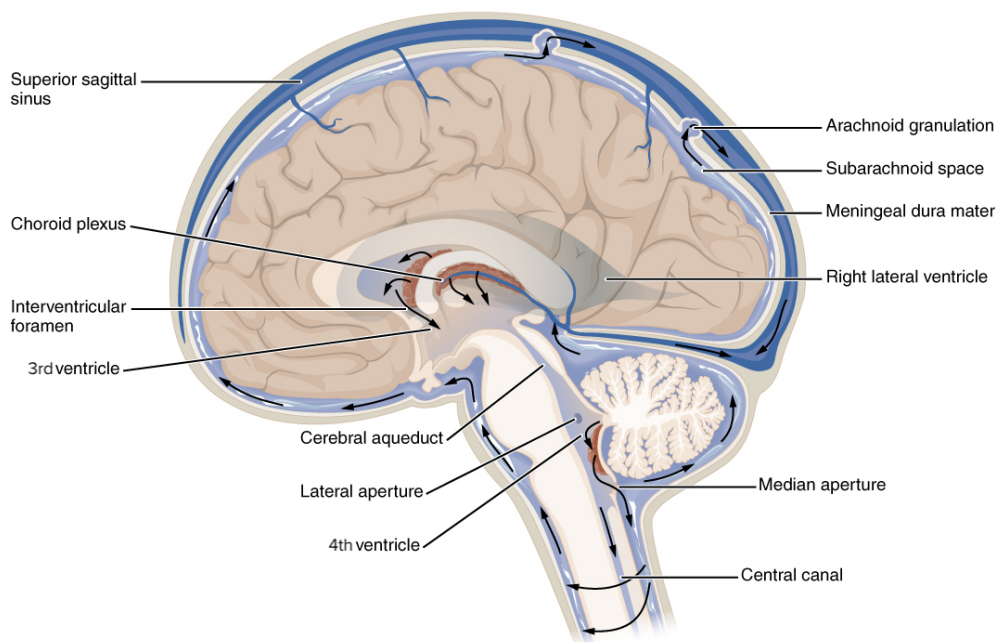
### Key words

Idiopathic Intracranial Hypertension, Genomic Association Study, Whole Genome Sequencing

# 1. Introducción

## 1.1. Contexto

La presión intracraneal (PIC) se define como la presión que existe dentro de la cavidad del cráneo y que se genera por el volumen combinado del tejido cerebral, el líquido cefalorraquídeo (LCR) y el flujo sanguíneo (Imagen 1). Esta presión, que se encuentra típicamente entre 7 y 15 mmHg en adultos sanos, es crucial para la función normal del sistema nervioso central, ya que contribuye a mantener la perfusión adecuada del cerebro, manteniéndolo así oxigenado y facilitando la eliminación de desechos (Steiner & Andrews, 2006)



**Imagen 1. Anatomía del sistema ventricular y flujo del LCR en la cavidad craneal.** El LCR, cuyo principal función es la de proteger el cerebro, es producido por el plexo coroideo en los ventrículos cerebrales. desde donde se distribuye por el espacio subaracnoideo, cavidad que separa la aracnoides de la piamadre, rodeando el cerebro y la médula espinal, para finalmente ser reabsorbido por el sistema venoso a través de las granulaciones aracnoideas que atraviesan la dura madre. En tonos beis esta representado el tejido cerebral, en azul claro el LCR y en color tanto rojo como azul oscuro el flujo sanguíneo.

En condiciones basales, la PIC tiende a mantenerse constante gracias a un delicado equilibrio entre los volúmenes del tejido cerebral, el LCR y el flujo sanguíneo del interior de la cavidad craneal. Este equilibrio se explica gracias a la Ley de Monro-Kellie, que establece que debido a la naturaleza rígida del cráneo el volumen total de los componentes de su interior es fijo; por lo tanto, cualquier aumento en uno de los volúmenes de los componentes anteriormente citados debe ser compensado por una

reducción en otro para mantener la PIC dentro de un rango normal. Sin embargo, si la compensación es insuficiente o no es posible reducir alguno de los volúmenes, generalmente de LCR o sangre, la presión intracraneal inevitablemente aumentará (Mokri, 2001). De esta manera, diferentes condiciones o patologías pueden derivar en un incremento de la PIC:

- Aumento del volumen cerebral fruto de un edema o inflamación a causa de un traumatismo craneal grave o alguna condición metabólica como una insuficiencia renal grave que conduzca a una encefalopatía urémica (Arnold et al., 2016), o la hiperamonemia fruto de una enfermedad hepática crónica como la cirrosis o la hepatitis (Scott et al., 2013)
- Procesos intracraneales expansivos que compriman el tejido cerebral y reduzcan el volumen intracraneal como un tumor cerebral, un absceso o un hematoma ya sea fruto de un traumatismo craneal o una aneurisma (Rodríguez-Boto et al., 2015).
- Alteraciones del flujo sanguíneo cerebral que aumenten el volumen de sangre intracraneal como la trombosis de los senos venosos o presiones venosas centrales elevadas a causa de una insuficiencia cardíaca (Geisbüsch et al., 2021).
- Diferentes enfermedades como por ejemplo la meningitis (Tariq et al., 2017) o la encefalitis (Kumar et al., 2009) que además de poder aumentar el volumen de tejido cerebral a causa de una inflamación generalizada también pueden alterar los procesos de producción y reabsorción de LCR al modificar la composición citoquímica de este o al producir una inflamación de las meninges interfiriendo así con su circulación normal.
- Otras casusas menos comunes abarcarían diversas malformaciones congénitas que producen deformidades craneales o que comprometen de alguna forma los flujos normales de LCR y sangre, como por ejemplo la malformación de Chiari (Santos-Bueso et al., 2015) o la malformación de Dandy-Walker (Spennato et al., 2011) que obstruyen el drenaje del LCR produciendo hidrocefalia, acumulación de LCR dentro de la cavidad craneal, y consecuentemente aumentan la PIC.

Como hemos mencionado en el párrafo anterior, el incremento de la PIC es una condición multifactorial que en la mayoría de los casos suele estar ligada a una causa o patología conocida fácilmente diagnosticable ya sea mediante técnicas de imagen médica, como la tomografía computerizada y la resonancia magnética, o pruebas de laboratorio tales como análisis de sangre o la punción lumbar, un procedimiento médico en el que se inserta una aguja en la parte baja de la columna vertebral, específicamente en el espacio subaracnoideo, para extraer LCR, permitiendo así medir tanto su presión como analizar su composición citoquímica. Sin embargo un pequeño porcentaje de la población sufre de esta condición sin una causa aparente, es entonces cuando hablamos de hipertensión intracraneal idiopática (HII) también conocida históricamente como *pseudotumor cerebri*.

## **1.2. Presentación clínica de la HII**

Las personas que padecen de HII sufren síntomas muy similares a aquellos que padecen de un proceso intracraneal expansivo, como un tumor cerebral, siendo el dolor de cabeza constante el síntoma más habitual seguido de alteraciones visuales tales como oscurecimiento de la visión, diplopia, fotosensibilidad, visión borrosa, escotomas, y en los casos más graves, ceguera transitoria además de dolor o presión en la parte posterior de los globos oculares. (Giuseffi et al., 1991) Estos síntomas visuales y oftalmológicos son debidos al papiledema, una inflamación del nervio óptico, y a la inflamación del sexto nervio craneal, el cual se encarga del movimiento de la musculatura ocular, fenómenos que ocurren de manera muy habitual como respuesta a una PIC elevada (Friedman, 2014)

También se destacan otros síntomas como el tinnitus pulsátil, dolores cervicales, mareos recurrentes en ocasiones acompañados de náuseas y, menos frecuentemente, ya que para que ocurran la PIC elevada debe mantenerse durante largos periodos de tiempo, otorrea o rinorrea, como consecuencia de una erosión ósea de la base del cráneo, capaz de provocar fugas de LCR (Tam & Gilbert, 2019) y alteraciones cognitivas leves, fruto de una perfusión cerebral deficiente, mantenida en el tiempo que suelen verse reflejadas en tiempos de reacción más lentos, pérdidas de memoria, fatiga mental y problemas a la hora de concentrarse (Kleinschmidt et al., 2000).



Para que una persona sea diagnosticada con HII, debe presentar alguno de los signos y síntomas relacionados con el aumento de la PIC mencionados en el párrafo anterior, entre los cuales siempre debe figurar el papiledema. Se requiere que la presión de apertura del LCR, medida mediante punción lumbar, sea igual o superior a 250 mm H<sub>2</sub>O y que su composición no presente anomalías citológicas ni químicas que puedan relacionar este incremento de la PIC con otra patología. Finalmente, es fundamental que los estudios de neuroimagen sean normales, lo que permite descartar otras patologías cerebrales que puedan causar estos síntomas aunque será habitual encontrar hallazgos que reafirmen la PIC elevada como lo son el aplanamiento posterior de los globos oculares, estenosis del seno transversal y el engrosamiento de la vaina del nervio óptico como consecuencia del papiledema (Friedman et al., 2013).

### **1.3. Epidemiología**

La HII es una enfermedad rara con una incidencia aproximada de tan solo un caso por cada cien mil habitantes, que se ha reportado tanto en personas adultas como en infantes, aunque principalmente se focaliza en personas de entre veinte y cincuenta años (Craig et al., 2001), representando el sexo masculino únicamente el 10% de los casos (Bruce et al. 2009). El principal factor de riesgo es la obesidad, con una asociación dosis-respuesta que relaciona directamente el aumento de índice de masa corporal con un mayor riesgo de sufrir HII (Andrews et al., 2014), siendo de esta manera las mujeres jóvenes en edad reproductiva y con sobrepeso el grupo poblacional más afectado por esta enfermedad. Durante la última década otros factores de riesgo como el uso de determinados fármacos del grupo de las tetraciclinas, el uso de anticonceptivos, el embarazo o la apnea obstructiva del sueño también han sido señalados como factores de riesgo para el desarrollo de esta enfermedad aunque sin una evidencia tan clara, al menos hasta la fecha (Bioussé, 2012).

Es debido a su clara relación con el sobrepeso por lo que la prevalencia de la HII está en aumento en los últimos años, especialmente en países occidentales donde el índice de obesidad entre sus habitantes es alto, como Estados Unidos o Reino Unido (Chai et al., 2014), sin embargo cabe destacar que aunque la obesidad es un factor de riesgo

clave, existen casos de HII en personas con un peso normal según el IMC lo que sugiere que otros factores ya sea genéticos, ambientales o una combinación de ambos también juegan un papel importante en el desarrollo de esta enfermedad. Un ejemplo de esto es la población asiática, donde la prevalencia de HII es menor que en poblaciones occidentales, muy posiblemente debido a menores índices de obesidad entre sus integrantes, pero sin embargo, los casos de HII en personas sin sobrepeso son significativamente más comunes (Hsu et al., 2024).

#### **1.4. Pronóstico y tratamiento**

Aunque la HII puede presentarse en forma de episodios agudos en la mayoría de los casos, esta condición tiende a ser crónica y persistente, empeorando significativamente con el tiempo si no es correctamente tratada, produciendo así daños irreversibles y afectando notablemente la calidad de vida de quienes la padecen (Thaller et al., 2023).

Entre estos daños irreversibles destacan la pérdida total de la visión, como consecuencia de lesiones en el nervio óptico (Corbett et al., 1982), así como el deterioro cognitivo leve irreversible (Grech et al., 2021). Además, aunque la HII no está directamente relacionada con la mortalidad de quienes la sufren, los dolores de cabeza y mareos cada vez más intensos dificultan a los afectados realizar actividades cotidianas e incluso dormir. Esto se traduce en tasas más altas de ansiedad y depresión dentro del grupo de afectados lo que, en consecuencia lleva a un mayor riesgo de suicidio (Hermes et al., 2020).

El diagnóstico temprano de la HII es crítico pues no solo previene de daños irreversibles sino que mejora notablemente la calidad de vida de los pacientes, sin embargo y por desgracia, este proceso puede ser difícil y prolongado en el tiempo pues la naturaleza inespecífica de los síntomas de la HII pueden confundirse con otros trastornos neurológicos, además, los pacientes no siempre son conscientes de las alteraciones visuales que sufren, especialmente al inicio, por lo que no pueden reportarlas. Sumado a esto, la confirmación de la HII requiere de una combinación entre estudios de imagen médica, y una punción lumbar que junto con la realización de otras pruebas inespecíficas o repetitivas utilizadas para descartar otras causas del incremento de la PIC pueden retrasar aún más el diagnóstico (Chagot et al., 2017)

Entre los tratamientos actuales frente a la HII encontramos diferentes fármacos como la Acetazolamida y la Furosamida, un inhibidor de la anhidrasa carbónica y un diurético, respectivamente, que reducen eficientemente la producción de LCR y consecuentemente la PIC. Otros fármacos como el Topiramato, un anticonvulsivo, también ha demostrado ser eficaz controlando la HII además de mejorar la función visual en algunos pacientes. La selección del medicamento dependerá de factores como la respuesta inicial del paciente, las comorbilidades y la gravedad de los síntomas. (Piper et al., 2015). En pacientes con obesidad será fundamental también controlar su peso dado que la HII está significativamente asociada con esta condición. Reducciones moderadas de entre el 5-10% del peso han demostrado mejoría en los síntomas de personas afectadas, de esta manera, es recomendable, seguir dietas bajas en sodio y realizar ejercicio físico de manera regular. (Abbott et al., 2023)

En los casos más graves o donde no se observen mejorías con los tratamientos anteriores, podrán valorarse intervenciones quirúrgicas invasivas como la derivación lumboperitoneal, que consiste en la inserción de un tubo hueco en la columna vertebral o el cráneo para drenar el exceso de LCR, aliviando de esta manera los síntomas de la HII (Daou et al., 2020). La fenestración de vaina del nervio óptico, cuya finalidad es la de aliviar la inflamación del nervio al crear una pequeña abertura en la vaina que lo recubre y que está especialmente indicado para aquellos pacientes con deterioros visuales severos o que avanzan con gran rapidez, estabilizando de esta manera los síntomas visuales y evitando la ceguera total (Nithyanandam et al., 2008).

Por último, y de nuevo relacionado con la obesidad, encontraríamos la cirugía bariátrica, un conjunto de procedimientos quirúrgicos diseñados para modificar el sistema digestivo de los pacientes con el objetivo de reducir la cantidad de alimentos que una persona puede consumir, ayudándoles a bajar de peso rápidamente si no ha sido posible con métodos convencionales (Handley et al., 2015)

## 1.5. Patogénesis y etiología de la HII

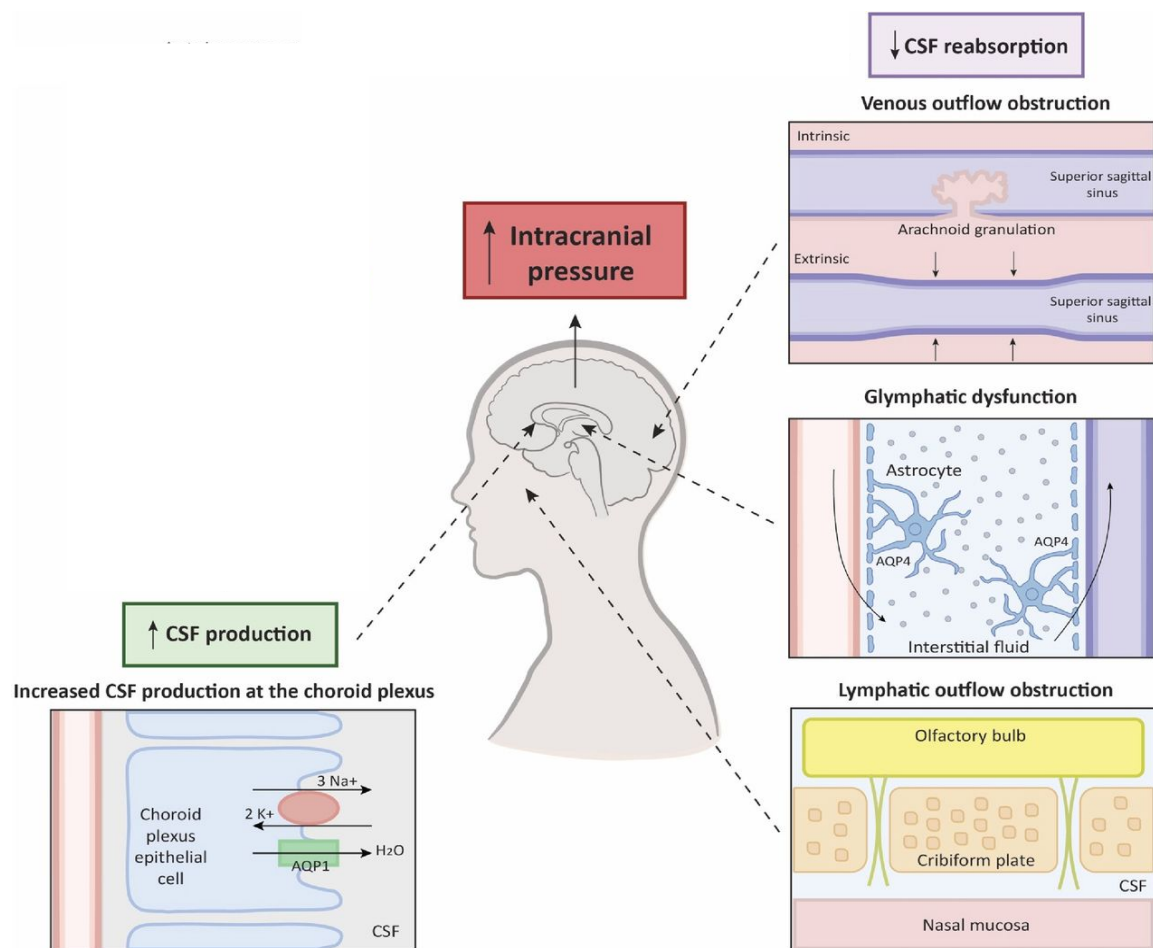
A pesar de contar con tratamientos y criterios diagnósticos bien establecidos, la patogénesis de la HII continúa siendo una incógnita, lo que limita la disponibilidad de dianas terapéuticas efectivas para combatirla y dificulta la rapidez en su diagnóstico.

### 1.5.1. Homeostásis del LCR

Diversos estudios presumen que la etiología implica una alteración en la homeostasis del LCR ya sea por un aumento en la producción de este o deficiencias en su drenaje. La secreción de LCR es constante y la mayor parte tiene lugar en el plexo coroideo, una estructura vascularizada localizada en los ventrículos del cerebro cuyas células epiteliales mueven los iones del plasma sanguíneo hacia el interior del espacio ventricular usando bombas sodio-potasio ATPasa ( $\text{Na}^+/\text{K}^+$ -ATPasa), de esta manera crean un gradiente osmótico que mueven también el agua del plasma sanguíneo hacia el espacio ventricular donde se mezclará con otros solutos secretados por el mismo plexo coroideo convirtiéndose así en LCF el cual fluirá hacia el espacio subaracnoideo y la médula espinal (Pollay et al., 1985). Este movimiento de agua es gracias a la acción de la proteína AQP1 perteneciente a la familia de las aquoporinas (Benarroch, 2016).

Debido a su constante producción el LCR debe también drenarse de forma constante para evitar su acumulación y un aumento excesivo de la PIC. La principal vía de drenaje son las granulaciones aracnoideas, protuberancias de la aracnoides que se extienden hacia el sistema venoso permitiendo así que el LCR fluya lentamente hacia el sistema circulatorio. Recientemente se ha descubierto una segunda vía de drenaje conocida como sistema glial-linfático, la cual está implicada en la eliminación de desechos metabólicos del LCR. Este sistema se basa en el flujo del LCR desde el espacio subaracnoideo a los espacios perivascuales donde gracias a la acción de los astrocitos, unas células gliales que rodean los vasos sanguíneos llenas de prolongaciones conocidas como pies perivascuales, el LCR y diferentes sustancias de desecho son transportados al espacio intersticial donde serán drenados hacia el sistema venoso o el sistema linfático. (Cogollo et al., 2023) Esto es posible gracias a la acción de la proteína AQP4 presente en los pies

perivasculares de los astrocitos la cual facilita el paso de agua y el LCR desde el espacio perivesicular al espacio intersticial (Peng et al., 2023). La proteína AQP4 no solo está involucrada en el sistema glial-linfático sino que desempeña un papel crucial en la regulación hídrica del cerebro, siendo la acuaporina más abundante en el sistema nervioso central. De esta manera, se apunta a anomalías en las vías de drenaje del LCR o alteraciones en el funcionamiento de los canales involucrados en la dinámica de LCR, como las proteínas de la familia de las aquaporinas, como una posible causa de la HII (Imagen 2), aunque todavía sin un resultado prometedor. (Ekizoglu et al., 2012, Stiebel-Kalish et al., 2013). Mas adelante se abordará la posible implicación de la  $\text{Na}^+/\text{K}^+$ -ATPasa.



**Imagen 2. Mecanismos fisiopatológicos propuestos para el incremento de la PIC basados en la alteración del flujo del LCR.** Las hipótesis basadas en un aumento de la producción de LCR se apoyan en la idea de la sobreexpresión o incremento de la actividad de las proteínas AQP1 o  $\text{Na}^+/\text{K}^+$ -ATPasa, mientras que aquellas que defienden que se deben a problemas en la reabsorción del LCR proponen obstrucciones en las vías de drenaje, así como anomalías en el funcionamiento de la proteína AQP4.

### 1.5.2. Desregulación hormonal

Debido a que la prevalencia de la HII es significativamente mayor en mujeres en edad reproductiva algunos estudios hipotetizan que las hormonas femeninas podrían ser un factor causal de la enfermedad sin embargo un estudio realizado en el año 2019 liderado por O'Reilly identificó una firma única de andrógenos en personas con HII que incluía niveles de testosterona significativamente más altos que en individuos con obesidad o síndrome de ovario poliquístico además de identificarse una mayor actividad de la 5alfa-reductasa, una enzima involucrada en el metabolismo de andrógenos y glucocorticoides.

Se ha demostrado que el tratamiento crónico con testosterona en ratas incrementa su tasa de secreción de LCR ya que el plexo coroideo presenta receptores de andrógenos lo que respaldaría el rol de los andrógenos en la patogénesis de la HII al modular la dinámica del LCR (Wardman et al., 2023) (Imagen 3). En consonancia con esto cada vez hay más reportes de personas transgénero en terapia con testosterona que reportar una presión intracraneal elevada (Hornby et al., 2017). En contraste, se ha observado que hombres con deficiencia de testosterona son más propensos a desarrollar HII (Valcamonico et al., 2014). De esta manera, se hipotetiza que los niveles anormales de andrógenos, más específicamente, concentraciones de testosterona similares a la de hombres con deficiencia de andrógenos y mujeres con exceso de andrógenos, podrían aumentar el riesgo de desarrollar HII (Hornby et al., 2018)

### 1.5.3. Perfil metabólico

Recientemente un estudio examinó el metaboloma de personas con HII donde se pudo observar que las concentraciones de formilpiruvato, acilpiruvato y malepiruvato fueron significativamente más bajas en el LCR y más altas en el plasma sanguíneo de los individuos diagnosticados con HII frente a al grupo control. Dichos niveles se normalizaron tras una intervención terapéutica de pérdida de peso. Además a través de un análisis de enriquecimiento de rutas metabólicas se identificaron perturbaciones en las vías de metabolismo de lípidos así como de algunos aminoácidos como la arginina y la prolina (Alimajstorovic et al., 2023). La

investigación adicional en un futuro de estas rutas metabólicas podría aportar más información sobre la etiología de la enfermedad y el rol de la obesidad (Imagen 3).

Otro estudio realizado por el mismo grupo investigó los efectos metabólicos de dos tipos diferentes de intervenciones bariátricas, la gastrectomía en manga y el bypass gástrico de Roux, en pacientes con HII. Los pacientes que se sometieron a un bypass gástrico de Roux experimentaron una reducción de la PIC un 50% mayor en comparación con las que se sometieron a una gastrectomía en manga a pesar de una pérdida de peso similar en ambos grupos. El grupo sometido al bypass gástrico de Roux mostró un aumento en la secreción de péptido similar al glucacón-1 o GLP1 tras ingerir alimentos. Se sugirió que este aumento en la secreción de GLP1 podría impulsar la reducción de la PIC (Alimajstorovic et al., 2023). Esto es respaldado por la observación de receptores de GLP-1 en el plexo coroideo de roedores a los cuales, tras administrarles exendina-4, un agonista del receptor GLP-1 se observó una reducción rápida y sostenida de la PIC durante el tratamiento, presumiblemente mediante la inhibición de la actividad de la  $\text{Na}^+/\text{K}^+$ -ATPasa (Botfield et al., 2017). Así mismo, la pérdida de peso en los pacientes del estudio se correlacionó con cambios en el metabolismo de lípidos concretamente de ceramidas, glicerofosfolípidos, y lisoglicerofosfolípidos que también podrían estar involucrados en la regulación de la PIC en personas con HII.

#### **1.5.4. Mediadores pro-inflamatorios y sistema inmunitario**

Estudios han sugerido que el tejido adiposo, más abundante en personas con sobrepeso, no solo actúa como un depósito de energía sino que también secreta diversas adipocinas y citocinas pro-inflamatorias que pueden contribuir a una inflamación sistémica leve (Fain, 2006). Debido a esta premisa y a estudios anteriores que demostraron niveles alterados de citocinas y adipocinas en pacientes con HII (Ball et al., 2009; Edwards et al., 2013) un grupo de investigación en 2017 liderado por Samanci B. estudió el rol de varios mediadores inflamatorios en pacientes con HII y en un grupo control y observaron que los pacientes diagnosticados de HII presentaban niveles significativamente más altos de interleucina (IL)  $-1\beta$  y niveles más bajos de IL-8 y factor de necrosis tumoral alfa



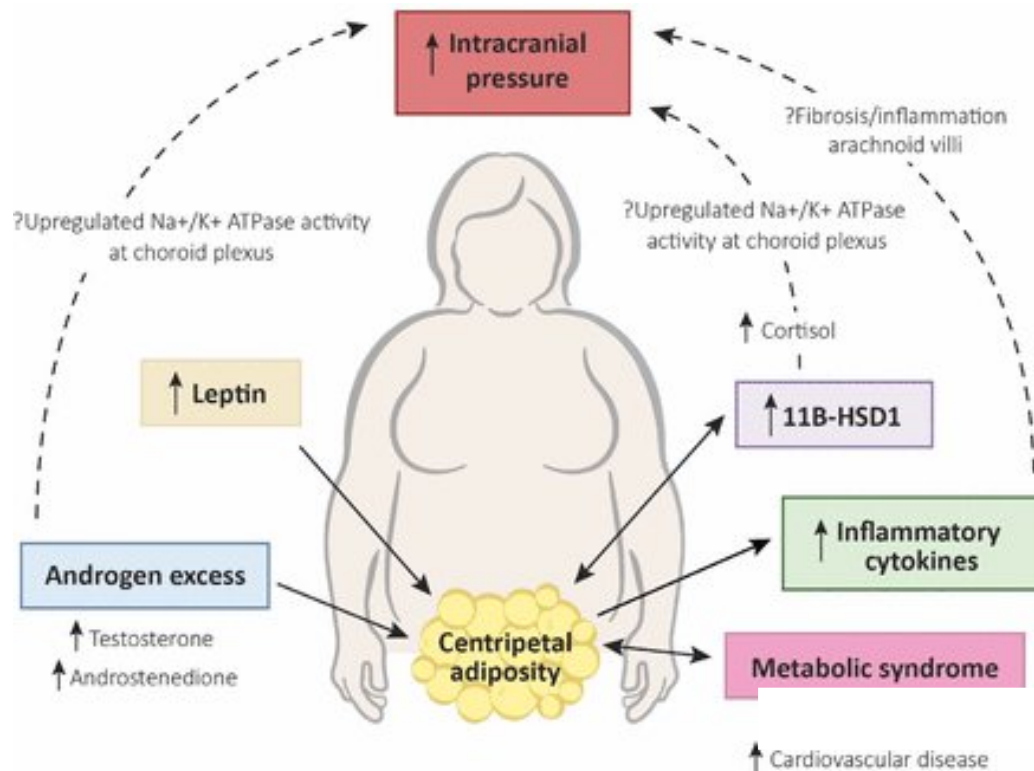
(TNF- $\alpha$ ) en comparación con los controles, además no se encontró una correlación significativa entre esos mediadores inflamatorios y otros factores como la edad o el índice de masa corporal, por lo tanto esta falta de correlación con la adiposidad indica que estas citocinas no están directamente mediados por el tejido adiposo sino por mecanismos inmunológicos intrínsecos.

Estas citocinas proinflamatorias podrían contribuir a una neuroinflamación crónica leve lo que alteraría la función de la barrera hematoencefálica o las meninges, alterando la homeostasis y finalmente conduciendo a disfunciones en el sistema nervioso central (Sawada et al., 1995).

Otros estudios como el de Sincalir et. Al 2008 plantean la hipótesis de que concentraciones elevadas de citocinas proinflamatorias como IL-1 $\beta$ , TNF- $\alpha$  e IL-6 pueden aumentar la actividad de la enzima 11 $\beta$ -hidroxiesteroide deshidrogenasa tipo 1 (11 $\beta$ -HSD1) la cual regula la conversión de cortisona a cortisol en tejidos específicos como el tejido adiposo y el plexo coroideo (Tomlinson et al., 2001, 2004). Un exceso de cortisol generado localmente en el plexo coroideo, debido a una mayor actividad de 11 $\beta$ -HSD1, podría estimular la producción de LCR, contribuyendo así al incremento en la PIC. Este aumento ocurriría porque el cortisol intensifica la actividad de transportadores iónicos, como la Na<sup>+</sup>/K<sup>+</sup>-ATPasa en las células del plexo coroideo, lo cual genera un gradiente osmótico mucho mayor que facilitaría el paso de agua hacia los ventrículos cerebrales, aumentando así el volumen de LCR (Markey et al., 2016). Además este mecanismo podría desencadenar un ciclo de retroalimentación negativa, ya que el cortisol inducido por las citocinas proinflamatorias también modula diferentes respuestas inflamatorias, promoviendo la inflamación continua en el tejido cerebral. Esta inflamación sostenida afectaría la dinámica del LCR, contribuyendo así al aumento crónico de la PIC (Sorrells & Sapolsky, 2007) (Imagen 3).

Por último destacar que el lupus eritematoso sistémico, una enfermedad autoinmune, se han asociado con el desarrollo de HII afianzando aún más la hipótesis de que las alteraciones inmunológicas y la inflamación sistémica podrían contribuir en la fisiopatología de la enfermedad (Green et al., 1995, Maloney, 2013).





**Imagen 3. Mecanismos fisiopatológicos propuestos para el incremento de la PIC causados o agravados por la obesidad.** Las mayor parte de las hipótesis que relacionan la HII con la obesidad apuntando a un aumento en la actividad de la Na<sup>+</sup>/K<sup>-</sup>-ATPasa bajo determinadas circunstancias que que son causadas o exacerbadadas debido a una acumulación excesiva de tejido adiposo

## 1.6. Estudios anteriores que relacionan la HII con variantes genéticas

Aunque desde el siglo pasado se tiene constancia de que la HII muy probablemente tenga una base genética, debido a la aparición de esta enfermedad dentro de un mismo núcleo familiar con relativa frecuencia, especialmente en hermanos gemelos (Buchheit et al., 1969, Traviesa et al., 1976, Johnston & Morgan, 1991) actualmente son pocos los estudios estudios que se han realizados con el objetivo de determinar la causa o predisposición genética de esta enfermedad.

En el año 2013 un estudio realizado por Kerty et al. secuenció el gen de AQP4 de 26 individuos con HII en busca de variantes genéticas asociadas a con la enfermedad debido a la implicación de esta proteína con la homeostásis del LCR y la regulación de hídrica en cerebro. Se encontraron un total de 24 variantes genéticas diferentes con las cuales se realizó un estudio de asociación genética comparandolas con un grupo control de 52 personas sanas, sin embargo tras realizar los análisis estadísticos pertinentes, ninguna pareció estar asociada con el desarrollo de la enfermedad.

Recientemente la Universidad de Valladolid publicó una tesis doctoral realizada por (Tellería-Oriols et al., 2023) en la que con el objetivo de encontrar una asociación genética con la HII se estudiaron cinco polimorfismo de un solo nucleótido (SNPs) localizados en los genes AQP1 y AQP4 que otros estudios previos habían demostraron su asociación con otras patologías, validando así su consecuencia funcional. Tras genotipar 72 pacientes afectados y comparar tanto las frecuencias alélicas como genotípicas con un grupo control de personas sanas obtenido de la base de datos de *Ensembl* se determinó que tan solo uno de los SNPs, rs3763040, localizado en el gen AQP4 y que había sido previamente relacionado con un avance más rápido en el deterioro cognitivo causado por el Alzheimer, podría estar correlacionado con el desarrollo de la HII en población caucásica española.

Por último, y dejando de lado las proteínas de la familia de las acuaporinas, en 2019 un equipo liderado por Kuehn et al. realizó el primer estudio GWAS con personas diagnosticadas con HII. Participaron 95 afectados y 95 controles, todos de sexo femenino pero de diferentes etnias, aunque mayoritariamente de origen caucásico, cuyas muestras fueron genotipadas usando el chip de microarrays *Infinium HumanCoreExome v1-0 BeadChips* de Illumina que permitía la evaluación de hasta 538,448 variantes genéticas con cierta relevancia en investigación genética y, distribuidas por todo el genoma. El limitado número de participantes, para ser un estudio de GWAS, no permitió encontrar asociaciones extremadamente fuertes entre la HII y las variantes genéticas analizadas aunque cabe destacar que se detectaron loci en los cromosomas 5, 13 y 14 altamente sugestivos debido a la gran cantidad de polimorfismos estadísticamente significativos que contenían, destacando genes como LINC00359 y FOX3N, relacionados con los ARN no codificantes de cadena larga comunmente involucrados en la regulación de la expresión génica (Kung et al., 2013; Li et al., 2017), aunque se recalca que los resultados obtenidos deben ser estudiados con mayor profundidad ya que al no haber podido encontrar un solo gen o loci asociado con la HII todo apunta a que podría ser una enfermedad causada por mutaciones en varios genes o que diferentes alelos de un mismo gen puedan contribuir a su aparición. Además debido al uso de un chip de microarrays un porcentaje alto del genoma no ha sido estudiado pudiendo perderse un gran número de variantes de interés.

## 1.7. La secuenciación de genoma completo (WGS) en la práctica clínica

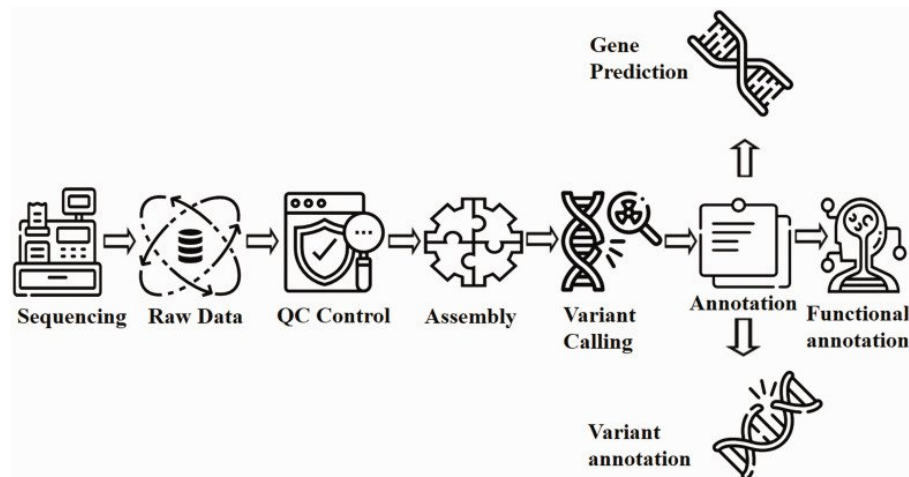
La secuenciación de genoma completo (WGS, por sus siglas en inglés) ha revolucionado el campo de la medicina, proporcionando una herramienta esencial para la identificación de variantes genéticas asociadas con enfermedades, tanto monogénicas como poligénicas, lo que ha mejorado significativamente el diagnóstico clínico (Splinter et al., 2018).

La historia de la WGS se remonta a los primeros años del siglo XXI, impulsada por el Proyecto del Genoma Humano, que culminó en 2003 con la primera secuenciación completa del genoma humano. Sin embargo, el avance decisivo ocurrió en 2008 con la introducción de tecnologías de secuenciación de próxima generación (NGS), que no solo hicieron que la WGS fuera más accesible económicamente, sino también más rápida y eficiente en su ejecución. Este cambio permitió que la WGS se integrara en la práctica clínica, abriendo nuevas posibilidades para el diagnóstico de enfermedades genéticas. Un hito significativo tuvo lugar en 2010, cuando se utilizó la WGS por primera vez con éxito en un entorno clínico al identificar variantes genéticas asociadas a un mayor riesgo de padecer enfermedades cardíacas (Ashley et al., 2010)

Además de la posibilidad de detectar variantes en regiones no codificantes, una de las principales ventajas de la WGS es su capacidad de actuar como un recurso a largo plazo para los pacientes ya que los datos generados pueden reanalizarse y reinterpretarse con el tiempo, lo que resulta especialmente valioso en enfermedades cuyas manifestaciones clínicas pueden evolucionar. (Nisar et al., 2021)

La diferencia fundamental entre la WGS y otros tipos de análisis de NGS radica en la cantidad de datos generados, mientras que los análisis de paneles y exomas producen aproximadamente entre 0.15 - 5 GB de datos en bruto, un análisis de WGS genera alrededor de 30 GB. El análisis bioinformático de estos datos es un proceso intensivo que requiere una considerable cantidad de recursos computacionales (Grealey et al., 2022), representando un desafío en términos de infraestructura y costos. Para abordar el procesamiento de estos grandes volúmenes de datos, se han desarrollado flujos de trabajo estandarizados (Imagen 4), destacando el uso de la herramienta de análisis genómico de código abierto Genome Analysis Toolkit (GATK), que se ha convertido en

la base de muchos centros clínicos y académicos (Van der Auwera et al., 2013) o pipeline managers, como Nextflow, que permiten diseñar, ejecutar y automatizar pipelines de análisis de datos complejos (Di Tommaso et al., 2017). Asimismo, existen soluciones comerciales optimizadas por hardware, como DRAGEN™, que buscan acelerar el análisis de datos de WGS (Behera et al., 2024). El uso de clústeres de supercomputación y el análisis en paralelo también ha resultado esencial para garantizar un análisis rápido y eficiente (Zhou et al., 2024).



**Imagen 4. Diagrama de flujo de una pipeline bioinformática general de análisis de WGS.** Los datos crudos generados a partir del proceso de secuenciación se someten a un exhaustivo proceso de control de calidad y filtrado. Posteriormente, las lecturas filtradas se ensamblan mediante ensamblaje comparativo usando un genoma de referencia para formar secuencias contiguas (contigs). Los contigs se conectan mediante "scaffolding" para obtener ensamblajes preliminares. A continuación, el genoma ensamblado se examina para detectar variantes y se anota con el fin de identificar su localización en los genes, determinar su función y cuantificar su impacto en las proteínas resultantes.

El diagnóstico de enfermedades genéticas raras presenta desafíos únicos, dado que estas condiciones a menudo no se identifican fácilmente a través de fenotipos clínicos. Por ello, es fundamental determinar la mutación exacta que causa la enfermedad. Aunque la secuenciación dirigida puede proporcionar resultados rápidos al centrarse en variantes patogénicas conocidas, no puede detectar la heterogeneidad genética ni identificar nuevas causas o mecanismos de la enfermedad. (Belkadi et al., 2015). La secuenciación de exoma completo (WES) se centra en la región codificante del genoma, que representa aproximadamente el 2%. Si bien ha demostrado ser eficaz en la identificación de causas genéticas en trastornos neurológicos (Soden et al., 2014) y del espectro autista, (Tammimies et al., 2015) presenta limitaciones significativas al centrarse exclusivamente en el exoma ya que no puede detectar variaciones patogénicas en el 98% del genoma restante.

A pesar de los desafíos que enfrenta la WGS, como el procesamiento de grandes volúmenes de datos y la necesidad de una infraestructura computacional robusta, esta se perfila como una herramienta crucial en el diagnóstico de enfermedades raras de elevada complejidad que no solo beneficia a los pacientes al proporcionarles información genética detallada, —o nos permiten conocer sino que también les ofrece opciones de tratamiento más personalizadas y efectiva (Bagger et al., 2024)

## 2. Objetivos

El objetivo principal sobre el que se construye este trabajo es identificar, a través de la secuenciación de genoma completo, nuevas variantes genéticas en población europea asociadas con la HII que puedan ser utilizadas como potenciales biomarcadores diagnósticos o de riesgo, así como para esclarecer la fisiopatología de la enfermedad.

De forma secundaria y paralelamente al principal objetivo también se busca:

- Validar en nuestra población de estudio los resultados obtenidos por otros artículos científicos que relacionan variantes genéticas con la HII.
- Desarrollar una pipeline bioinformática de software libre y escalable a un mayor número de muestras que nos permita obtener una serie de variantes genéticas candidatas partir de archivos FASTQ.

## 3. Material y métodos

### 3.1. Población del estudio

Se procedió a un estudio descriptivo prospectivo de pacientes afectados de HII pertenecientes a la Asociación de Enfermos y Familiares con Hipertensión Intracraneal Idiopática (ADEFHIC) entre 2023 y 2024 donde se seleccionaron un total de 25 individuos. Todos los participantes del grupo de afectados, fueron caucásicos de nacionalidad española y las edades estuvieron comprendidas entre los 20 y 62 años con una edad media de 41 años, de entre los cuales 23 fueron mujeres y 2 fueron hombres Como criterios de inclusión se tomaron en cuenta la mayoría de edad, el diagnóstico de HII por medio de punción lumbar y la realización de estudios de neuroimagen que descartasen cualquier otra patología cerebral responsable del

incremento de PIC y finalmente la firma del consentimiento informado para la autorización del estudio. Como criterios de exclusión se tuvo en cuenta la minoría de edad, la ausencia de un diagnóstico claro de HII, la revocación del consentimiento informado y la presencia de familiares directos (padre, madre, hijo o hermano) entre el resto de participantes del estudio, tanto en el grupo de afectados como en el grupo control, que será detallado más adelante . Este último criterio busca minimizar la consanguinidad dentro del estudio, dado el tamaño reducido de la muestra, para así evitar la sobrerrepresentación de variantes genéticas que no estén directamente relacionadas con la HII y puedan generar ruido.

Como grupo control se seleccionaron otras 25 personas cuyos genomas habían sido secuenciados en Laboratorios Larrasa por iniciativa propia, ya que deseaban acceder a sus servicios de medicina personalizada. A estos individuos se les informó de la posibilidad de incluir sus datos genéticos en el estudio de forma totalmente altruista y anónima y dieron su consentimiento informado para este propósito. Todos los participantes del grupo de afectados, fueron caucásicos de nacionalidad española y las edades estuvieron comprendidas entre los 18 y 75 años con una edad media de años, de 38 años entre los cuales 17 fueron mujeres y 8 fueron hombres Los criterios de inclusión para el grupo control fueron la mayoría de edad, pertenecer a la etnia caucásica, no presentar síntomas de incremento de la PIC ni haber sido diagnosticados de HII, no tener ninguna relación de parentesco directo con ningún participante del estudio y la firma del consentimiento informado para la autorización del estudio.

### **3.2.Aspectos éticos y legales**

Todas las muestras fueron obtenidas previa firma de consentimiento informado conforme a la legislación vigente en materia de protección de datos (Ley Orgánica 3/2018, de 5 de diciembre, de protección de datos con carácter personal y garantía de los derechos digitales), la Declaración de Helsinki de la Asociación Médica Mundial (AMM) y los principios éticos para las investigaciones médicas en seres humanos (última enmienda en octubre de 2013) y de estudios clínicos en España. (Anexo 1)



A fecha, 7 de noviembre de 2024 el estudio fue presentado ante la Comisión de Investigación de la Escuela de Doctorado e Investigación de la Universidad Europea y se encuentra en trámites de ser aprobado y autorizado por el Comité Ético de Investigación Clínica del Hospital Clínico Universitario de Getafe. (Anexo 2)

### 3.3. Obtención de los datos

Para la secuenciación de genoma completo de los participantes, se utilizó saliva recolectada mediante el tubo Oragene 600, la extracción de ADN se llevó a cabo usando el Kit Quick-DNA™ Magbead Plus. Cada muestra de ADN fue cuantificada y ajustada a una concentración de entre 3,3 y 16,6 ng/μL para cumplir con los requisitos de calidad en los pasos posteriores. La preparación de librerías se realizó mediante el Kit Illumina PCR-Free DNA Prep en un sistema de pipeteo automatizado Fluent, siguiendo cuidadosamente las especificaciones del fabricante. Posteriormente, la secuenciación de genoma completo se efectuó con el kit NovaSeq 6000 Reagent Kit en la plataforma NovaSeq 6000, con una cobertura promedio de 40x, garantizando una alta precisión y profundidad de lectura. Para asegurar la calidad de los datos, se aplicó un control de calidad mediante FASTQC, seguido de un preprocesamiento de los archivos FASTQ. Este preprocesamiento fue realizado por empleados de Laboratorios Larrasa y conforme a los parámetros de calidad establecidos por el laboratorio, asegurando así la integridad y confiabilidad de los resultados finales (Imagen 5).



**Imagen 5. Representación del flujo de trabajo efectuado para la obtención del genoma completo del grupo de afectados por HII.** La realización de todos estos procedimientos, desde el contacto con ADEFHIC hasta el filtrado de calidad de los reads arrojados por el secuenciador, en su mayoría de wet lab, fueron llevados a cabo por personal de Laboratorios Larrasa entre noviembre de 2023 y julio de 2024 por lo que los archivos FASTQ de buena calidad fueron el punto de partida de este trabajo.

Los datos de secuenciación se obtuvieron en tres ciclos de secuenciación diferentes realizados entre noviembre de 2023 y julio de 2024 en las que se incluyeron tanto muestras como controles. Por último destacar que, aunque no se utilizaron en este trabajo debido a cuestiones de tamaño poblacional y potencia estadística, también se recogieron datos clínicos y descriptivos de los participantes en el estudio tales como la edad, el sexo o el índice de masa corporal (IMC) con el objetivo de observar diferencias entre grupos poblacionales de interés. Así pues, estos datos nunca fueron revelados ni relacionados con ninguna de las muestras empleadas en el estudio aunque podrán utilizarse en un futuro si el tamaño de la poblacional es aumentado considerablemente.

### 3.4. Software y bases de datos empleadas

Para poder ejecutar el análisis de datos genómicos, que será detallado más adelante en, así como para llevar a cabo tanto el análisis estadístico de los resultados como la visualización de los mismos se hicieron uso de diferentes programas de código abierto y bases de datos de libre acceso.

#### 3.4.1. Software

Todas las herramientas software empleadas en este trabajo, a excepción de una, que es totalmente prescindible y fue utilizada con el único objetivo de agilizar el proceso en términos de tiempo, son de código abierto y están disponibles de manera gratuita facilitando la reproducibilidad de este trabajo.

- FASTQC. Herramienta de software diseñada para realizar controles de calidad en datos de secuenciación de alto rendimiento. Fue empleada para verificar la buena calidad de los archivos FASTQ de los que partía el análisis.
- Samtools. Conjunto de herramientas diseñado para manipular y analizar archivos de alineamiento siendo fundamental en el análisis de datos de secuenciación. En este trabajo Samtools fue utilizado específicamente para indexar el genoma de referencia mediante la herramienta *samtools faidx*, proceso esencial para asegurar un acceso eficiente a las posiciones del genoma de referencia durante los análisis que hacen uso de este archivo.
- BWA (Burrows-Wheeler Aligner). Herramienta de software diseñada para mapear, de forma rápida y precisa, lecturas cortas contra una secuencia de



referencia, en nuestro caso, el genoma de referencia humano hg38. Se empleó el algoritmo BWA-MEM (Maximal Exact Matches) que tolera pequeñas diferencias entre la lectura y la secuencia de referencia contra la que se mapea siendo adecuado para la detección de SNPs y pequeñas inserciones o deleciones.

- GATK (Genome Analysis Toolkit). Conjunto de herramientas de software desarrollado por Broad Institute y Harvard, diseñado para analizar grandes volúmenes de datos genómicos, optimizar la calidad de los mismos y facilitar el proceso conocido como llamada de variantes o variant calling que consiste en identificar variantes genéticas en un genoma mediante la comparación de secuencias de ADN de una muestra con una secuencia de referencia (Van der Auwera et al., 2013). En este trabajo fue empleada la última versión hasta la fecha (GATK 4.3.0.0) lanzada en mayo de 2023 y se utilizó para marcar lecturas duplicadas que puedan introducir sesgos (MarkDuplicatesSpark), recalibrar los valores de calidad de las bases para mejorar la precisión de los datos del alineamiento efectuado con BWA (BaseRecalibrator y ApplyBQSR) y para efectuar el variant calling de variantes germinales (HaplotypeCaller).
- Exomiser. Software de anotación y priorización de variantes genómicas basado en Java y desarrollado por Wellcome Trust Sanger Institute, diseñado para identificar variantes genéticas potencialmente patogénicas asociadas con enfermedades. Exomiser integra información de múltiples bases de datos, lo que incluye datos sobre variantes genómicas, información fenotípica, frecuencia poblacional de variantes y estudios previos sobre genes asociados a enfermedades. Además utiliza ontologías estandarizadas para clasificar y comparar fenotipos. Dentro de Exomiser destaca *hiPhivePrioritiser*, una herramienta fundamental para la priorización de variantes, esta utiliza datos de fenotipos tanto de seres humanos como de otras especies, como ratón, para establecer similitudes entre manifestaciones de enfermedades en diferentes modelos biológicos ayudando a identificar genes que puedan estar relacionados con condiciones patológicas específicas (Robinson et al., 2014). Además, no considera solo el gen en cuestión sino que también toma en cuenta genes

cercanos y su interactoma (Smedley et al., 2014), permitiendo modelar la influencia que la disrupción de un gen tiene en el contexto de toda la red de interacciones genéticas, facilitando así la identificación de variantes que pueden ser relevantes en enfermedades complejas. Esta herramienta asigna una puntuación a los genes basándose en los análisis y comparaciones realizadas permitiendo así priorizar aquellos que tienen una mayor probabilidad de estar involucrados en las enfermedades estudiadas (Zemojtel et al., 2014). Para este trabajo se usó la última versión de Exomiser (Exomiser 14.0.0) lanzada en febrero de 2024.

- Nextflow. Herramienta software de gestión de flujos de trabajo de código abierto que permite diseñar y ejecutar pipelines de análisis de datos complejos de forma eficiente y reproducible mediante la automatización y paralelización de los diferentes procesos implicados. Está diseñado para ser altamente flexible permitiendo definir flujos de trabajo en un lenguaje de scripting simple, basado en Groovy, lo que facilita la modificación y reutilización de pipelines sin tener que realizar cambios significativos donde cada etapa involucra el uso de diferentes herramientas software. Además puede ser ejecutado tanto en servidores locales como en plataformas en la nube o clústeres de supercomputación. (Rashid et al., 2024) En este trabajo fue empleado para diseñar y automatizar el flujo de trabajo de los datos genómicos analizados con el fin de crear una pipeline de análisis de código abierto que pueda ser empleada en el futuro.
- DRAGEN germline. Herramienta software desarrollada por la empresa Illumina, la única de software propietario en este trabajo y por tanto no disponible de forma gratuita, dirigida al análisis de datos de secuenciación genómica en el contexto de la genética humana e investigación médica (Behera et al., 2024). Esta herramienta está optimizada para la realización de filtrado de calidad, mapeo de lecturas, variant calling de variantes germinales, tanto de genoma completo como de paneles de genes y anotación de variantes de forma rápida y eficiente gracias al uso de algoritmos de análisis avanzados, machine learning y hardware especializado permitiendo así obtener resultados de alta calidad y precisión en un tiempo mucho menor en comparación con otras herramientas (Alganmi &

Abusamra, 2023). DRAGEN germline fue usado en este trabajo para analizar parcialmente algunas de las muestras empleadas, con el único objetivo de reducir los tiempos de análisis de determinados procesos como el mapeo de secuencias y el variant calling, que con herramientas de software libre como BWA y GATK eran considerablemente más altos.

Destacar también que además de las herramientas anteriormente enumeradas se hizo uso de RStudio para el procesamiento y análisis estadístico de los resultados arrojados por Exomiser así como para la visualización de dichos resultados haciendo uso de librerías como *data.table*, diseñada para la manipulación eficientes y análisis de grandes volúmenes de datos, *ggplot2*, diseñada para la generación de gráficos, o *clusterProfiler* para la realización de análisis de enriquecimiento de genes. También se hizo uso de Python para la visualización de resultados y generación de gráficos haciendo uso de la librería *Matplotlib*

### 3.4.2. Bases de datos

Muchas de las herramientas software empleadas hacen uso de bases de datos o archivos de referencia para poder ejecutar los análisis. Todas estas bases de datos son de libre acceso y pueden ser descargadas en un servidor local de forma totalmente gratuita para hacer uso de ellas.

- Archivos de referencia genómica y dbSNP. Secuencia de referencia del genoma humano, versión hg38, en formato FASTA utilizada para el alineamiento de lecturas durante el proceso de mapeo con BWA-MEM, y base de datos de variantes conocidas en humanos dbSNP (Single Nucleotide Polymorphism Database) que contiene una colección extensa de variantes genéticas que se han documentado en la literatura científica, fue empleado durante el proceso de recalibración de calidad de las lecturas realizado por GATK y el proceso de anotación de variantes llevado a cabo por Exomiser.
- REMM. Base de datos que evalúa el impacto patogénico de variantes en regiones no codificantes, principalmente reguladoras, del genoma proporcionando puntuaciones en relación con su posible impacto funcional. Fue empleada por Exomiser durante el proceso de anotación de variantes

- CADD data. Base de datos específica necesaria para implementar la herramienta CADD en Exomiser la cual viene integrada en este y proporciona puntuajes que indican la probabilidad de que una variante tenga un efecto perjudicial en la función génica. Contiene información sobre todas las variantes SNVs posibles en el genoma hg38, así como sobre 80 millones de variantes InDels, permitiendo un análisis detallado de todas las variantes del genoma.
- Base de datos de información fenotípica estandarizada. Base de datos creada por Monarch Initiative que clasifica términos fenotípicos asociados a enfermedades humanas. Facilita la comparación entre el perfil fenotípico del paciente y las características de diferentes enfermedades para mejorar la priorización de variantes. Empleada por Exomiser, más concretamente *hiPhivePrioritiser* durante el proceso de priorización de variantes
- gnomAD. Base de datos de frecuencias alélicas integrada en Exomiser basada en el análisis de grandes cohortes de individuos de diferentes poblaciones. Utilizada para filtrar variantes comunes y detectar variantes raras con mayor probabilidad de ser patogénicas. Se emplearon los datos referentes a la cohorte caucásicas no finlandesas ya que todos los participantes del estudio pertenecían dicho grupo poblacional.

### 3.4.3. Instalación y puesta a punto

Se creó un entorno de trabajo utilizando el gestor de paquetes y entornos MiniConda para gestionar todas las dependencias necesarias para llevar a cabo el análisis (ver Anexo 3). Todas las herramientas software, a excepción de Exomiser, fueron descargados a través de MiniConda desde los canales bioconda y conda-forge. Exomiser, por su parte, fue instalado y descargado siguiendo las instrucciones de la documentación oficial disponible en su página web, junto con las bases de datos necesarias para llevar a cabo el proceso de anotación y priorización de variantes (Wellcome Trust Sanger Institute, n.d) Todas las bases de datos y archivos de referencia, a excepción de gnomAD que se encontraba integrada en Exomiser, fueron descargadas (Anexo 4) y almacenadas en un disco duro externo con una capacidad de 4 terabytes, desde el cual se realizaron los análisis y se almacenaron los archivos generados durante el proceso debido al gran tamaño de estos

### 3.5. Estructura de la pipeline bioinformática

El flujo de trabajo utilizado en el análisis genómico desarrollado en este estudio, y que será detallado a continuación, se implementó en la plataforma de código abierto Nextflow con el objetivo de automatizar y garantizar su reproducibilidad. El script completo de la pipeline está disponible en un repositorio público en GitHub (Anexo 5) Además, se dispone de un script específico para la automatización del proceso de anotación y priorización de variantes con Exomiser (Anexo 6), empleado en muestras previamente procesadas y analizadas con, DRAGEN se encargó de tareas como el mapeo, marcado de duplicados, recalibración de la calidad y variant calling, logrando una reducción significativa en el tiempo de obtención de resultados. Previamente, se verificó que los resultados obtenidos con DRAGEN fueran muy similares a los generados con herramientas de software libre, con la ventaja de un menor tiempo de procesamiento.

#### 3.5.1. Control de calidad

Referido en el script de Nextflow como *process RunFastQC*. Se realizó un control de calidad de las cincuenta muestras empleadas en el estudio, representadas por cien archivos FASTQ, con un par de archivos por cada muestra. El objetivo fue verificar su buena calidad utilizando la herramienta FastQC, que genera un informe en formato HTML donde se pueden consultar diferentes análisis y métricas de calidad. Todas las muestras analizadas presentaron alta calidad y, por lo tanto, fueron utilizadas en los análisis subsecuentes.

Se consideraron parámetros de calidad básicos (imagen 6), como la longitud de las secuencias, que variaron entre 32 y 157 pb, la calidad media por base (imagen 7), y la puntuación de calidad promedio por secuencia, que en todos los casos mostró un valor máximo en el Phred Score superior a 30 (Formula 1), lo que corresponde a una precisión superior al 99.9% de que la base en la secuencia esta correctamente identificada (Imagen 7). Además, se evaluó el contenido de bases indeterminadas, que fue cero en todos los casos, así como el porcentaje de citosina y guanina para descartar la presencia de cualquier tipo de contaminación o artefacto.

$$Q = -10 \log_{10}(P)$$

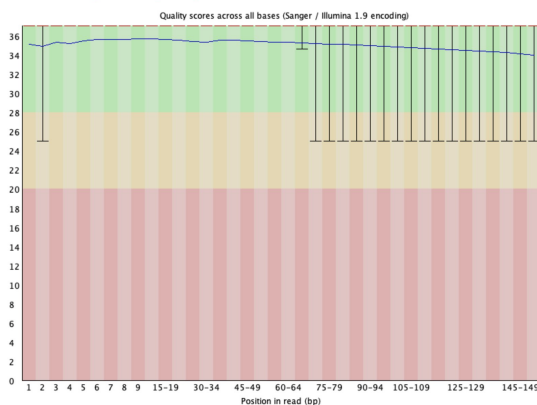
**Fórmula 1. Fórmula utilizada para el cálculo del Phred Score.** La calidad de cada base en una secuencia se mide mediante una puntuación Phred (Q), donde el valor P representa la probabilidad de que una base esté mal identificada.

**Basic Statistics**

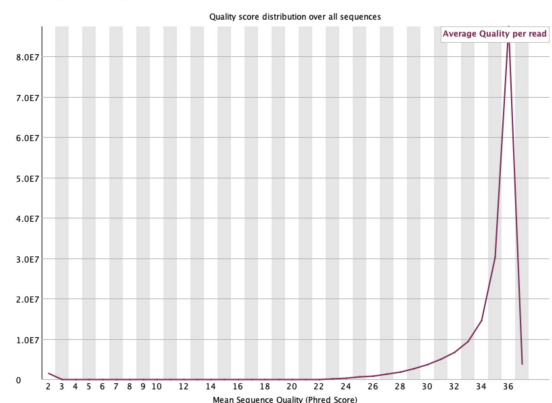
Measure	Value
Filename	L23-0030_S2_L001_R1_001.fastq.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	171811125
Total Bases	25.3 Gbp
Sequences flagged as poor quality	0
Sequence length	35-151
%GC	41

**Imagen 6. Parametros básicos de uno de los archivos FASTQ empleados en el análisis.** Captura de pantalla del archivo HTML generado por la herramienta FASTQC donde se puede observar que el archivo FASTQ contiene más de 170 millones de secuencias de ADN o lecturas de entre 35 y 151 pares de bases entre las cuales ninguna fue catalogada como de baja calidad poniendo así en evidencia la efectividad del preprocesamiento de calidad efectuado.

**Per base sequence quality**



**Per sequence quality scores**



**Imagen 7. Gráficas de calidad media por base (a) y calidad promedio por secuencia (b) de uno de los archivos FASTQ empleados en el análisis.** Captura de pantalla del archivo HTML generado por la herramienta FASTQC donde se puede observar que el calidad Phred por base en todos los casos fue igual o superior a 25 (menos de 1% de probabilidad de que la base este mal identificada) y que la mayoría de las secuencias presentaban un valor de calidad Phred promedio superior a 30 (menos de un 0,1% de probabilidad de error) indicando que los datos de secuenciación son de alta calidad con una gran parte de las secuencias siendo especial confiables.

### 3.5.2. Mapeo

Referido en el script de nextflow como *process BwaMem*. Utilizando el software BWA-MEM se procedió a alinear las lecturas de cada muestra contra el genoma humano de referencia hg38, previamente indexado con Samtools. Como resultado, se generó un archivo de alineamiento en formato SAM para cada muestra.

### 3.5.3. Eliminación de duplicados y recalibración del valor de calidad

Referido en el script de nextflow como *process MarkDuplicates*, *process BaseRecalibrator* y *process ApplyBQSR*. El siguiente paso consiste en identificar y marcar las lecturas duplicadas de los archivos SAM haciendo uso de la herramienta *MarkDuplicatesSpark* incluida en GATK. Los duplicados son secuencias de ADN idénticas que se han secuenciado más de una vez debido a la amplificación durante el proceso de preparación de la biblioteca. Estos fragmentos redundantes introducen sesgos en el análisis y pueden afectar la precisión de las variantes detectadas. Al marcar los duplicados, se indica a los algoritmos de análisis que ignoren el peso de estos duplicados en la identificación de variantes que se llevará a cabo más adelante, evitando así sobreestimaciones debido a la repetición de los mismos fragmentos de ADN. Tras ejecutar *MarkDuplicatesSpark* todos los archivos de alineamiento en formato SAM fueron convertidos en archivos de alineamiento en formato BAM con las lecturas ordenadas por su posición genómica y los duplicados marcados.

Seguidamente se procedió a recalibrar el valor de calidad de las bases haciendo uso de las herramientas *BaseRecalibrator* y *ApplyBQSR* incluidas en GATK. Con la herramienta *BaseRecalibrator*, se calcularon las estadísticas de recalibración a partir de las lecturas alineadas y un conjunto de variantes conocidas almacenadas en la base de datos de variantes conocidas en humanos dbSNP (Single Nucleotide Polymorphism Database), lo que permitió ajustar las puntuaciones de calidad de cada base. Este paso es crucial, ya que los secuenciadores a menudo presentan errores sistemáticos que pueden influir en las puntuaciones de calidad originales. Posteriormente, con la herramienta *ApplyBQSR*, se aplicaron los datos de recalibración obtenidos del análisis anterior, generando un nuevo archivo BAM con las puntuaciones de calidad de cada base recalibradas. Este ajuste mejorará la precisión del proceso de variant calling que se realizará a continuación reduciendo la cantidad de falsos positivos, resultando así en un conjunto de datos más confiable para análisis posteriores.

### 3.5.4. Variant calling

Referido en el script de nextflow como *process HaplotypeCaller* El proceso de llamada de variantes se realizó utilizando la herramienta *HaplotypeCaller* incluida en GATK. Esta herramienta permite identificar y extraer variantes genéticas al comparar las lecturas alineadas de una muestra en un archivo BAM con un genoma de referencia al evaluar cada posición en el genoma y detectar si existen variantes en forma de SNPs y pequeñas inserciones o deleciones. Los archivos de salida, en formato Variant Call Format (VCF), contienen todas las variantes detectadas para cada muestra, incluyendo información detallada sobre el tipo de variante, su posición en el genoma y diversas métricas de calidad que indican la fiabilidad de estas variantes. Entre estas métricas de calidad, destaca QUAL (Formula 2), que representa la confianza general en la autenticidad de una variante específica expresada en una escala logarítmica de probabilidad. La cual se calcula en función de la probabilidad de que la variante detectada sea real, donde un valor alto implica una mayor fiabilidad de la variante y uno más bajo una mayor probabilidad de que se trate de un falso positivo. Finalmente, para facilitar el almacenamiento y uso de estos archivos VCF en análisis posteriores, se les aplicó una compresión adicional con el comando *bgzip*.

$$\text{QUAL} = -10 \times \log_{10}(p_{\text{error}})$$

**Formula 2. Cálculo de la métrica de calidad QUAL.** La calidad general de una variante se determina empleando una escala logarítmica de probabilidad, donde el valor  $p_{\text{error}}$  representa la probabilidad estimada de que la variante sea un falso positivo. Este valor es calculado por GATK en función de varios parámetros de las lecturas alineadas en el archivo BAM, tales como la profundidad de cobertura, la calidad de mapeo de cada secuencia, la presencia de la variante en ambas direcciones de lectura, su posición en la secuencia alineada y la calidad de las bases que soportan dicha variante. Estos factores se integran mediante algoritmos de machine learning y modelos estadísticos que proporcionan una estimación confiable de  $p_{\text{error}}$ .

### 3.5.5. Anotación y priorización de variantes con *hiPhivePrioritiser*

Referido en el script de nextflow como *process runExomiser*. El último paso de la pipeline de análisis genómico consistió en la anotación y priorización de las variantes almacenadas en los archivos VCF de cada muestra mediante el uso de Exomiser el cual utiliza un archivo YAML como plantilla de configuración de análisis, el cual puede personalizarse para ajustarse a los objetivos y necesidades específicos de cada caso particular y según nuestros intereses (ver Anexo 7).



Para este análisis, el archivo YAML fue configurado para:

- Tomar la versión hg38 del genoma humano como referencia.
- Aplicar un filtro de calidad que solo considera variantes confiables, teniendo en cuenta únicamente aquellas con un valor de QUAL superior a 20, lo que se traduce en una probabilidad de ser un falso positivo inferior al 1%.
- Habilitar el uso de *hiPhivePrioritiser* empleando el término HPO específico para la HII (HP:0002516). Esta herramienta asigna a cada gen un valor entre 0 y 1, donde valores más altos reflejan una mayor asociación con el fenotipo HPO proporcionado. Para este análisis, se estableció un umbral de priorización superior a 0.5, este valor es el recomendado por el propio Exomiser para mejorar la precisión en la identificación de variantes con relevancia fenotípica significativa.
- Incluir información de frecuencia poblacional usando la base de datos gnomAD usando exclusivamente la cohorte caucásica no finlandesa, grupo poblacional al que pertenecen las muestras analizadas. No se aplicaron filtros sobre este valor, permitiendo que tanto variantes comunes como raras se tuvieran en cuenta en el análisis.
- Activar las herramientas de predicción de patogenicidad CADD, REMM, PolyPhen y SIFT integradas en exomiser para predecir y anotar el impacto funcional de las variantes.

Además, por defecto, Exomiser anota las variantes de acuerdo con su clase funcional, de acuerdo a los criterios de las recomendaciones ACMG (American College of Medical Genetics) e incluye su identificador rs, así como otra información de interés como entradas de ClinVar cuando están disponibles, la posición genómica exacta y el nombre del gen en el que se encuentran

Tras ejecutar exomiser con esta configuración YAML se obtuvo un archivo por muestra, con variantes priorizadas y anotadas, en formato TSV para facilitar su análisis y visualización posterior en Rstudio. Como se ha comentado anteriormente este proceso se automatizó además de manera aislada con el objetivo de anotar y priorizar las muestras que ya habían sido procesadas utilizando la herramienta DRAGEN en lugar de la pipeline completa.

### 3.5.6. Anotación dirigida de variantes. Panel de genes de AQP y estudio GWAS

Con el fin de validar en nuestra población de estudio los resultados de publicaciones anteriores que relacionan la HII con variantes genéticas específicas, concretamente los de Tellería-Orriols et al. (2023) y Kuehn et al. (2019), se configuró un nuevo archivo YAML (Anexo 8) para que Exomiser priorizara los genes con mayor asociación a la enfermedad según estos estudios. En lugar de utilizar *hiPhivePrioritiser* con el término HPO asociado a la HII, esta nueva versión del archivo YAML replicó los filtros de calidad y herramientas de anotación de variantes del archivo anterior, pero se deshabilitó la funcionalidad de *hiPhivePrioritiser* y se agregó un panel de genes de interés que incluyó AQP1, AQP2, y los 23 genes asociados con las 15 variantes que mayor asociación mostraron con HII en el estudio de GWAS realizado por Kuehn et al. En este caso el número de genes es superior al número de variantes ya que algunas de estas son intergénicas y al encontrarse entre dos genes se tuvieron en cuenta ambos. La información sobre los genes y las 15 variantes candidatas se obtuvieron consultando la base de datos pública GWAS Catalog correspondiente a dicho estudio.

De la misma forma que en el caso anterior, tras ejecutar *exomiser*, con esta nueva configuración YAML, se obtuvo un archivo TSV por muestra, con variantes priorizadas y anotadas, todas pertenecientes a los genes incluidos en el panel especificado en el archivo YAML.

### 3.6. Procesamiento y análisis estadístico de variantes priorizadas

Tras la ejecución de la pipeline de análisis genómico se obtuvieron como resultado un total de cien archivos en formato TSV, cincuenta correspondientes a la priorización con *hiPhivePrioritiser* y otros cincuenta correspondientes a la priorización dirigida con el panel de genes de interés. Estos archivos fueron procesados con un script de R en formato Rmarkdown (Anexo 9) cuyo objetivo fue el de determinar la frecuencia de aparición, de cada una de las variantes recogidas por el proceso de priorización en el grupo de afectados por HII y en el grupo control obteniéndose así dos grandes dataset de variantes anotadas donde además se incluía la frecuencia de aparición en el grupo de casos y grupo control que posteriormente serían sometidas a un análisis

estadístico para observar si hay una diferencia significativa que pudiera asociar dichas variantes con la enfermedad. Se determinó la frecuencia de aparición binaria, aparece o no aparece, sin tener en cuenta la frecuencia alélica

Con respecto al dataset obtenido del conjunto de archivos TSV resultantes de la priorización con *hiPhivePrioritiser*, se llevó a cabo un preprocesamiento para eliminar aquellas variantes que podrían generar ruido debido a su falta de información. Este filtrado incluyó:

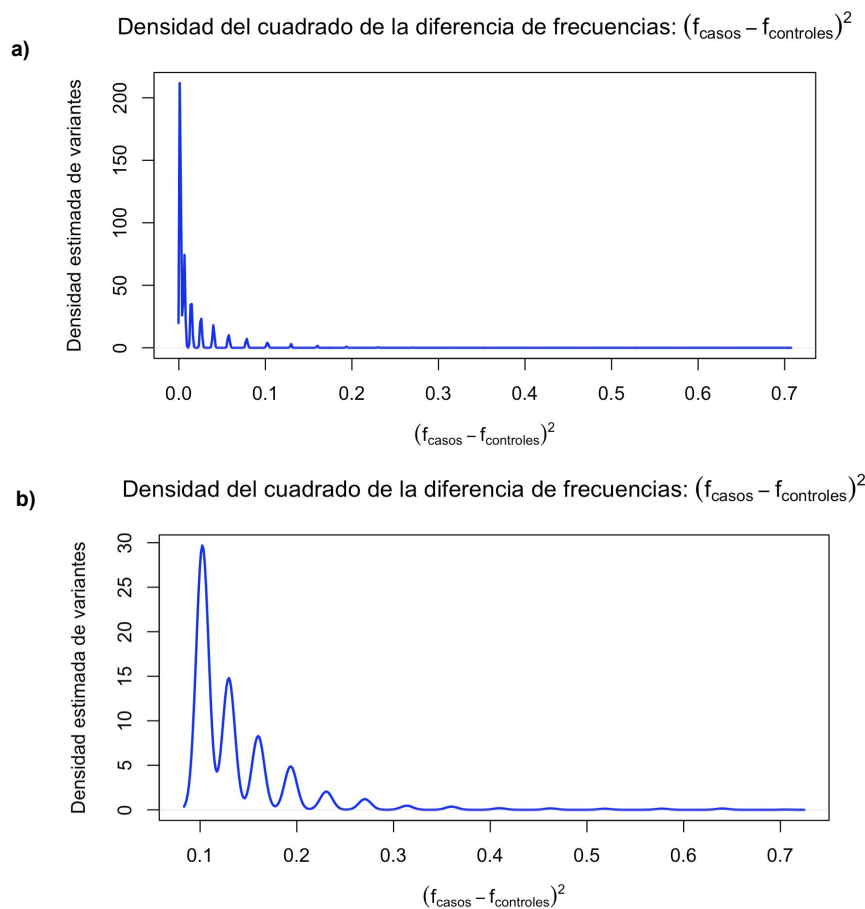
- Variantes con mayor frecuencia de aparición en el grupo control que en el grupo de afectados.
- Variantes con poca frecuencia de aparición en ambos grupos. Se decidió eliminar todas aquellas variantes que tenían una frecuencia de aparición igual o menor al 20% en ambos grupos ya que esto indicaba que muy posiblemente no estuvieran asociadas con la HII al no estar sobrerrepresentadas.
- Variantes con diferencias mínimas en frecuencia de aparición entre los dos grupos. Se calculó el cuadrado de la diferencia de frecuencias entre el grupo de casos y el grupo de controles (Formula 3) y se eliminó aquellas variantes cuya diferencia al cuadrado fue inferior a 0,1 lo que indicaba que la varianza era muy cercana a cero y que por tanto estas variantes no resultaban informativas en un estudio de asociación genómica y generarían ruido en los análisis estadísticos posteriores (Imagen x)

$$(frecuencia_{casos} - frecuencia_{control})^2 = \text{cuadrado de la diferencia}$$

**Formula 3. Cálculo del cuadrado de la diferencia de frecuencias.** Este valor se utilizó para filtrar variantes cuya frecuencia de aparición es similar entre el grupo de casos y el grupo de controles, eliminando así aquellas que no aportan información significativa. El cuadrado de la diferencia puede tomar valores de 0 a 1, donde valores cercanos a 0 indican una similitud alta en la frecuencia de aparición entre los grupos siendo el valor 0 indicativo de que la variante aparece con exactamente la misma frecuencia en ambos grupos.

Tras realizar este filtrado en base a la frecuencia de aparición pasamos de más de 400 mil variantes diferentes a aproximadamente 10 mil (Imagen 8). Seguidamente se realizó otro filtrado en base a los criterios de las recomendaciones ACMG, un conjunto de pautas estándar para la interpretación de variantes genéticas publicadas en 2015 y que ayudan a categorizar variantes en función de su potencial impacto clínico, filtrando todas aquellas clasificadas como benignas o probablemente

benignas y únicamente teniendo en cuenta aquellas de significado incierto, probablemente patogénicas y patogénicas (Richards et al., 2015). Se realizó un tercer y último filtrado en base a la clase funcional de las variantes eliminando aquellas variantes sinónimas e intergénicas, ya son las que tienen un menor potencial de afectar la función de los genes relevantes en la enfermedad. Tras estos dos filtrados se mantuvieron cerca de 8.500 variantes permitiendo así focalizar el análisis en aquellas con un mayor potencial patogénico y relevancia clínica.



**Imagen 8. Distribución de la densidad del valor de  $(f_{\text{casos}} - f_{\text{controlos}})^2$  antes (a) y después (b) de aplicar el filtrado de variantes basado en la frecuencia de aparición.** Como se puede observar, previo al filtrado, la cantidad de variantes no informativas debido a una frecuencia de aparición similar en ambos grupos representaba el porcentaje mayoritario de variantes introduciendo grandes cantidades de ruido por lo que se decidió filtrarlas.

Debido a la sobrerrepresentación de variantes localizadas en regiones no codificantes, especialmente en intrones, frente a variantes exónicas se optó por analizarlas en dos grupos diferenciados:

- **Variantes exónicas**. Donde se incluyen variantes missense, variantes de ganancia o pérdida de codón de inicio o terminación y variantes en regiones codificantes que induzcan cambios en el marco de lectura. Representado por un total de 37 variantes diferentes.
- **Variantes en regiones no codificantes**. Donde se incluyen variantes localizadas en las regiones UTRs (5' y 3'), en zonas de splicing, en intrones, en regiones upstream o downstream, así como en regiones reguladoras, como promotores o enhancers. Representado por 8421 variantes diferentes.

Esta estrategia se tomó para no desestimar la posible implicación de las variantes que impactan directamente en las regiones codificantes, cuya relevancia clínica y funcional difiere considerablemente de aquellas localizadas en regiones no codificantes.

Como último paso se realizó un análisis estadístico utilizando el test exacto de Fisher, dado los tamaños de muestra reducidos y las bajas frecuencias de aparición de algunas de las variantes. Este análisis se llevó a cabo de manera paralela en cada uno de los grupos previamente mencionados, con el objetivo de determinar si existía una diferencia estadísticamente significativa entre el grupo control y el grupo de afectados. Tras el test exacto de Fisher a cada variante se le asignó un p-value indicador del grado de asociación con la HII el cual fue corregido mediante el método de Holm–Bonferroni con el objetivo reducir el número de falsos positivos.

Con respecto al dataset obtenido del conjunto de archivos TSV resultantes de la priorización con el panel de genes de interés no se le realizó ningún filtrado de variantes en base a su frecuencia de aparición, clasificación clínica o funcional, únicamente se filtraron variantes de interés por su identificador rs y se realizó el test exacto de fisher con cada variante, asignándole un p-value, con el que determinar si había una diferencia significativa entre el grupo control y el grupo de afectados y estas estaban correlacionadas con la HII tal y como se había indicado en el correspondiente estudio del que se habían sido extraídos.

## 4. Resultados

### 4.1. Estudio de asociación genómica WGS

#### 4.1.1. Variantes localizadas en zonas codificantes

Con respecto al grupo de variantes con potencial de afectar de forma directa la secuencia del producto génico, tras la aplicación del test de Fisher y corrección del p-value tan solo 3 de las 37 variantes iniciales mostraron aparecer significativamente con mayor frecuencia en el grupo de afectados por HII que en el grupo control, utilizando un umbral de significancia de 0.1 (Tabla 1). Siendo la variante rs1279780267, localizada en el gen ASAH2, aquella con una asociación más fuerte ( $p\text{-value} = 6.15 \times 10^{-3}$ ).

rs ID	Posición genómica	Gen	Freq poblacional	Clase funcional	Freq HII	Freq Control	p-value ajustado
rs1279780267	10-50189586-G-T	ASAH2	0.5052 *	missense variant	0.76	0.20	0.00615
rs1050118	6-31269999-C-T	HLA-C	0.0000605	missense variant	0.60	0.12	0.03253
rs1131500	6-31355203-C-T	HLA-B	0.01659	missense variant	0.68	0.20	0.05045

**Tabla 1. Variantes con potencial de afectar de forma directa la secuencia del producto génico, significativamente asociadas con la HII en nuestro grupo de estudio.** El asterisco rojo sobre la frecuencia poblacional de una de las variantes indica que, según la base de datos gnomAD, esta variante esta cubierta en menos del 50% de los individuos analizados que la conforman, lo que sugiere que la estimación de su frecuencia poblacional podría no ser precisa y por lo tanto, debe interpretarse con cautela.

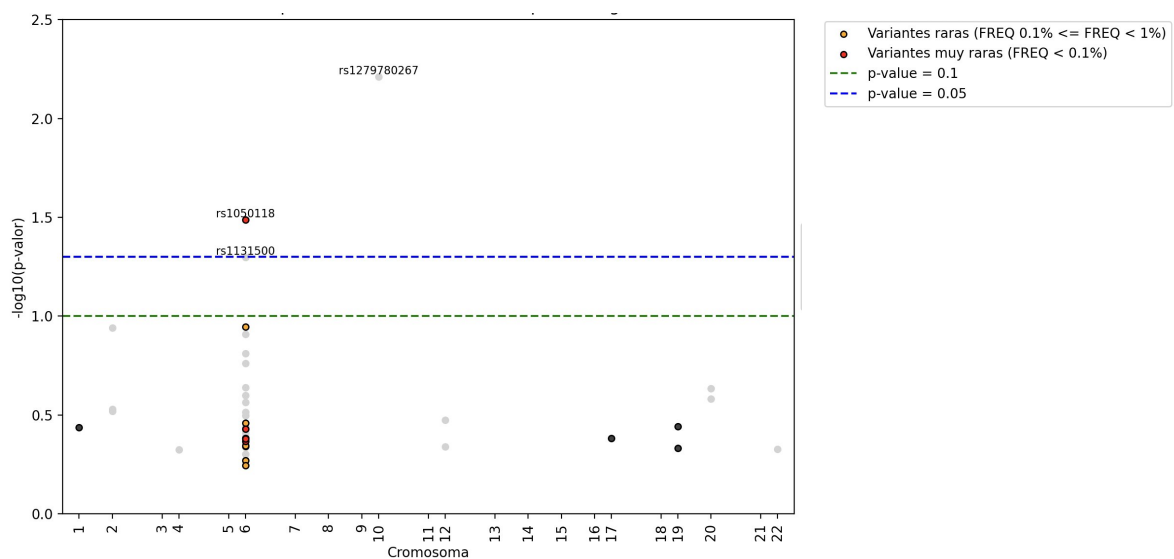
Las predicciones obtenidas mediante herramientas de evaluación de patogenicidad e impacto funcional para estas tres variantes fueron mixtas (Tabla 2). CADD fue la herramienta más estricta, asignando puntuaciones superiores a 0.5 a las tres variantes, lo que indica una muy posible asociación con algún rasgo fenotípico o enfermedad. En contraste, otras herramientas como SIFT y REVEL fueron más permisivas, calificando a las tres variantes como probablemente benignas o tolerables.

rs ID	CADD	PolyPhen	SIFT	REVEL
rs1279780267	0.9954	0.787	0.993	0.195
rs1050118	0.8490	0.05	0.997	0.044
rs1131500	0.5503	0.002	-	0.011

**Tabla 2. Puntuaciones asignadas por las herramientas de predicción de patogenicidad a las variantes exónicas asociadas con la HII en nuestra población de estudio.** Las puntuaciones en color rojo indican una alta posibilidad de patogenicidad, las puntuaciones en amarillo reflejan una posibilidad intermedia, y las puntuaciones en verde sugieren que la variante es probablemente benigna o que no tiene ningún efecto.

Con el objetivo de identificar de forma rápida y visual loci altamente sugestivos de estar correlacionados con la HII debido a la presencia de un número elevado de variantes, se representaron las 37 variantes exónicas en un gráfico tipo Manhattan, donde también se resaltaron aquellas variantes con una baja frecuencia poblacional según

la base de datos gnomAD. Estas variantes se señalaron ya que su rareza en la población puede indicar un mayor potencial patogénico estando asociadas a fenotipos específicos o enfermedades raras. (Imagen 9). Se observó un número elevado de variantes, especialmente variantes con una frecuencia de aparición en la población inferior al 1 %, en el cromosoma seis, concretamente en los genes que codifican el complejo de histocompatibilidad humano (HLA), aunque muchas de estas variantes no alcanzaron una significancia estadística.



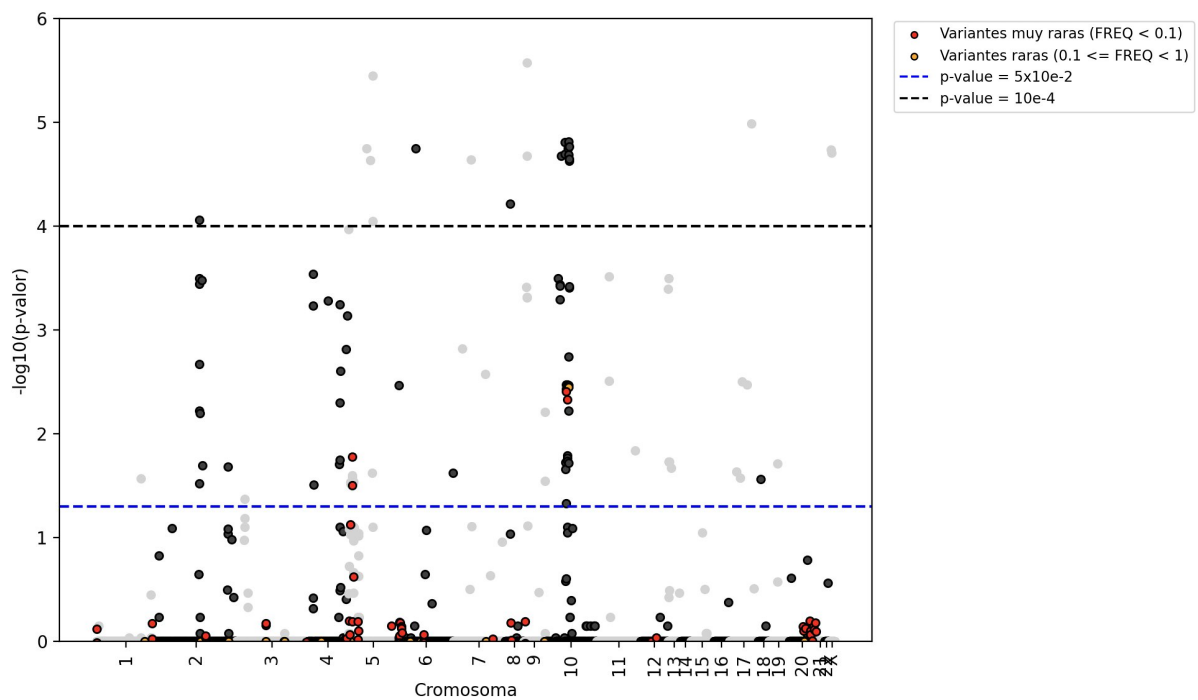
**Imagen 9. Gráfico tipo Manhattan donde se recogen todas las variantes exónicas incluidas en el estudio.** Las variantes, representadas como puntos, se sitúan en el gráfico en función de su posición cromosómica y su p-value asociado. A mayor significancia estadística mayor altura de la variante en el gráfico, indicando una asociación más fuerte con el HII. Las variantes raras con una cobertura inferior al 50% en los individuos que conforman la base de datos gnomAD no fueron resaltadas ya que como la misma base de datos advierte, la estimación de su frecuencia poblacional puede no ser precisa y por lo tanto, debe interpretarse con cautela, por lo que puede haber un número mayor de variantes raras.

#### 4.1.2. Variantes localizadas en zonas no codificantes

Enfocándonos en el grupo de variantes no codificantes, que representa con diferencia, la mayoría de variantes a analizar, tras aplicar el test de Fisher y corregir el p-value, se identificaron un total de 92 variantes de las 8421 iniciales que mostraron una frecuencia de aparición significativamente mayor en el grupo de afectados por HII que en el grupo control, utilizando un umbral de significancia de 0.05. Las variantes identificadas se pueden consultar en el Anexo 10, junto con las puntuaciones obtenidas por las herramientas de predicción de patogénicidad REMM y CADD, las cuales, en muchos casos presentaron resultados mixtos. Destacan genes como MARCHF1, PDE4D o ASAH2 los cuales acumulan una gran cantidad de variantes estadísticamente significativas, en especial ASAH2 con cerca de treinta, algunas de ellas presentes en todos los afectados.

Se representaron todas las variantes no codificantes en un gráfico tipo Manhattan y se

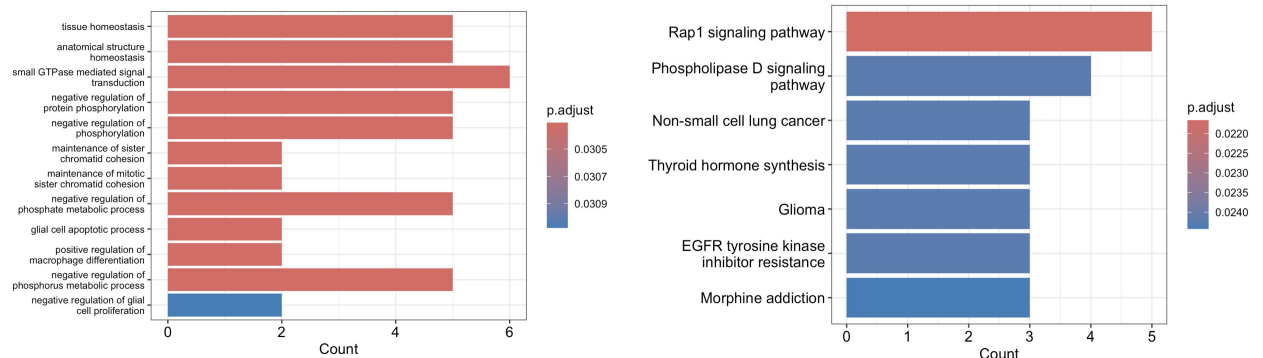
señalaron aquellas con una frecuencia poblacional baja según la base de datos gnomAD (Imagen 10). Se observaron varios loci sugestivos de estar correlacionados con la HII con una alta concentración de variantes estadísticamente significativas en los cromosomas 2, 4, 5 y 10, destacando este último sobre el resto debido a la acumulación de variantes en el gen *ASAH2* localizado en este cromosoma. Dejando de lado el gen *ASAH2* destacan también otras variantes, con una muy alta asociación con la HII presentes en genes como *ADCY2* y *CNTNAP3* ( $p\text{-value} = 3.16 \times 10^{-6}$ ) o *NF1* y *IL92* ( $p\text{-value} = .89 \times 10^{-5}$ )



**Imagen 10. Gráfico tipo Manhattan donde se recogen todas las variantes no codificadas incluidas en el estudio.** Las variantes, representadas como puntos, se sitúan en el gráfico en función de su posición cromosómica y su p-value asociado. A mayor significancia estadística mayor altura de la variante en el gráfico, indicando una asociación más fuerte con al HII. Las variantes raras con una cobertura inferior al 50% en los individuos que conforman la base de datos gnomAD no fueron resaltadas como tales ya que como la misma base de datos advierte, que la estimación de su frecuencia poblacional puede no ser precisa y por lo tanto, debe interpretarse con cautela, por lo que puede haber un número mayor de variantes raras.

También se realizó un análisis de enriquecimiento utilizando términos de Gene Ontology (GO) y la base de datos KEGG, con el objetivo de identificar funciones moleculares y rutas de señalización en las que los genes asociados pudieran estar implicados de manera relevante (Imagen 11). Este análisis reveló un enriquecimiento en funciones relacionadas con la homeostasis de tejidos y la regulación de la señalización intracelular, concretamente en procesos mediados por pequeñas GTPasas y en la regulación negativa de la fosforilación de proteínas. Paralelo a esto, se identificaron rutas de señalización clave en el enriquecimiento con KEGG, entre las que destacan la vía de señalización de Rap1 y la vía de señalización de la fosfolipasa D.





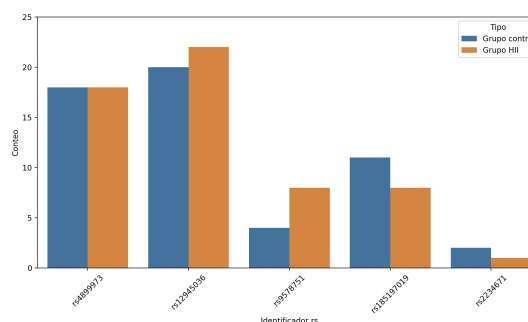
**Imágen 11. Análisis de enriquecimiento de términos GO (a) y rutas KEGG (b) de los genes con variantes asociadas a la HII.** La gráfica (a) muestra los procesos biológicos, mientras que la gráfica (b) resalta las rutas metabólicas y de señalización más asociadas o en las que participan los genes de interés. Gráficos generados utilizando el paquete *clusterProfiler* en R para el análisis de enriquecimiento GO y vías KEGG. El análisis GO se realizó sobre la ontología 'BP' (biological process) con un valor de  $p < 0.05$ , mientras que el análisis KEGG se realizó sobre el organismo humano ('hsa') con un valor de  $p < 0.05$ .

## 4.2. Validación del estudio de acuaporinas

Al analizar la frecuencia de aparición de las variantes estudiadas por Tellería-Oriols et al. en nuestra población de estudio, encontramos que, al igual que en su publicación, las variantes rs2075575, rs3763043 y rs1049305 no mostraron una correlación significativa con el grupo de afectados por HII. Sin embargo, la variante rs3763040, que Tellería-Oriols et al. señalaron como posible inductora de la HII, tampoco se presentó significativamente más en el grupo de afectados en comparación con el grupo de controles (Tabla 3) (Imagen 12), de hecho, en ambos grupos la frecuencia de aparición de esta variante fue similar a la reportada por la base de datos gnomAD para población caucásica no finlandesa.

rs ID	Posición genómica	Gen	Freq poblacional	Clase funcional	Freq HII	Freq Control	p-value
rs2075575	18-26866562-G-A	AQP4	0.427	upstream gene	0.40	0.60	0.258
rs3763040	18-26864410-G-A	AQP4	0.211	coding transcript intron	0.20	0.28	0.741
rs3763043	18-26855854-C-T	AQP4	0.321	3' UTR exon	0.44	0.36	0.773
rs1049305	7-30924207-G-C	AQP1	0.383	3' UTR exon	0.56	0.48	0.778

**Tabla 3 . Variantes genéticas estudiadas por Tellería-Oriols et al. 2023 anotadas y analizadas en nuestra población de estudio.**



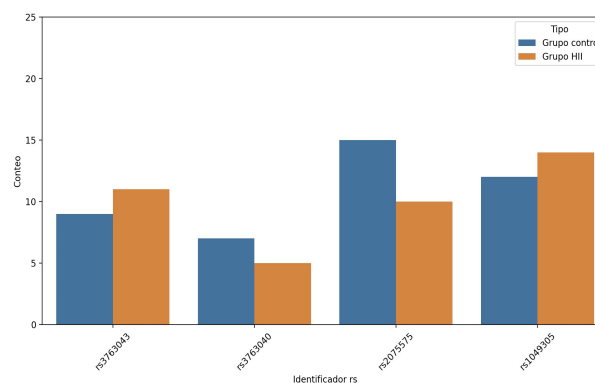
**Imagen 12. Conteo de cada uno de las variantes genéticas estudiadas por Tellería-Oriols et al. 2023 en nuestra población de estudio**

### 4.3. Validación del estudio de GWAS

Al analizar la frecuencia de aparición de las 15 variantes con la asociación más fuerte a la HII según Kuehn et al. (2019), encontramos que tan solo 5 de estas variantes estaban presentes en nuestra población de estudio, además, ninguna de ellas mostró una frecuencia de aparición significativamente mayor en el grupo de afectados por HII que en el grupo de control (Tabla 4) (Imagen 13).

rs ID	Posición genómica	Gen	Freq poblacional	Clase funcional	Freq HII	Freq Control	p-value
rs9578751	13-24500352-C-T	PARP4	0.117	missense variant	0.32	0.16	0.321
rs185197019	8-72537054-G-A	KCNB2	0.223	upstream gene	0.32	0.44	0.561
rs12945036	17-230930-C-G	RPH3AL	0.581	coding transcript intron	0.88	0.80	0.702
rs2234671	2-218164385-C-G	CXCR1	0.056	missense variant	0.04	0.08	1
rs4899973	14-89272011-T-C	FOXN3	0.424	coding transcript intron	0.72	0.72	1

**Tabla 4. Variantes genéticas relevantes en la publicación de Kuehn et al. 2019 anotadas y analizadas en nuestra población de estudio**



**Imagen 13. Conteo de cada uno de las variantes genéticas relevantes en la publicación Kuehn et al. 2019 en nuestra población de estudio**

## 5. Discusión

En este trabajo se analizó si existe una asociación entre la HII y variantes genéticas que puedan explicar la predisposición de algunos pacientes a desarrollar esta condición siendo el uso del WGS y el de un priorizador de variantes como *hiPhivePrioritiser*, que tiene en cuenta tanto el interactoma como el efecto funcional en diferentes modelos biológicos de las variantes analizadas, su principal fortaleza y punto distintivo frente a publicaciones anteriores en las cuales se analizaban únicamente pequeñas fracciones del genoma, ya sea mediante la secuenciación de genes concretos o el uso de chips de micro-arrays de alta densidad.

Es importante destacar que, si bien los resultados obtenidos son prometedores y pueden sentar las bases para el entendimiento de los mecanismos genéticos implicados en la HII, la población de estudio es pequeña, lo que limita la potencia estadística y requiere que los hallazgos

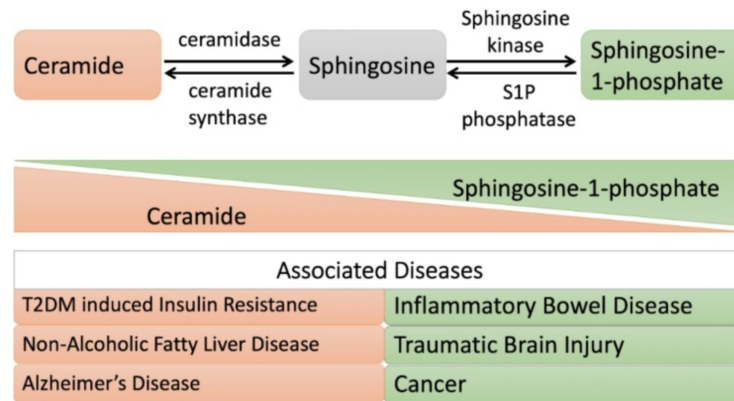
obtenidos sean interpretados con cierta cautela. Este trabajo constituye el primer paso en un estudio mucho más amplio, en el cual se planea aumentar considerablemente la población de estudio, aún así, las conclusiones obtenidas hasta ahora pueden ser reveladoras.

Ninguno de los loci o genes sugestivos de estar involucrados en el desarrollo de la HII hallados en este estudio se corresponde con los indicados en estudios anteriores aunque se refuerza la idea de que la HII se trata de una condición poligénica de herencia compleja causada por mutaciones en varios genes que puedan contribuir en diferente medida en su aparición. Las variantes localizadas en los genes de la familia de las acuaporina propuestos por Tellería-Oriols et al. y más concretamente la variante rs3763040 localizada en AQP4 no estuvieron correlacionadas con la HII en nuestra población de estudio, por lo que, aunque esta familia de proteínas puedan verse afectadas durante el desarrollo de la enfermedad debido a su implicación en la homeostasis del LCR, los hallazgos obtenidos en este estudio nos indican que la base de la fisiopatología de la HII no parece deberse a variantes genéticas en este par de genes, respaldando así los resultados obtenidos por Kerty et al 2013 donde tras secuenciar el gen AQP4 de un grupo de pacientes diagnosticados con HII no se halló ninguna variante causal.

En relación con las variantes genéticas identificadas como altamente correlacionadas con la HII por Kuehn et al. (2019), solo cinco de las quince reportadas en el GWAS catalogado asociado a dicho estudio aparecieron en nuestra población de estudio. Esta discrepancia podría deberse a la pequeña muestra poblacional disponible en nuestra investigación o a su alta homogeneidad poblacional en comparación con la población utilizada en el estudio de Kuehn et al. (2019), que incluye individuos de diversas etnias. Ninguna de las cinco variantes identificadas mostró una asociación significativamente más fuerte en el grupo afectado por HII, con lo cual, no podemos afirmar que ninguna de ellas esté implicada con el desarrollo de la enfermedad en nuestra población de estudio.

Centrándonos en los resultados obtenidos en el estudio de asociación genómica WGS, más concretamente en las variantes exónicas, si bien ninguna de las que mostraron una correlación significativa con la HII presentaron una asociación tan fuerte como alguna de las variantes localizadas en zonas no codificantes, en las que profundizaremos más adelante, su presencia no debe desestimarse pues son variantes que tienen un efecto directo sobre la secuencia del producto génico de sus respectivos genes. Por encima de todas, destaca rs1279780267 ( $p$ -value =  $6.15 \times 10^{-3}$ ) localizada en el gen *ASAH2* que codifica la ceramidasa neutra, una pieza clave en el metabolismo lipídico, la cual cataliza la hidrólisis de las ceramidas para producir un ácido graso y una esfingosina, la cual puede ser fosforilada posteriormente por una esfingosina quinasa (SphK) para producir esfingosina-1-fosfato (S1P)

Tanto la ceramida como la S1P son lípidos bioactivos que regulan diversas funciones celulares en diferentes rutas de señalización por lo que niveles alterados en alguno de los dos se han relacionado con diferentes patologías tales como el Alzheimer, la diabetes tipo 2, algunos tipos de cáncer o enfermedades intestinales inflamatorias (Imagen 14) (Parveen et al., 2019).



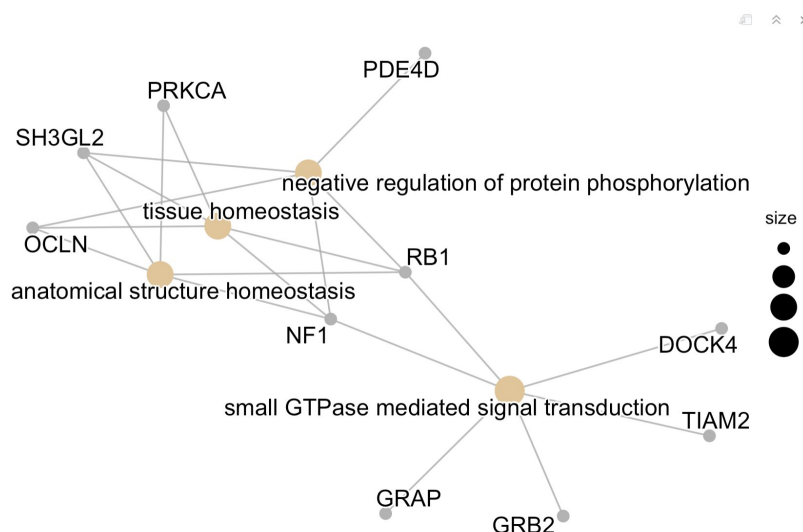
**Imagen 14 .Metabolismo de la esfingosina y su implicación en diferentes enfermedades.** El desequilibrio entre ceramida y S1P puede orientar la respuesta celular hacia la apoptosis o la proliferación, respectivamente, lo que contribuye al desarrollo de diversas patologías.

Diversos estudios han relacionado la S1P con vías proinflamatorias, lo que la convierte en un factor clave a tener en cuenta en la modulación de la inflamación. Por ejemplo, la deficiencia de esfingosina-1-fosfatasa, la principal enzima responsable del catabolismo de S1P, da lugar a una respuesta inflamatoria exacerbada, probablemente debido a la acumulación de S1P (Allende et al., 2011). La SphK y la S1P son esenciales para la producción, inducida por TNF- $\alpha$ , de prostaglandina E2R, un mediador lipídico que desempeña un papel crucial en los procesos inflamatorio como la fiebre o la vasodilatación (Pettus et al., 2003). También se ha observado que los niveles de S1P están elevados en diversas condiciones inflamatorias crónicas, como la esclerosis múltiple (Kułakowska et al., 2010) o el asma (Ammit et al., 2001), lo que pone en manifiesto una participación clave en la exacerbación de la inflamación en estos trastornos. La S1P juega también un papel crucial en la regulación de la expresión de RANKL, una molécula pro inflamatoria implicada en la artritis reumatoide (Takeshita et al., 2012) y en la modulación de la permeabilidad vascular (Pettus et al., 2003) Por otro lado, se ha demostrado como la S1P también participa activamente en la regulación de la esteroidogénesis, estimulando la biosíntesis de cortisol (Lucki et al., 2011; Rábano et al., 2003).

De esta manera, debido a su estrecha relación con vías proinflamatorias, su rol en la modulación de procesos que dan lugar a inflamaciones crónicas leves, así como su participación en la biosíntesis de cortisol, que como se menciona en el apartado 1.5.4, está asociada a un aumento de la actividad de la Na<sup>+</sup>/K<sup>+</sup>-ATPasa en las células del plexo coroideo, incrementando

así la producción de LCR y su papel clave en el metabolismo lipídico, *ASAH2* se perfila como uno de los genes más susceptibles de estar implicado en el desarrollo de HII. Además, este gen presenta un número elevado de variantes significativas en sus regiones intrónicas, con una fuerte asociación con el grupo de afectados por HII ( $p\text{-value} = 1.89 \times 10^{-5}$ ).

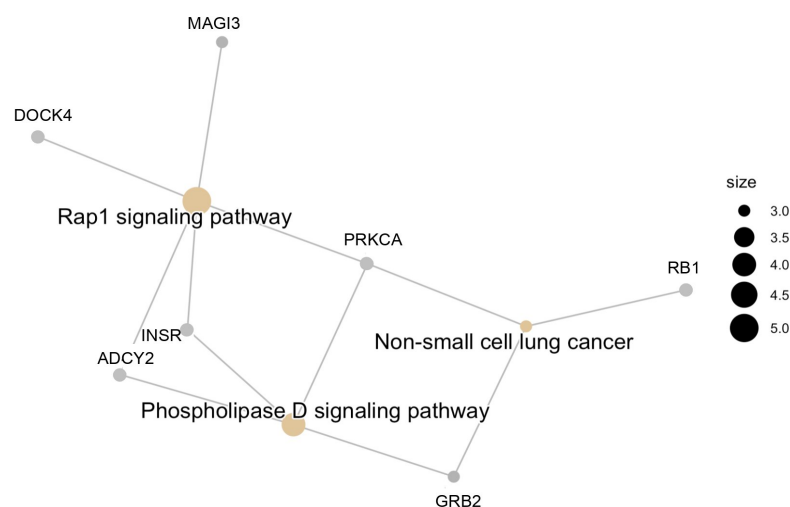
Los análisis de enriquecimiento también sugieren que un desequilibrio en el metabolismo de las esfingosinas podría ser clave en la patogénesis de la HII. Un número elevado de variantes altamente correlacionadas con la HII en nuestra población de estudio se localiza en regiones intrónicas de genes implicados en la regulación negativa de la fosforilación (Imagen 15). Esta regulación es esencial para evitar la hiperactivación de ciertas vías celulares, y en ella participan tanto proteínas quinasas, que añaden grupos fosfato, como fosfatasas, que los eliminan. Si la regulación negativa de la fosforilación se ve comprometida, podría descontrolarse la actividad de SphK o producirse una deficiencia en la actividad de la esfingosina-1-fosfatasa, lo que llevaría a una acumulación de S1P que podría impactar en la respuesta inflamatoria y la biosíntesis de cortisol.



**Imagen 15 .Genes asociados a los procesos biológicos más enriquecidos en el análisis de términos GO realizado sobre variantes no codificantes.** El gráfico de red (cnetplot), generado con el paquete *clusterProfiler* en R, donde se visualiza la relación entre los genes y los términos GO más enriquecidos. Los nodos representan los términos GO y los genes, mientras que las conexiones indican cómo los genes se asocian con los diferentes términos funcionales. El análisis GO se realizó sobre la ontología 'BP' (biological process), utilizando un valor de  $p < 0.05$  como umbral de significancia

Los análisis también mostraron un enriquecimiento en variantes localizadas en genes involucrados en la ruta de señalización de Rap1, una GTPasa pequeña perteneciente a la familia de las proteínas Ras las cuales actúan como interruptores moleculares, regulando la activación e inactivación de diversas vías de señalización intracelulares, a través de la adición y eliminación de grupos fosfato. Estudios confirman que Rap1 modula procesos clave en la respuesta inmune y la inflamación, pues se ha observado que promueve la producción de

citoquinas a través de la activación del factor de transcripción kappa B (NFκB,) favoreciendo un ambiente proinflamatorio que conduce al desarrollo y progresión de enfermedades como la aterosclerosis.(Cai et al., 2015). También se observó, aunque de manera menos significativa, un enriquecimiento de variantes intrónicas localizadas en genes involucrados en la ruta de señalización de la fosfolipasa D. Se ha demostrado que S1P está implicado en la activación de esta vía, promoviendo la producción de citoquinas proinflamatorias como IL-8 en células epiteliales y macrófagos. Esto sugiere que la interacción entre S1P y la ruta de fosfolipasa D podría ser un factor clave en la regulación de la inflamación y en el desarrollo de ciertas enfermedades inflamatorias crónicas como el asma (Cummings et al., 2002)



**Imagen 16 . Genes asociados a las rutas de señalización más enriquecidas en el análisis de rutas KEGG realizado sobre variantes no codificantes.** El gráfico de red (cnetplot), generado con el paquete clusterProfiler en R, visualiza la relación entre los genes y las vías KEGG más enriquecidas. Los nodos representan las vías KEGG y los genes, mientras que las conexiones muestran cómo los genes se asocian con las diferentes vías funcionales. El análisis KEGG se realizó sobre el organismo humano ('hsa') con un valor de  $p < 0.05$  como umbral de significancia

Todos estos hallazgos, junto con la identificación de variantes raras significativamente asociadas al grupo de pacientes afectados por HII en la región codificante de los genes HLA, que alteran su zona de unión a antígenos o epítipo (Busch et al., 2019) , así como en las regiones intrónicas del gen IL9, que codifica para una citocina que influye en la regulación del sistema inmunológico, especialmente en la activación de células asociadas con reacciones alérgicas e inflamatorias (Neurath & Finotto, 2016) (Deng et al., 2017) respaldan la hipótesis de que la HII es una condición originada por alteraciones inmunológicas sugiriendo que esta enfermedad no se debe únicamente a alteraciones en la homeostasis LCR, sino que también se debe a un componente inmunológico o de inflamación crónica subyacente, abriendo la puerta tanto a nuevas dianas teraputicas como al desarrollo de herramientas diagnósticas más precisas.

## 6. Conclusiones

A través de la realización de este trabajo y el subsecuente análisis de los resultados obtenidos se pudieron extraer las siguientes conclusiones:

- Las regiones no codificantes del genoma albergan una gran variabilidad genética que debe tenerse en cuenta para comprender los mecanismos y la predisposición a diversas enfermedades.
- La HII presenta un posible componente inmunológico o de inflamación crónica que puede explicar gran parte de su etiología y se respalda con estudios anteriores
- Alteraciones en el metabolismo lipídico de las esfingosinas, ocasionadas por variantes genéticas, ya sea en su síntesis a través de la ceramidasa neutra ASAH2 como en su fosforilación para producir S1P, podrían promover vías proinflamatorias y la síntesis de cortisol, contribuyendo así a la sintomatología de la HII.
- Los resultados obtenidos, aunque prometedores, deben interpretarse con cautela debido a la limitada población de estudio. En estudios futuros será necesario ampliar la muestra y analizar en detalle las variantes con mayor asociación a la HII para comprender su impacto funcional y su posible implicación en el desarrollo de la enfermedad.



## Bibliografía

1. Abbott, S., Chan, F., Tahrani, A. A., Wong, S. H., Campbell, F. E. J., Parmar, C., Pournaras, D. J., Denton, A., Sinclair, A. J., & Mollan, S. P. (2023). Weight Management Interventions for Adults With Idiopathic Intracranial Hypertension: A Systematic Review and Practice Recommendations. *Neurology*, 101(21), e2138-e2150. <https://doi.org/10.1212/WNL.0000000000207866>
2. Alganmi, N., & Abusamra, H. (2023). Evaluation of an optimized germline exomes pipeline using BWA-MEM2 and Dragen-GATK tools. *PloS One*, 18(8), e0288371. <https://doi.org/10.1371/journal.pone.0288371>
3. Alimajstorovic, Z., Mitchell, J. L., Yiangou, A., Hancox, T., Southam, A. D., Grech, O., Ottridge, R., Winder, C. L., Tahrani, A. A., Tan, T. M., Mollan, S. P., Dunn, W. B., & Sinclair, A. J. (2023). Determining the role of novel metabolic pathways in driving intracranial pressure reduction after weight loss. *Brain Communications*, 5(5), fcad272. <https://doi.org/10.1093/braincomms/fcad272>
4. Alimajstorovic, Z., Mollan, S. P., Grech, O., Mitchell, J. L., Yiangou, A., Thaller, M., Lyons, H., Sassani, M., Seneviratne, S., Hancox, T., Jankevics, A., Najdekr, L., Dunn, W., & Sinclair, A. J. (2023). Dysregulation of Amino Acid, Lipid, and Acylpyruvate Metabolism in Idiopathic Intracranial Hypertension: A Non-targeted Case Control and Longitudinal Metabolomic Study. *Journal of Proteome Research*, 22(4), 1127-1137. <https://doi.org/10.1021/acs.jproteome.2c00449>
5. Allende, M. L., Bektas, M., Lee, B. G., Bonifacino, E., Kang, J., Tuymetova, G., Chen, W., Saba, J. D., & Proia, R. L. (2011). Sphingosine-1-phosphate Lyase Deficiency Produces a Pro-inflammatory Response While Impairing Neutrophil Trafficking\*. *Journal of Biological Chemistry*, 286(9), 7348-7358. <https://doi.org/10.1074/jbc.M110.171819>
6. Ammit, A. J., Hastie, A. T., Edsall, L. C., Hoffman, R. K., Amrani, Y., Krymskaya, V. P., Kane, S. A., Peters, S. P., Penn, R. B., Spiegel, S., & Panettieri, R. A. (2001). Sphingosine 1-phosphate modulates human airway smooth muscle cell functions that promote inflammation and airway remodeling in asthma. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, 15(7), 1212-1214. <https://doi.org/10.1096/fj.00-0742fje>
7. Andrews, L. E., Liu, G. T., & Ko, M. W. (2014). Idiopathic intracranial hypertension and obesity. *Hormone Research in Paediatrics*, 81(4), 217-225. <https://doi.org/10.1159/000357730>
8. Arnold, R., Issar, T., Krishnan, A. V., & Pussell, B. A. (2016). Neurological complications in chronic kidney disease. *JRSM Cardiovascular Disease*, 5, 2048004016677687. <https://doi.org/10.1177/2048004016677687>



9. Ashley, E. A., Butte, A. J., Wheeler, M. T., Chen, R., Klein, T. E., Dewey, F. E., Dudley, J. T., Ormond, K. E., Pavlovic, A., Morgan, A. A., Pushkarev, D., Neff, N. F., Hudgins, L., Gong, L., Hodges, L. M., Berlin, D. S., Thorn, C. F., Sangkuhl, K., Hebert, J. M., ... Altman, R. B. (2010). Clinical assessment incorporating a personal genome. *Lancet* (London, England), 375(9725), 1525-1535. [https://doi.org/10.1016/S0140-6736\(10\)60452-7](https://doi.org/10.1016/S0140-6736(10)60452-7)
10. Bagger, F. O., Borgwardt, L., Jespersen, A. S., Hansen, A. R., Bertelsen, B., Kodama, M., & Nielsen, F. C. (2024). Whole genome sequencing in clinical practice. *BMC Medical Genomics*, 17(1), 39. <https://doi.org/10.1186/s12920-024-01795-w>
11. Ball, A. K., Sinclair, A. J., Curnow, S. J., Tomlinson, J. W., Burdon, M. A., Walker, E. A., Stewart, P. M., Nightingale, P. G., Clarke, C. E., & Rauz, S. (2009). Elevated cerebrospinal fluid (CSF) leptin in idiopathic intracranial hypertension (IIH): Evidence for hypothalamic leptin resistance? *Clinical Endocrinology*, 70(6), 863-869. <https://doi.org/10.1111/j.1365-2265.2008.03401.x>
12. Behera, S., Catreux, S., Rossi, M., Truong, S., Huang, Z., Ruehle, M., Visvanath, A., Parnaby, G., Roddey, C., Onuchic, V., Cameron, D. L., English, A., Mehtalia, S., Han, J., Mehio, R., & Sedlazeck, F. J. (2024). Comprehensive and accurate genome analysis at scale using DRAGEN accelerated algorithms. *bioRxiv: The Preprint Server for Biology*, 2024.01.02.573821. <https://doi.org/10.1101/2024.01.02.573821>
13. Behera, S., Catreux, S., Rossi, M., Truong, S., Huang, Z., Ruehle, M., Visvanath, A., Parnaby, G., Roddey, C., Onuchic, V., Finocchio, A., Cameron, D. L., English, A., Mehtalia, S., Han, J., Mehio, R., & Sedlazeck, F. J. (2024). Comprehensive genome analysis and variant detection at scale using DRAGEN. *Nature Biotechnology*. <https://doi.org/10.1038/s41587-024-02382-1>
14. Belkadi, A., Bolze, A., Itan, Y., Cobat, A., Vincent, Q. B., Antipenko, A., Shang, L., Boisson, B., Casanova, J.-L., & Abel, L. (2015). Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proceedings of the National Academy of Sciences of the United States of America*, 112(17), 5473-5478. <https://doi.org/10.1073/pnas.1418631112>
15. Benarroch, E. E. (2016). Choroid plexus--CSF system: Recent developments and clinical correlations. *Neurology*, 86(3), 286-296. <https://doi.org/10.1212/WNL.0000000000002298>
16. Biousse, V. (2012). Idiopathic intracranial hypertension: Diagnosis, monitoring and treatment. *Revue Neurologique*, 168(10), 673-683. <https://doi.org/10.1016/j.neurol.2012.07.018>
17. Botfield, H. F., Uldall, M. S., Westgate, C. S. J., Mitchell, J. L., Hagen, S. M., Gonzalez, A. M., Hodson, D. J., Jensen, R. H., & Sinclair, A. J. (2017). A glucagon-like peptide-1 receptor agonist reduces intracranial pressure in a rat model of hydrocephalus. *Science Translational Medicine*, 9(404), eaan0972. <https://doi.org/10.1126/scitranslmed.aan0972>

18. Bouffard, M. A. (2020). Fulminant Idiopathic Intracranial Hypertension. *Current Neurology and Neuroscience Reports*, 20(4), 8. <https://doi.org/10.1007/s11910-020-1026-8>
19. Bruce, B. B., Kedar, S., Van Stavern, G. P., Monaghan, D., Acierno, M. D., Braswell, R. A., Preechawat, P., Corbett, J. J., Newman, N. J., & Biousse, V. (2009). Idiopathic intracranial hypertension in men. *Neurology*, 72(4), 304-309. <https://doi.org/10.1212/01.wnl.0000333254.84120.f5>
20. Buchheit, W. A., Burton, C., Haag, B., & Shaw, D. (1969). Papilledema and Idiopathic Intracranial Hypertension. *New England Journal of Medicine*, 280(17), 938-942. <https://doi.org/10.1056/NEJM196904242801707>
21. Busch, R., Kollnberger, S., & Mellins, E. D. (2019). HLA associations in inflammatory arthritis: Emerging mechanisms and clinical implications. *Nature Reviews Rheumatology*, 15(6), 364-381. <https://doi.org/10.1038/s41584-019-0219-5>
22. Cai, Y., Sukhova, G. K., Wong, H. K., Xu, A., Tergaonkar, V., Vanhoutte, P. M., & Tang, E. H. C. (2015). Rap1 induces cytokine production in pro-inflammatory macrophages through NFκB signaling and is highly expressed in human atherosclerotic lesions. *Cell Cycle*, 14(22), 3580. <https://doi.org/10.1080/15384101.2015.1100771>
23. Chagot, C., Blonski, M., Machu, J.-L., Bracard, S., Lacour, J.-C., & Richard, S. (2017). Idiopathic Intracranial Hypertension: Prognostic Factors and Multidisciplinary Management. *Journal of Obesity*, 2017, 5348928. <https://doi.org/10.1155/2017/5348928>
24. Chai, N. C., Scher, A. I., Moghekar, A., Bond, D. S., & Peterlin, B. L. (2014). Obesity and headache: Part I--a systematic review of the epidemiology of obesity and headache. *Headache*, 54(2), 219-234. <https://doi.org/10.1111/head.12296>
25. Cogollo, J. A. M., Ávila, C. A. G., & Mejía, F. E. S. (2023). Sistema glinfático: Aspectos anatómicos, fisiológicos e implicaciones clínicas. *Acta Neurológica Colombiana*, 39(2), Article 2. <https://doi.org/10.22379/anc.v39i2.835>
26. Colman, B. D., Boonstra, F., Nguyen, M. N., Raviskanthan, S., Sumithran, P., White, O., Hutton, E. J., Fielding, J., & Walt, A. van der. (2024). Understanding the pathophysiology of idiopathic intracranial hypertension (IIH): A review of recent developments. *Journal of Neurology, Neurosurgery & Psychiatry*, 95(4), 375-383. <https://doi.org/10.1136/jnnp-2023-332222>
27. Contreras-Martin, Y., & Bueno-Perdomo, J. H. (2015). Idiopathic intracranial hypertension: Descriptive analysis in our setting. *Neurologia (Barcelona, Spain)*, 30(2), 106-110. <https://doi.org/10.1016/j.nrl.2013.08.009>

28. Corbett, J. J., Savino, P. J., Thompson, H. S., Kansu, T., Schatz, N. J., Orr, L. S., & Hopson, D. (1982). Visual loss in pseudotumor cerebri. Follow-up of 57 patients from five to 41 years and a profile of 14 patients with permanent severe visual loss. *Archives of Neurology*, 39(8), 461-474. <https://doi.org/10.1001/archneur.1982.00510200003001>
29. Craig, J. J., Mulholland, D. A., & Gibson, J. M. (2001). Idiopathic intracranial hypertension; incidence, presenting features and outcome in Northern Ireland (1991-1995). *The Ulster Medical Journal*, 70(1), 31-35.
30. Cummings, R. J., Parinandi, N. L., Zaiman, A., Wang, L., Usatyuk, P. V., Garcia, J. G. N., & Natarajan, V. (2002). Phospholipase D activation by sphingosine 1-phosphate regulates interleukin-8 secretion in human bronchial epithelial cells. *The Journal of Biological Chemistry*, 277(33), 30227-30235. <https://doi.org/10.1074/jbc.M111078200>
31. Daou, B. J., Sweid, A., Weinberg, J. H., Starke, R. M., Sergott, R. C., Doermann, A. S., Hauge, J., Zanaty, M., Chalouhi, N., Gooch, R., Herial, N., Zarzour, H., Jabbour, P., Rosenwasser, R. H., & Tjoumakaris, S. (2020). Effect of Shunting on Visual Outcomes and Headache in Patients with Idiopathic Intracranial Hypertension. *World Neurosurgery*, 142, e73-e80. <https://doi.org/10.1016/j.wneu.2020.05.186>
32. Deng, Y., Wang, Z., Chang, C., Lu, L., Lau, C. S., & Lu, Q. (2017). Th9 cells and IL-9 in autoimmune disorders: Pathogenesis and therapeutic potentials. *Human Immunology*, 78(2), 120-128. <https://doi.org/10.1016/j.humimm.2016.12.010>
33. Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nature Biotechnology*, 35(4), 316-319. <https://doi.org/10.1038/nbt.3820>
34. Ducros, A., & Biousse, V. (2015). Headache arising from idiopathic changes in CSF pressure. *The Lancet Neurology*, 14(6), 655-668. [https://doi.org/10.1016/S1474-4422\(15\)00015-0](https://doi.org/10.1016/S1474-4422(15)00015-0)
35. Edwards, L. J., Sharrack, B., Ismail, A., Tench, C. R., Gran, B., Dhungana, S., Brettschneider, J., Tumani, H., & Constantinescu, C. S. (2013). Increased levels of interleukins 2 and 17 in the cerebrospinal fluid of patients with idiopathic intracranial hypertension. *American Journal of Clinical and Experimental Immunology*, 2(3), 234-244.
36. Ekizoglu, E., Içoz, S., Tuzun, E., Birisik, O., Kocasoy-Orhan, E., Akman-Demir, G., & Baykan, B. (2012). Aquaporin-4 antibodies are not present in patients with idiopathic intracranial hypertension. *Cephalalgia: An International Journal of Headache*, 32(3), 198-202. <https://doi.org/10.1177/0333102411434167>
37. Fain, J. N. (2006). Release of interleukins and other inflammatory cytokines by human adipose tissue is enhanced in obesity and primarily due to the nonfat cells. *Vitamins and Hormones*, 74, 443-477. [https://doi.org/10.1016/S0083-6729\(06\)74018-3](https://doi.org/10.1016/S0083-6729(06)74018-3)

38. Franke, K. R., & Crowgey, E. L. (2020). Accelerating next generation sequencing data analysis: An evaluation of optimized best practices for Genome Analysis Toolkit algorithms. *Genomics & Informatics*, 18(1), e10. <https://doi.org/10.5808/GI.2020.18.1.e10>
39. Friedman, D. I. (2014). Papilledema and Idiopathic Intracranial Hypertension. *Continuum: Lifelong Learning in Neurology*, 20(4 Neuro-ophthalmology), 857. <https://doi.org/10.1212/01.CON.0000453314.75261.66>
40. Friedman, D. I., Liu, G. T., & Digre, K. B. (2013). Revised diagnostic criteria for the pseudotumor cerebri syndrome in adults and children. *Neurology*, 81(13), 1159-1165. <https://doi.org/10.1212/WNL.0b013e3182a55f17>
41. Fyrst, H., & Saba, J. D. (2010). An update on sphingosine-1-phosphate and other sphingolipid mediators. *Nature Chemical Biology*, 6(7), 489-497. <https://doi.org/10.1038/nchembio.392>
42. Geisbüsch, C., Herweh, C., Gumbinger, C., Ringleb, P. A., Möhlenbruch, M. A., & Nagel, S. (2021). Chronic intracranial hypertension after cerebral venous and sinus thrombosis – frequency and risk factors. *Neurological Research and Practice*, 3, 28. <https://doi.org/10.1186/s42466-021-00127-y>
43. Giuseffi, V., Wall, M., Siegel, P. Z., & Rojas, P. B. (1991). Symptoms and disease associations in idiopathic intracranial hypertension (pseudotumor cerebri): A case-control study. *Neurology*, 41(2 ( Pt 1)), 239-244. [https://doi.org/10.1212/wnl.41.2\\_part\\_1.239](https://doi.org/10.1212/wnl.41.2_part_1.239)
44. Grealey, J., Lannelongue, L., Saw, W.-Y., Marten, J., Méric, G., Ruiz-Carmona, S., & Inouye, M. (2022). The Carbon Footprint of Bioinformatics. *Molecular Biology and Evolution*, 39(3), msac034. <https://doi.org/10.1093/molbev/msac034>
45. Grech, O., Clouter, A., Mitchell, J. L., Alimajstorovic, Z., Ottridge, R. S., Yiangou, A., Roque, M., Tahrani, A. A., Nicholls, M., Taylor, A. E., Shaheen, F., Arlt, W., Lavery, G. G., Shapiro, K., Mollan, S. P., & Sinclair, A. J. (2021). Cognitive performance in idiopathic intracranial hypertension and relevance of intracranial pressure. *Brain Communications*, 3(3), fcab202. <https://doi.org/10.1093/braincomms/fcab202>
46. Green, L., Vinker, S., Amital, H., Amir, T., Bar-Dayana, Y., Levi, Y., & Schoenfeld, Y. (1995). Pseudotumor cerebri in systemic lupus erythematosus. *Seminars in Arthritis and Rheumatism*, 25(2), 103-108. [https://doi.org/10.1016/S0049-0172\(95\)80023-9](https://doi.org/10.1016/S0049-0172(95)80023-9)
47. Handley, J. D., Baruah, B. P., Williams, D. M., Horner, M., Barry, J., & Stephens, J. W. (2015). Bariatric surgery as a treatment for idiopathic intracranial hypertension: A systematic review. *Surgery for Obesity and Related Diseases*, 11(6), 1396-1403. <https://doi.org/10.1016/j.soard.2015.08.497>

48. Hermes, S. M., Miller, N. R., Waslo, C. S., Benes, S. C., & Tanne, E. (2020). Mortality among patients with idiopathic intracranial hypertension enrolled in the IH Registry. *Neurology*, 95(7), e921-e929. <https://doi.org/10.1212/WNL.00000000000009312>
49. Hornby, C., Mollan, S. P., Botfield, H., O'Reilly, M. W., & Sinclair, A. J. (2018). Metabolic Concepts in Idiopathic Intracranial Hypertension and Their Potential for Therapeutic Intervention. *Journal of Neuro-Ophthalmology: The Official Journal of the North American Neuro-Ophthalmology Society*, 38(4), 522-530. <https://doi.org/10.1097/WNO.0000000000000684>
50. Hornby, C., Mollan, S. P., Mitchell, J., Markey, K. A., Yangou, A., Wright, B. L. C., O'Reilly, M. W., & Sinclair, A. J. (2017). What Do Transgender Patients Teach Us About Idiopathic Intracranial Hypertension? *Neuro-Ophthalmology (Aeolus Press)*, 41(6), 326-329. <https://doi.org/10.1080/01658107.2017.1316744>
51. Hsu, H.-T., Cheng, H.-C., Hou, T.-W., Tzeng, Y.-S., Fuh, J.-L., Chen, S.-P., Chen, W.-T., Lee, W.-J., Pai, Y.-W., Lee, Y.-C., Lirng, J.-F., Wang, S.-J., & Wang, Y.-F. (2024). Idiopathic intracranial hypertension in Asians: A retrospective dual-center study. *The Journal of Headache and Pain*, 25(1), 144. <https://doi.org/10.1186/s10194-024-01852-w>
52. Johnston, I., & Morgan, M. K. (1991). A Familial Coincidence of Pseudotumor Cerebri and Communicating Hydrocephalus. *Neurosurgery*, 28(5), 727.
53. Kerty, E., Heuser, K., Indahl, U. G., Berg, P. R., Nakken, S., Lien, S., Omholt, S. W., Ottersen, O. P., & Nagelhus, E. A. (2013). Is the brain water channel aquaporin-4 a pathogenetic factor in idiopathic intracranial hypertension? Results from a combined clinical and genetic study in a Norwegian cohort. *Acta Ophthalmologica*, 91(1), 88-91. <https://doi.org/10.1111/j.1755-3768.2011.02231.x>
54. Kleinschmidt, J. J., Digre, K. B., & Hanover, R. (2000). Idiopathic intracranial hypertension: Relationship to depression, anxiety, and quality of life. *Neurology*, 54(2), 319-324. <https://doi.org/10.1212/wnl.54.2.319>
55. Kuehn, M. H., Mishra, R., Deonovic, B. E., Miller, K. N., McCormack, S. E., Liu, G. T., Kupersmith, M. J., Wall, M., & NORDIC IIHTT Study Group. (2019). Genetic Survey of Adult-Onset Idiopathic Intracranial Hypertension. *Journal of Neuro-Ophthalmology: The Official Journal of the North American Neuro-Ophthalmology Society*, 39(1), 50-55. <https://doi.org/10.1097/WNO.0000000000000648>
56. Kułakowska, A., Żendzian-Piotrowska, M., Baranowski, M., Konończuk, T., Drozdowski, W., Górski, J., & Bucki, R. (2010). Intrathecal increase of sphingosine 1-phosphate at early stage multiple sclerosis. *Neuroscience Letters*, 477(3), 149-152. <https://doi.org/10.1016/j.neulet.2010.04.052>
57. Kumar, G., Kalita, J., & Misra, U. K. (2009). Raised intracranial pressure in acute viral encephalitis. *Clinical Neurology and Neurosurgery*, 111(5), 399-406. <https://doi.org/10.1016/j.clineuro.2009.03.004>



58. Kung, J. T. Y., Colognori, D., & Lee, J. T. (2013). Long Noncoding RNAs: Past, Present, and Future. *Genetics*, 193(3), 651. <https://doi.org/10.1534/genetics.112.146704>
59. Li, W., Zhang, Z., Liu, X., Cheng, X., Zhang, Y., Han, X., Zhang, Y., Liu, S., Yang, J., Xu, B., He, L., Sun, L., Liang, J., & Shang, Y. (2017). The FOXN3-NEAT1-SIN3A repressor complex promotes progression of hormonally responsive breast cancer. *The Journal of Clinical Investigation*, 127(9), 3421-3440. <https://doi.org/10.1172/JCI94233>
60. Liang, H., Xu, C., Liu, T., Huang, S., & Hu, S. (2019). Spontaneous intracranial hypotension in Hashimoto's thyroiditis: A case report. *Medicine*, 98(18), e15476. <https://doi.org/10.1097/MD.00000000000015476>
61. Lucki, N. C., Li, D., & Sewer, M. B. (2011). Sphingosine-1-Phosphate Rapidly Increases Cortisol Biosynthesis and the Expression of Genes Involved in Cholesterol Uptake and Transport in H295R Adrenocortical Cells. *Molecular and Cellular Endocrinology*, 348(1), 165. <https://doi.org/10.1016/j.mce.2011.08.003>
62. Maloney, K. (2013). Idiopathic intracranial hypertension as an initial presentation of systemic lupus erythematosus. *BMJ Case Reports*, 2013, bcr2013010223. <https://doi.org/10.1136/bcr-2013-010223>
63. Markey, K. A., Uldall, M., Botfield, H., Cato, L. D., Miah, M. A. L., Hassan-Smith, G., Jensen, R. H., Gonzalez, A. M., & Sinclair, A. J. (2016). Idiopathic intracranial hypertension, hormones, and 11 $\beta$ -hydroxysteroid dehydrogenases. *Journal of Pain Research*, 9, 223-232. <https://doi.org/10.2147/JPR.S80824>
64. Marshall, C. R., Bick, D., Belmont, J. W., Taylor, S. L., Ashley, E., Dimmock, D., Jobanputra, V., Kearney, H. M., Kulkarni, S., Rehm, H., & Medical Genome Initiative. (2020). The Medical Genome Initiative: Moving whole-genome sequencing for rare disease diagnosis to the clinic. *Genome Medicine*, 12(1), 48. <https://doi.org/10.1186/s13073-020-00748-z>
65. Mokri, B. (2001). The Monro-Kellie hypothesis: Applications in CSF volume depletion. *Neurology*, 56(12), 1746-1748. <https://doi.org/10.1212/wnl.56.12.1746>
66. Neurath, M. F., & Finotto, S. (2016). IL-9 signaling as key driver of chronic inflammation in mucosal immunity. *Cytokine & Growth Factor Reviews*, 29, 93-99. <https://doi.org/10.1016/j.cytogfr.2016.02.002>
67. Nisar, H., Wajid, B., Shahid, S., Anwar, F., Wajid, I., Khatoon, A., Sattar, M. U., & Sadaf, S. (2021). Whole-genome sequencing as a first-tier diagnostic framework for rare genetic diseases. *Experimental Biology and Medicine*, 246(24), 2610. <https://doi.org/10.1177/15353702211040046>

68. Nithyanandam, S., Manayath, G. J., & Battu, R. R. (2008). Optic nerve sheath decompression for visual loss in intracranial hypertension: Report from a tertiary care center in South India. *Indian Journal of Ophthalmology*, 56(2), 115-120. <https://doi.org/10.4103/0301-4738.39115>
69. Parveen, F., Bender, D., Law, S.-H., Mishra, V. K., Chen, C.-C., & Ke, L.-Y. (2019). Role of Ceramidases in Sphingolipid Metabolism and Human Diseases. *Cells*, 8(12), 1573. <https://doi.org/10.3390/cells8121573>
70. Peng, S., Liu, J., Liang, C., Yang, L., & Wang, G. (2023). Aquaporin-4 in glymphatic system, and its implication for central nervous system disorders. *Neurobiology of Disease*, 179, 106035. <https://doi.org/10.1016/j.nbd.2023.106035>
71. Pettus, B. J., Bielawska, A., Spiegel, S., Roddy, P., Hannun, Y. A., & Chalfant, C. E. (2003). Ceramide Kinase Mediates Cytokine- and Calcium Ionophore-induced Arachidonic Acid Release\*. *Journal of Biological Chemistry*, 278(40), 38206-38213. <https://doi.org/10.1074/jbc.M304816200>
72. Piper, R. J., Kalyvas, A. V., Young, A. M., Hughes, M. A., Jamjoom, A. A., & Fouyas, I. P. (2015). Interventions for idiopathic intracranial hypertension. *Cochrane Database of Systematic Reviews*, 8. <https://doi.org/10.1002/14651858.CD003434.pub3>
73. Pollay, M., Hisey, B., Reynolds, E., Tomkins, P., Stevens, F. A., & Smith, R. (1985). Choroid plexus Na<sup>+</sup>/K<sup>+</sup>-activated adenosine triphosphatase and cerebrospinal fluid formation. *Neurosurgery*, 17(5), 768-772. <https://doi.org/10.1227/00006123-198511000-00007>
74. Rábano, M., Peña, A., Brizuela, L., Marino, A., Macarulla, J. M., Trueba, M., & Gómez-Muñoz, A. (2003). Sphingosine-1-phosphate stimulates cortisol secretion. *FEBS Letters*, 535(1), 101-105. [https://doi.org/10.1016/S0014-5793\(02\)03882-6](https://doi.org/10.1016/S0014-5793(02)03882-6)
75. Rashid, U., Wu, C., Shiller, J., Smith, K., Crowhurst, R., Davy, M., Chen, T.-H., Carvajal, I., Bailey, S., Thomson, S., & Deng, C. H. (2024). AssemblyQC: A Nextflow pipeline for reproducible reporting of assembly quality. *Bioinformatics (Oxford, England)*, 40(8), btae477. <https://doi.org/10.1093/bioinformatics/btae477>
76. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H. L., & ACMG Laboratory Quality Assurance Committee. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 17(5), 405-424. <https://doi.org/10.1038/gim.2015.30>

77. Robinson, P. N., Köhler, S., Oellrich, A., Project, S. M. G., Wang, K., Mungall, C. J., Lewis, S. E., Washington, N., Bauer, S., Seelow, D., Krawitz, P., Gilissen, C., Haendel, M., & Smedley, D. (2014). Improved exome prioritization of disease genes through cross-species phenotype comparison. *Genome Research*, 24(2), 340. <https://doi.org/10.1101/gr.160325.113>
78. Rodríguez-Boto, G., Rivero-Garvía, M., Gutiérrez-González, R., & Márquez-Rivas, J. (2015). Basic concepts about brain pathophysiology and intracranial pressure monitoring. *Neurología (English Edition)*, 30(1), 16-22. <https://doi.org/10.1016/j.nrleng.2012.09.002>
79. Roviezzo, F., Brancaleone, V., Gruttola, L. D., Vellecco, V., Bucci, M., D'Agostino, B., Cooper, D., Sorrentino, R., Perretti, M., & Cirino, G. (2011). Sphingosine-1-Phosphate Modulates Vascular Permeability and Cell Recruitment in Acute Inflammation In Vivo. *Journal of Pharmacology and Experimental Therapeutics*, 337(3), 830-837. <https://doi.org/10.1124/jpet.111.179168>
80. Samancı, B., Samancı, Y., Tüzün, E., Altıokka-Uzun, G., Ekizoğlu, E., İçöz, S., Şahin, E., Küçükali, C. İ., & Baykan, B. (2017). Evidence for potential involvement of pro-inflammatory adipokines in the pathogenesis of idiopathic intracranial hypertension. *Cephalalgia: An International Journal of Headache*, 37(6), 525-531. <https://doi.org/10.1177/0333102416650705>
81. Santos-Bueso, E., Porta-Etessam, J., Díaz-Valle, D., Benítez-del-Castillo, J. M., Gegúndez-Fernández, J. A., Vinuesa-Silva, J. M., & García-Sánchez, J. (2015). Hipertensión intracraneal idiopática en paciente con malformación de Chiari tipo I. *Archivos de la Sociedad Española de Oftalmología*, 90(4), 190-194. <https://doi.org/10.1016/j.ofal.2014.04.006>
82. Sawada, M., Suzumura, A., & Marunouchi, T. (1995). Cytokine network in the central nervous system and its roles in growth and differentiation of glial and neuronal cells. *International Journal of Developmental Neuroscience: The Official Journal of the International Society for Developmental Neuroscience*, 13(3-4), 253-264. [https://doi.org/10.1016/0736-5748\(94\)00076-f](https://doi.org/10.1016/0736-5748(94)00076-f)
83. Scott, T. R., Kronsten, V. T., Hughes, R. D., & Shawcross, D. L. (2013). Pathophysiology of cerebral oedema in acute liver failure. *World Journal of Gastroenterology: WJG*, 19(48), 9240. <https://doi.org/10.3748/wjg.v19.i48.9240>
84. Sm, A., & Nj, R. (2001). Cytokines and acute neurodegeneration. *Nature Reviews Neuroscience*, 2(10). <https://doi.org/10.1038/35094583>
85. Smedley, D., Köhler, S., Czeschik, J. C., Amberger, J., Bocchini, C., Hamosh, A., Veldboer, J., Zemojtel, T., & Robinson, P. N. (2014). Walking the interactome for candidate prioritization in exome sequencing studies of Mendelian diseases. *Bioinformatics*, 30(22), 3215-3222. <https://doi.org/10.1093/bioinformatics/btu508>



86. Soden, S. E., Saunders, C. J., Willig, L. K., Farrow, E. G., Smith, L. D., Petrikin, J. E., LePichon, J.-B., Miller, N. A., Thiffault, I., Dinwiddie, D. L., Twist, G., Noll, A., Heese, B. A., Zellmer, L., Atherton, A. M., Abdelmoity, A. T., Safina, N., Nyp, S. S., Zuccarelli, B., ... Kingsmore, S. F. (2014). Effectiveness of exome and genome sequencing guided by acuity of illness for diagnosis of neurodevelopmental disorders. *Science Translational Medicine*, 6(265), 265ra168. <https://doi.org/10.1126/scitranslmed.3010076>
87. Sorrells, S. F., & Sapolsky, R. M. (2007). An inflammatory review of glucocorticoid actions in the CNS. *Brain, Behavior, and Immunity*, 21(3), 259-272. <https://doi.org/10.1016/j.bbi.2006.11.006>
88. Spennato, P., Mirone, G., Nastro, A., Buonocore, M. C., Ruggiero, C., Trischitta, V., Aliberti, F., & Cinalli, G. (2011). Hydrocephalus in Dandy-Walker malformation. *Child's Nervous System: ChNS: Official Journal of the International Society for Pediatric Neurosurgery*, 27(10), 1665-1681. <https://doi.org/10.1007/s00381-011-1544-4>
89. Splinter, K., Adams, D. R., Bacino, C. A., Bellen, H. J., Bernstein, J. A., Cheatle-Jarvela, A. M., Eng, C. M., Esteves, C., Gahl, W. A., Hamid, R., Jacob, H. J., Kikani, B., Koeller, D. M., Kohane, I. S., Lee, B. H., Loscalzo, J., Luo, X., McCray, A. T., Metz, T. O., ... Undiagnosed Diseases Network. (2018). Effect of Genetic Diagnosis on Patients with Previously Undiagnosed Disease. *The New England Journal of Medicine*, 379(22), 2131-2139. <https://doi.org/10.1056/NEJMoa1714458>
90. Steiner, L. A., & Andrews, P. J. D. (2006). Monitoring the injured brain: ICP and CBF. *British Journal of Anaesthesia*, 97(1), 26-38. <https://doi.org/10.1093/bja/ael110>
91. Stiebel-Kalish, H., Eyal, S., & Steiner, I. (2013). The role of aquaporin-1 in idiopathic and drug-induced intracranial hypertension. *Medical Hypotheses*, 81(6), 1059-1062. <https://doi.org/10.1016/j.mehy.2013.10.002>
92. Sudhakar, P. (2021). Commentary: The role of inflammation in idiopathic intracranial hypertension. *Indian Journal of Ophthalmology*, 69(6), 1506. [https://doi.org/10.4103/ijo.IJO\\_252\\_21](https://doi.org/10.4103/ijo.IJO_252_21)
93. Takeshita, H., Kitano, M., Iwasaki, T., Kitano, S., Tsunemi, S., Sato, C., Sekiguchi, M., Azuma, N., Miyazawa, K., Hla, T., & Sano, H. (2012). Sphingosine 1-phosphate (S1P)/S1P receptor 1 signaling regulates receptor activator of NF- $\kappa$ B ligand (RANKL) expression in rheumatoid arthritis. *Biochemical and Biophysical Research Communications*, 419(2), 154-159. <https://doi.org/10.1016/j.bbrc.2012.01.103>
94. Tam, E. K., & Gilbert, A. L. (2019). Spontaneous cerebrospinal fluid leak and idiopathic intracranial hypertension. *Current Opinion in Ophthalmology*, 30(6), 467-471. <https://doi.org/10.1097/ICU.0000000000000603>

95. Tammimies, K., Marshall, C. R., Walker, S., Kaur, G., Thiruvahindrapuram, B., Lionel, A. C., Yuen, R. K. C., Uddin, M., Roberts, W., Weksberg, R., Woodbury-Smith, M., Zwaigenbaum, L., Anagnostou, E., Wang, Z., Wei, J., Howe, J. L., Gazzellone, M. J., Lau, L., Sung, W. W. L., ... Fernandez, B. A. (2015). Molecular Diagnostic Yield of Chromosomal Microarray Analysis and Whole-Exome Sequencing in Children With Autism Spectrum Disorder. *JAMA*, 314(9), 895-903. <https://doi.org/10.1001/jama.2015.10078>
96. Tariq, A., Aguilar-Salinas, P., Hanel, R. A., Naval, N., & Chmayssani, M. (2017). The role of ICP monitoring in meningitis. *Neurosurgical Focus*, 43(5), E7. <https://doi.org/10.3171/2017.8.FOCUS17419>
97. Tellería-Oriols, J. J., López-Hernández, S., Vidriales-Vicente, I., & Rodríguez-Arias, C. A. (2023). Association between RS3763040 polymorphism of the AQP4 and idiopathic intracranial hypertension in a Spanish Caucasian population. *Translational Neuroscience*, 14(1), 20220309. <https://doi.org/10.1515/tnsci-2022-0309>
98. Thaller, M., Homer, V., Hyder, Y., Yiangou, A., Liczkowski, A., Fong, A. W., Virdee, J., Piccus, R., Roque, M., Mollan, S. P., & Sinclair, A. J. (2023). The idiopathic intracranial hypertension prospective cohort study: Evaluation of prognostic factors and outcomes. *Journal of Neurology*, 270(2), 851-863. <https://doi.org/10.1007/s00415-022-11402-6>
99. Tomlinson, J. W., Moore, J., Cooper, M. S., Bujalska, I., Shahmanesh, M., Burt, C., Strain, A., Hewison, M., & Stewart, P. M. (2001). Regulation of expression of 11beta-hydroxysteroid dehydrogenase type 1 in adipose tissue: Tissue-specific induction by cytokines. *Endocrinology*, 142(5), 1982-1989. <https://doi.org/10.1210/endo.142.5.8168>
100. Tomlinson, J. W., Walker, E. A., Bujalska, I. J., Draper, N., Lavery, G. G., Cooper, M. S., Hewison, M., & Stewart, P. M. (2004). 11beta-hydroxysteroid dehydrogenase type 1: A tissue-specific regulator of glucocorticoid response. *Endocrine Reviews*, 25(5), 831-866. <https://doi.org/10.1210/er.2003-0031>
101. Toscano, S., Lo Fermo, S., Reggio, E., Chisari, C. G., Patti, F., & Zappia, M. (2021). An update on idiopathic intracranial hypertension in adults: A look at pathophysiology, diagnostic approach and management. *Journal of Neurology*, 268(9), 3249-3268. <https://doi.org/10.1007/s00415-020-09943-9>
102. Toshniwal, S. S., Kinkar, J., Chadha, Y., Khurana, K., Reddy, H., Kadam, A., & Acharya, S. (2024). Navigating the Enigma: A Comprehensive Review of Idiopathic Intracranial Hypertension. *Cureus*, 16(3), e56256. <https://doi.org/10.7759/cureus.56256>
- Traviesa, D. C., Schwartzman, R. J., Glaser, J. S., & Savino, P. (1976). Familial benign intracranial hypertension. *Journal of Neurology, Neurosurgery & Psychiatry*, 39(5), 420-423. <https://doi.org/10.1136/jnnp.39.5.420>

103. Valcamonico, F., Arcangeli, G., Consoli, F., Nonnis, D., Grisanti, S., Gatti, E., Berruti, A., & Ferrari, V. (2014). Idiopathic intracranial hypertension: A possible complication in the natural history of advanced prostate cancer. *International Journal of Urology: Official Journal of the Japanese Urological Association*, 21(3), 335-337. <https://doi.org/10.1111/iju.12273>
104. Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, 43(1110), 11.10.1-11.10.33. <https://doi.org/10.1002/0471250953.bi1110s43>
105. Wardman, J. H., Jensen, M. N., Andreassen, S. N., Styrisshave, B., Wilhjelm, J. E., Sinclair, A. J., & MacAulay, N. (2023). Modelling idiopathic intracranial hypertension in rats: Contributions of high fat diet and testosterone to intracranial pressure and cerebrospinal fluid production. *Fluids and Barriers of the CNS*, 20(1), 44. <https://doi.org/10.1186/s12987-023-00436-1>
106. Zemojtel, T., Köhler, S., Mackenroth, L., Jäger, M., Hecht, J., Krawitz, P., Graul-Neumann, L., Doelken, S., Ehmke, N., Spielmann, M., Øien, N. C., Schweiger, M. R., Krüger, U., Frommer, G., Fischer, B., Kornak, U., Flöttmann, R., Ardeshirdavani, A., Moreau, Y., ... Robinson, P. N. (2014). Effective diagnosis of genetic disease by computational phenotype analysis of the disease-associated genome. *Science Translational Medicine*, 6(252). <https://doi.org/10.1126/scitranslmed.3009262>
107. Zhou, Y., Kathiresan, N., Yu, Z., Rivera, L. F., Yang, Y., Thimma, M., Manickam, K., Chebotarov, D., Mauleon, R., Chougule, K., Wei, S., Gao, T., Green, C. D., Zuccolo, A., Xie, W., Ware, D., Zhang, J., McNally, K. L., & Wing, R. A. (2024). A high-performance computational workflow to accelerate GATK SNP detection across a 25-genome dataset. *BMC Biology*, 22, 13. <https://doi.org/10.1186/s12915-024-01820-5>

## **Anexo**

### **Anexo 1. Consentimiento informado de participación en el estudio**

[https://liveuem-my.sharepoint.com/:f/g/personal/22342452\\_live\\_uem\\_es/EvKs8pi34rLsipz0iLjcBsX4JnWSkoJJERB7qfaEUvw?e=UcWxOU](https://liveuem-my.sharepoint.com/:f/g/personal/22342452_live_uem_es/EvKs8pi34rLsipz0iLjcBsX4JnWSkoJJERB7qfaEUvw?e=UcWxOU)

### **Anexo 2. Aprobación del Comité de ética y Comisión de Investigación de la Escuela de Doctorado e Investigación de la Universidad Europea**

[https://liveuem-my.sharepoint.com/:f/g/personal/22342452\\_live\\_uem\\_es/EhIKfXsuRgllkIco3ikkjmcBzIzQPJtISN28\\_GUA72HMWg?e=zlqaqS](https://liveuem-my.sharepoint.com/:f/g/personal/22342452_live_uem_es/EhIKfXsuRgllkIco3ikkjmcBzIzQPJtISN28_GUA72HMWg?e=zlqaqS)

### **Anexo 3. Archivo de configuración del entorno Conda**

[https://github.com/jofereX/TFM-HII-WGS/blob/main/TFM\\_nextflowpipeline.yaml](https://github.com/jofereX/TFM-HII-WGS/blob/main/TFM_nextflowpipeline.yaml)

### **Anexo 4. Links de descarga a las bases de datos y archivos de referencia empleados**

[https://liveuem-my.sharepoint.com/:w/g/personal/22342452\\_live\\_uem\\_es/EWG0tBao-S5IoMOxT2vAicoB9OyT5c7JMNNBLsBp4\\_mSFQ?e=Q2dYtk](https://liveuem-my.sharepoint.com/:w/g/personal/22342452_live_uem_es/EWG0tBao-S5IoMOxT2vAicoB9OyT5c7JMNNBLsBp4_mSFQ?e=Q2dYtk)

### **Anexo 5. Script de nextflow de la pipeline completa de análisis genómico de código**

[https://github.com/jofereX/TFM-HII-WGS/blob/main/nextflow\\_HII\\_complete\\_pipeline.nf](https://github.com/jofereX/TFM-HII-WGS/blob/main/nextflow_HII_complete_pipeline.nf)

### **Anexo 6. Script de nextflow del proceso de anotación y priorización de variantes con Exomiser**

[https://github.com/jofereX/TFM-HII-WGS/blob/main/process\\_runExomiser.nf](https://github.com/jofereX/TFM-HII-WGS/blob/main/process_runExomiser.nf)

### **Anexo 7. Archivo de configuración empleado por exomiser para la anotación y priorización de variantes a través de *hivePrioritiser***

[https://github.com/jofereX/TFM-HII-WGS/blob/main/HII\\_analysis\\_general\\_nofilter.yml](https://github.com/jofereX/TFM-HII-WGS/blob/main/HII_analysis_general_nofilter.yml)

### **Anexo 8. Archivo de configuración empleado por exomiser para la anotación y priorización de variantes a través de un panel de genes de interés**

[https://github.com/jofereX/TFM-HII-WGS/blob/main/HII\\_analysis\\_panel\\_AQP\\_GWAS\\_nofilter.yml](https://github.com/jofereX/TFM-HII-WGS/blob/main/HII_analysis_panel_AQP_GWAS_nofilter.yml)

### **Anexo 9. Archivo Rmarkdown empleado para el procesamiento y análisis estadístico de las variantes priorizadas**

[https://github.com/jofereX/TFM-HII-WGS/blob/main/Procesamiento\\_Variantes\\_TFM.Rmd](https://github.com/jofereX/TFM-HII-WGS/blob/main/Procesamiento_Variantes_TFM.Rmd)

### **Anexo 10. Tabla con las 92 variantes anotadas localizadas en regiones no codificantes del genoma que presentaron una mayor asociación con la HI en nuestra población de estudio**

[https://liveuem-my.sharepoint.com/:f/g/personal/22342452\\_live\\_uem\\_es/ElgyqLKg22tGhgWXmIY\\_RGYBE8SrM9zkvfq6vSR8pd5\\_A?e=v29d3h](https://liveuem-my.sharepoint.com/:f/g/personal/22342452_live_uem_es/ElgyqLKg22tGhgWXmIY_RGYBE8SrM9zkvfq6vSR8pd5_A?e=v29d3h)

## Declaración del uso de herramientas IA

Para la realización de este trabajo se empleó ChatGPT-4 para las siguientes tareas:

1. Busqueda y sugerencia de bibliografía. Ante la gran cantidad de publicaciones y con el objetivo de optimizar la búsqueda bibliográfica, especialmente en las primeras etapas de este trabajo, se empleó ChatGPT-4 para recibir recomendaciones acerca de temas concretos ayudando así a identificar publicaciones de interés. Algunos de los prompts utilizados fueron: “Recomiendame artículos científicos que hablen acerca de las bases genéticas de la hipertensión intracraneal idiopática” o “Artículos que traten la implicación de la IL9 en la inflamación crónica o enfermedades autoinmunes”
2. Traducción del lenguaje. Algunos de los artículos y publicaciones empleados en este trabajo fueron total o parcialmente traducidos usando ChatGPT-4 con el objetivo de facilitar la comprensión de fuentes en idiomas distintos al castellano. Los prompts utilizados fueron en su inmensa mayoría copias literales de secciones concretas de los artículos de interés
3. Parafraseo y reformulación de ideas originales. Durante la redacción de este trabajo en empleó ChatGPT-4 para parafrasear de manera eficiente ideas complejas y reformular textos de manera que se ajustaran al tono del trabajo. Indirectamente también fue utilizada para la corrección de faltas de ortografía. Por ejemplo, se utilizó el siguiente prompt:  
“necesito que des fluidez y corrija este párrafo: *Todos estos resultados, junto con la identificación de variantes poco frecuentes significativamente correlacionadas con el grupo de pacientes afectados por HII en la región codificante de los genes HLA, lo que modifica su (zona en la que se une a antígenos) así como en zonas intrónicas del gen IL9, respaldan la hipótesis de que la HII es una condición producida por alteraciones inmunológicas, donde la inflamación podría contribuir al desarrollo de la enfermedad, en lugar de ser causada únicamente por alteraciones en la producción y drenaje del LCR*”
4. Escritura de código en lenguaje R y Nextflow DSL. Para el análisis de datos genómicos y la automatización de ciertos procesos fue necesario escribir y ejecutar código en lenguajes de programación como R y Nextflow DSL, en este contexto, ChatGPT-4 desempeñó un papel muy importante como asistente en la escritura del código corrigiendo o señalando errores y en definitiva, facilitando la creación de scripts. Los prompts utilizados fueron en su mayoría bloques de código que generaban errores.