

Estudio Genómico de Schwannomas Esporádicos y Schwannomas asociados a Neurofibromatosis tipo 2

Nombre del estudiante: José Antonio Ramos Fernández

Plan de estudios: Máster Universitario en Bioinformática y Bioestadística

Área de trabajo final: Bioinformática translacional, análisis de datos y genómica del cáncer

Nombre Consultor/a: Elisabeth Castellanos Pérez

Nombre Profesor/a responsable de la asignatura: Carles Ventura Royo

Fecha de entrega: 05/06/2019



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

FICHA DEL TRABAJO FINAL

Título del trabajo:	<i>Estudio Genómico de Schwannomas Esporádicos y Schwannomas asociados a Neurofibromatosis tipo 2</i>
Nombre del autor:	<i>José Antonio Ramos Fernández</i>
Nombre del consultor/a:	<i>Elisabeth Castellanos Pérez</i>
Nombre del PRA:	<i>Carles Ventura Royo</i>
Fecha de entrega (mm/aaaa):	05/2019
Titulación:	<i>Máster Universitario en Bioinformática y Bioestadística</i>
Área del Trabajo Final:	<i>Bioinformática translacional, análisis de datos y genómica del cáncer</i>
Idioma del trabajo:	<i>Castellano</i>
Palabras clave	<i>Schwannoma, Neurofibromatosis tipo 2, Whole Exome Sequencing</i>

Resumen del Trabajo (máximo 250 palabras): *Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.*

Los schwannomas son tumores benignos formados por células de Schwann que rodean a los nervios mielínicos. Pueden surgir de manera esporádica o como consecuencia de ciertos síndromes genéticos (como neurofibromatosis tipo 2). Hoy día no existe ningún tratamiento eficaz para esta patología y los mecanismos moleculares implicados no están completamente entendidos. Se han encontrado mutaciones en el gen NF2 en diferentes tipos de schwannomas, por lo que es posible que tenga un papel central en su desarrollo, aunque otros genes podrían estar implicados. Las técnicas de NGS, como *whole exome sequencing*, pueden ser de gran utilidad para la búsqueda de posibles marcadores moleculares y dianas terapéuticas que puedan mejorar el diagnóstico y el tratamiento de los pacientes, así como nuestro entendimiento de los mecanismos moleculares implicados.

El objetivo principal de este trabajo fue detectar genes mutados que puedan favorecer el crecimiento de este tipo de tumores, utilizando datos obtenidos mediante WES de muestras de schwannomas esporádicos y asociados a neurofibromatosis tipo 2. Para ello, se utilizaron diferentes herramientas bioinformáticas mediante un protocolo diseñado específicamente para este trabajo.

Los resultados mostraron que los schwannomas esporádicos pueden requerir de más mutaciones para su desarrollo y que los genes mutados en ambos tipos de schwannomas participan en procesos biológicos diferentes, lo cual sugiere que cada tipo de schwannoma podría seguir mecanismos moleculares diferentes para su desarrollo. Además, el gen NF2, se encontró mutado en la mayoría de las muestras indicando que puede tener un papel central, aunque otros genes podrían estar implicados (como PABPC3).

Abstract (in English, 250 words or less):

Schwannomas are benign tumors composed by Schwann cells of nerves sheaths. They can arise spontaneously or as consequence of several genetic syndromes (for example neurofibromatosis type 2). Today there is not a completely effective treatment for this pathology and the molecular pathways involve are not fully understood. Mutations in NF2 gene have been found in different types of schwannomas and it could have a main role in the development of these tumors, although other genes may be implicated too. NGS technology, as for example, the whole exome sequencing, can be of great help in the seek for news molecular and therapeutic targets, improving the diagnostic and treatment of patients as well as our understanding of molecular pathways involved.

The principal aim of this project was to find mutated genes related with the growing of these kinds of tumors, using data obtained from WES of sporadic and neurofibromatosis type 2 related schwannomas samples. Different bioinformatics tools were used, and a specific pipeline was designed to perform the analysis.

Results suggest that sporadic schwannomas could need a greater number of mutations for their development and mutated gene found in sporadic and neurofibromatosis type 2 related schwannomas are involved in different biological process, suggesting that distinct molecular pathways are implicated in their development. In addition, NF2 gene was found altered in most of the samples showing it can have a main role, although other genes could have an important function, as PABPC3.

Índice

1. Introducción	5
1.1 Contexto y justificación del Trabajo	5
1.2 Objetivos del Trabajo	8
1.3 Enfoque y método seguido	9
1.4 Planificación del Trabajo	11
1.5 Breve resumen de productos obtenidos	12
1.6 Breve descripción de los otros capítulos de la memoria	14
2. Materiales y métodos	14
2.1 Obtención de los datos	14
2.2 Análisis	14
3. Resultados	18
3.1. Número de mutaciones	18
3.2 Rutas metabólicas enriquecidas	20
4. Discusión	22
5. Conclusiones	25
6. Glosario	28
7. Bibliografía	30
8. Anexos	35

Lista de figuras

Figura 1. Diagrama de Gantt con la planificación temporal del proyecto.	13
Figura 2. Número de mutaciones por cada grupo de Schwannoma.....	19
Figura 3. Perfil genético de los dos grupos de Schwannomas.	19
Figura 4. Rutas metabólicas enriquecidas en Schwannomas asociados a neurofibromatosis tipo 2.	21
Figura 5. Diez rutas metabólicas más enriquecidas de los Schwannomas esporádicos. 21	
Figura 6. Rutas metabólicas enriquecidas en los Schwannomas esporádicos.....	38

Lista de tablas

Tabla 1. Genes mutados más frecuentemente en cada grupo de schwannomas.	20
Tabla 2. Genes mutados previamente relacionados con Schwannomas.....	24
Tabla 3. Genes mutados en cada grupo de Schwannomas.	39

1. Introducción

1.1 Contexto y justificación del Trabajo

Los schwannomas son tumores benignos formados por células de Schwann que rodean los nervios mielínicos. En la mayoría de las ocasiones ocurren de forma aislada y esporádica, pero también, pueden presentarse como consecuencia de un síndrome hereditario, como puede ser el caso de neurofibromatosis tipo 2 (NF2), schwannomatosis o complejo de Carney. Están clasificados como tumores de grado I por la Organización Mundial de la Salud (OMS)¹ por lo que, muy raramente, sufren transformaciones malignas, casi siempre asociadas con algún síndrome genético.

Principalmente se manifiestan en adultos, con igual incidencia tanto en hombres como en mujeres. Suponen un 8.2% de todos los tumores del sistema nervioso. Su localización más común son los nervios periféricos y algunos nervios craneales como la rama vestibular del octavo nervio craneal o el trigémino, también se pueden encontrar en las raíces sensitivas de los nervios espinales. Raramente aparecen en múltiples sitios, salvo en caso de algunos síndromes genéticos como neurofibromatosis tipo 2.

Son tumores encapsulados que crecen alrededor de los nervios, por lo que la mayoría son asintomáticos. Pero, en algunos casos, pueden presentar sintomatología por compresión, provocando pérdida de función y sensibilidad de los nervios. Especialmente en los craneales, donde pueden originar sordera, pérdida de visión, tinnitus y vértigo. Si se produce cerca de algún componente óseo pueden debilitar la estructura del hueso, llegando a causar fracturas.²⁻⁵

Se han encontrado mutaciones en el gen NF2 tanto en schwannomas esporádicos como en schwannomas asociados a neurofibromatosis tipo 2 (síndrome hereditario provocado por la pérdida de función del gen NF2⁶) por lo que es posible que NF2 tenga un papel central en el desarrollo de estos tumores. Aunque es posible que otros eventos genéticos o epigenéticos pueden ser necesarios para el desarrollo y crecimiento de los schwannomas⁷⁻¹³. El gen NF2 codifica para una proteína llamada merlina, que se encarga de unir el citoesqueleto a las proteínas de la membrana plasmática y contribuye

a la inhibición de la proliferación por contacto¹⁴. Mutaciones en este gen también han sido asociadas con otros tipos de tumores (hígado, tiroides o tracto urinario, por ejemplo)¹⁵.

Hoy en día, la cirugía y la radioterapia son las técnicas más utilizadas para el tratamiento de los schwannomas. En el caso de schwannomas asociados a neurofibromatosis tipo 2, la radioterapia es menos efectiva, debido a que los tumores pueden surgir en múltiples localizaciones y algunos de ellos pueden no recibir la dosis de radiación adecuada. Por otro lado, durante la cirugía, dependiendo del tamaño del tumor, la localización y de la experiencia del cirujano, el nervio auditivo y el nervio facial pueden verse afectados, pudiendo causar pérdida de audición y parálisis facial. Hasta la fecha, aunque se han realizado avances, no existe un tratamiento basado en quimioterapéuticos completamente eficaz^{7,16}.

Por lo tanto, debido a que nuestro entendimiento de las alteraciones moleculares de los schwannomas es limitado y, además, hoy día no existe un tratamiento eficaz para ellos, se hace necesaria la búsqueda de posibles marcadores moleculares y dianas terapéuticas que puedan mejorar el diagnóstico y el tratamiento de los pacientes. En este aspecto pueden entrar en juego las técnicas de secuenciación de nueva generación (NGS).

Desde la secuenciación de Sanger, que sólo permite la secuenciación de un pequeño fragmento de ADN, con un elevado coste económico, la tecnología ha avanzado enormemente en la última década, permitiendo la secuenciación simultánea de un gran número de pequeños fragmentos con una reducción importante en el precio y el tiempo invertidos. Las tecnologías de NGS han permitido la secuenciación del genoma completo de una persona en cuestión de días, lo cual permite su uso en la práctica clínica¹⁷. Actualmente, es posible encontrar mutaciones en la secuencia de ADN que podrían estar relacionadas con ciertas patologías^{18,19}, incluido el cáncer²⁰, lo que puede suponer una mejora en el diagnóstico y el tratamiento de estas enfermedades.

Brevemente, el proceso de secuenciación seguido por estas técnicas consiste en la rotura del ADN en pequeños fragmentos que, mediante adaptadores son ligados a una superficie sólida. A continuación, se procede a la amplificación de las cadenas con el objetivo de aumentar la señal emitida durante la secuenciación. Por último, se van a

añadiendo los distintos dNTPs marcados mediante un fluoróforo diferente. A medida que se van uniendo los nucleótidos por complementariedad a la cadena molde, la señal del fluoróforo es detectada para identificar el nucleótido. En algunos casos, como la tecnología de Ion Torrent, los dNTPs son detectados mediante un cambio en el pH producido por una cascada enzimática que se activa cuando un nucleótido se une a la cadena molde.²¹

A pesar de los grandes avances en las técnicas de secuenciación, obtener la secuencia de un genoma completo sigue siendo un proceso largo y costoso que genera una gran cantidad de información para analizar. Esto ha llevado al desarrollo de ciertas técnicas, más baratas y rápidas, que permiten la secuenciación de zonas específicas del genoma, entre ellas se encuentran los paneles de genes (utilizados para secuenciar un conjunto de genes concretos) y *whole exome sequencing* (secuenciación del exoma completo).

La técnica de *whole exome sequencing* (WES) permite la secuenciación de la región codificante del genoma (exoma), que constituye en torno al 1% del genoma (unas 30Mb). Existen diferentes técnicas que permiten capturar y amplificar el exoma para facilitar su secuenciación, las cuales se describen a continuación²²:

- Hibridación en fase sólida: utiliza sondas complementarias a las secuencias de interés fijadas a una superficie sólida o filtros. Una vez se ha producido la hibridación, los fragmentos de ADN que no se han unido a las sondas son eliminados, quedando los fragmentos que corresponden a las zonas que se desean secuenciar.
- Hibridación en fase líquida: es similar a la anterior, salvo que las sondas no están unidas a una superficie sólida, en este caso, están unidas con una molécula de biotina. Tras la hibridación, las sondas se unen a una molécula de streptavidina magnética y se separan del ADN que no se ha unido a ellas, quedando el ADN enriquecido listo para secuenciar.
- Técnica de la polimerasa: en esta tecnología se utilizan sondas que actúan, a la vez, como cebadores en una reacción en cadena de la polimerasa (PCR), permitiendo, en un único paso la captura y amplificación de las secuencias de interés.

Se estima que el 85% de las mutaciones causantes de enfermedades se producen en el exoma, por lo que han surgido numerosos artículos en los que se han asociado diferentes alteraciones genéticas con distintos tipos de patologías utilizando WES^{23,24}. Gran cantidad de estudios previos ya han demostrado que la secuenciación del exoma puede ser útil para detectar posibles mutaciones dañinas en diferentes tipos de cáncer²⁵⁻³⁰, incluidos schwannomas esporádicos^{13,31}, lo que podría ayudar a mejorar el diagnóstico y tratamiento de los pacientes. Aunque hasta la fecha no se ha realizado ningún estudio de este tipo con schwannomas asociados a neurofibromatosis tipo 2.

En este estudio se pretende, utilizando datos obtenidos por WES de schwannomas esporádicos y schwannomas asociados a neurofibromatosis tipo 2:

1. Detectar genes mutados que ayuden a entender las alteraciones moleculares que influyen en el desarrollo de este tipo de tumores.
2. Comprobar si existen diferencias a nivel genético entre ambos tipos de schwannomas.

1.2 Objetivos del Trabajo

El proyecto consta de una serie de objetivos generales (A-B) divididos, a su vez, en objetivos específicos (1-5), los cuales se detallan a continuación:

A. Detectar posibles genes mutados en un conjunto de datos de 9 schwannomas de pacientes con neurofibromatosis tipo 2 y 22 schwannomas esporádicos obtenidos por WES.

1. Poner a punto un protocolo para analizar los datos obtenidos por WES.
2. Identificar las variantes de genes mutados utilizando el protocolo desarrollado en el objetivo anterior.

B. Comparar los resultados de ambos tipos de schwannomas.

3. Comprobar estadísticamente si existen diferencias entre el número de mutaciones por megabase de los schwannomas esporádicos y los schwannomas asociados a neurofibromatosis tipo 2.
4. Comparar los genes mutados y el tipo de mutación que aparecen más frecuentemente en cada tipo de schwannoma.

5. Observar si existen diferencias en las rutas metabólicas en las que participan los genes mutados en cada grupo de schwannomas.

1.3 Enfoque y método seguido

Tal y como se ha mencionado anteriormente, el uso de la secuenciación del exoma completo puede ser muy útil para la detección de genes mutados en muestras de schwannomas y así, ayudar a esclarecer los mecanismos moleculares que influyen en su desarrollo. El análisis de toda la información obtenida en experimentos de WES se compone de 4 etapas principales, las cuales se explican a continuación³²:

- *Mapping*: alineamiento de las secuencias con un genoma de referencia.
- *Variant calling*: identificación de las variaciones que existen entre la muestra secuenciada y el genoma de referencia. Existen varios tipos de variaciones que se pueden encontrar: cambios en un único nucleótido (SNVs), en múltiples nucleótidos, inserciones y deleciones (INDELs), cambios estructurales (SVs) y cambios en el número de copias (CNVs). Los experimentos de WES son muy efectivos en la detección de variantes de tipo SNV, pero pierden efectividad en la detección de variantes que implican un mayor número de nucleótidos.
- *Variant annotation*: consiste en asignar información biológica a las variaciones detectadas para establecer un enlace con el fenotipo del paciente. Se utilizan diferentes algoritmos y bases de datos para proporcionar información sobre: el gen en que se encuentran, el tipo de mutación, (sinónimas, sin sentido, cambio de sentido, *splicing*, cambio de la fase de lectura, etc.), frecuencia alélica en la población, efecto que causan en la secuencia proteica (benigno, dañino, etc.), conservación de la secuencia entre diferentes organismos o si existe alguna asociación conocida con patologías.
- *Priorización*: durante el proceso de *variant calling*, pueden detectarse hasta 100.000 variantes o más, de las cuales, la mayoría serán neutras o falsos positivos. En esta etapa, el objetivo es seleccionar aquellas variantes que tienen una alta probabilidad de ser dañinas o ser causante del fenotipo.

Las técnicas de NGS como WES generan gran cantidad de información, por lo que se necesitan de herramientas bioinformáticas que ayuden a manejar tal cantidad de datos y

a filtrar aquellos que no son útiles o no cumplen con los criterios de calidad mínimos. Existen disponibles gran cantidad de softwares que permiten realizar el alineamiento con el genoma de referencia, la identificación de variantes en la secuencia de ADN y la anotación de estas variantes^{33,34}.

El objetivo principal de este trabajo es detectar los genes mutados con alta probabilidad de ser dañinos en diferentes muestras de schwannomas, para lo que será necesario el uso de varias herramientas bioinformáticas con las que realizar las etapas descritas anteriormente. Se han utilizado los softwares *Burrows-Wheeler Aligner (BWA)*³⁵, *Strelka*³⁶, *Annovar*³⁷ y *SAMtools*³⁸ para realizar las etapas de *mapping*, *variant calling*, *variant annotation* y priorización respectivamente.

BWA es una herramienta ampliamente utilizada por la comunidad científica para realizar alineamientos con un genoma de referencia, que ha demostrado tener una alta velocidad y sensibilidad³⁹. Por otra parte, en la bibliografía existen varias revisiones en las que se comparan diferentes herramientas donde se puede ver que Strelka es una de las que presentan mayor sensibilidad y especificidad a la hora de identificar las alteraciones presentes en una muestra^{40,41}. Annovar, por su parte, es un programa ampliamente utilizado por la comunidad científica para la anotación de variantes, el cual dispone de acceso a multitud de bases de datos públicas, facilitando una rápida anotación de las mismas³⁷. SAMtools⁴² es un set de utilidades muy popular que se utiliza para la manipulación de archivos con el formato *.vcf* (formato de archivo más comúnmente utilizado para almacenar las variaciones de una secuencia y sus anotaciones⁴³) y que, entre otras funciones, permite realizar la priorización con unas pocas líneas de código.

Como base para el proceso de análisis, se utilizará un protocolo elaborado previamente por el grupo de investigación que incluye los pasos de *mapping*, *variant calling* y *variant annotation*, que se debió adaptar para el tipo de análisis que se va a realizar en este proyecto.

1.4 Planificación del Trabajo

El proyecto se ha ejecutado, principalmente, en 3 fases, cuyas tareas están descritas a continuación:

1. Desarrollar el protocolo a utilizar. Duración 4 semanas.
 - a. Revisar el protocolo utilizado por el grupo de investigación para la realización de análisis de datos de secuenciación por WES.
 - b. Realizar las modificaciones necesarias para realizar un análisis específico de los schwannomas esporádicos y asociados a neurofibromatosis tipo 2.
 - c. Realizar las pruebas necesarias para comprobar que el protocolo funciona correctamente.

2. Analizar los datos utilizando el protocolo desarrollado anteriormente. Duración 4 semanas.
 - a. Alinear los datos con un genoma de referencia (*mapping*).
 - b. Identificar alteraciones en las regiones codificantes del genoma (*variant calling*).
 - c. Anotar todas las variaciones encontradas (*variant annotation*).
 - d. Priorizar las variantes para encontrar aquellas que tengan mayor probabilidad de contribuir al desarrollo de los schwannomas (*priorización*).

3. Comparar los dos conjuntos de datos. Duración 3 semanas.
 - a. Calcular el número de mutaciones por megabase para cada grupo de schwannomas.
 - b. Comprobar mediante un análisis estadístico si existen diferencias en el número de mutaciones por megabase entre ambos grupos.
 - c. Construir un perfil genético de cada grupo de schwannomas donde se pueda ver los genes que aparecen mutados, su frecuencia y el tipo de mutación en cada muestra.
 - d. Obtener las rutas metabólicas en las que participan la mayoría de los genes encontrados en cada grupo.

- e. Comparar los perfiles genéticos y rutas metabólicas enriquecidas en cada grupo para ver si intervienen los mismos mecanismos moleculares en los dos grupos de schwannomas.

El tiempo total previsto para la ejecución del proyecto es de 11 semanas, desde 18/03/19 hasta 3/06/19. Para la correcta realización del proyecto en el plazo previsto deben alcanzarse una serie de hitos o metas en las fechas planeadas, las cuales se detallan a continuación:

- Puesta a punto del protocolo. 15/04/19.
- Identificación de las alteraciones (*mapping* y *variant calling*). 29/04/19.
- Identificación de los genes mutados (*variant annotation* y *priorización*). 13/05.
- Realización de la comparación entre los dos conjuntos de datos. 3/06/19.
- Finalización del proyecto. 4/06/19.
- Fin de elaboración de la memoria. 6/06/19.
- Fin de elaboración de la presentación. 13/06/19.

En la Figura 1. se puede ver la distribución por semanas de las tareas que se van a realizar durante el proyecto, los hitos a alcanzar, entregables y período de elaboración de la memoria y presentación.

1.5 Breve resumen de productos obtenidos

Al terminar la realización del proyecto, se han podido obtener los siguientes resultados:

- Genes mutados con alta probabilidad de ser dañinos en cada grupo de schwannomas con el porcentaje de muestras en que aparecen mutados.
- Perfil genético de cada grupo de schwannomas, con el tipo y número de mutaciones que aparecen en cada muestra.
- Comparación estadística entre el número de mutaciones por megabase de cada grupo de schwannomas.
- Rutas metabólicas enriquecidas en ambos grupos de schwannomas.

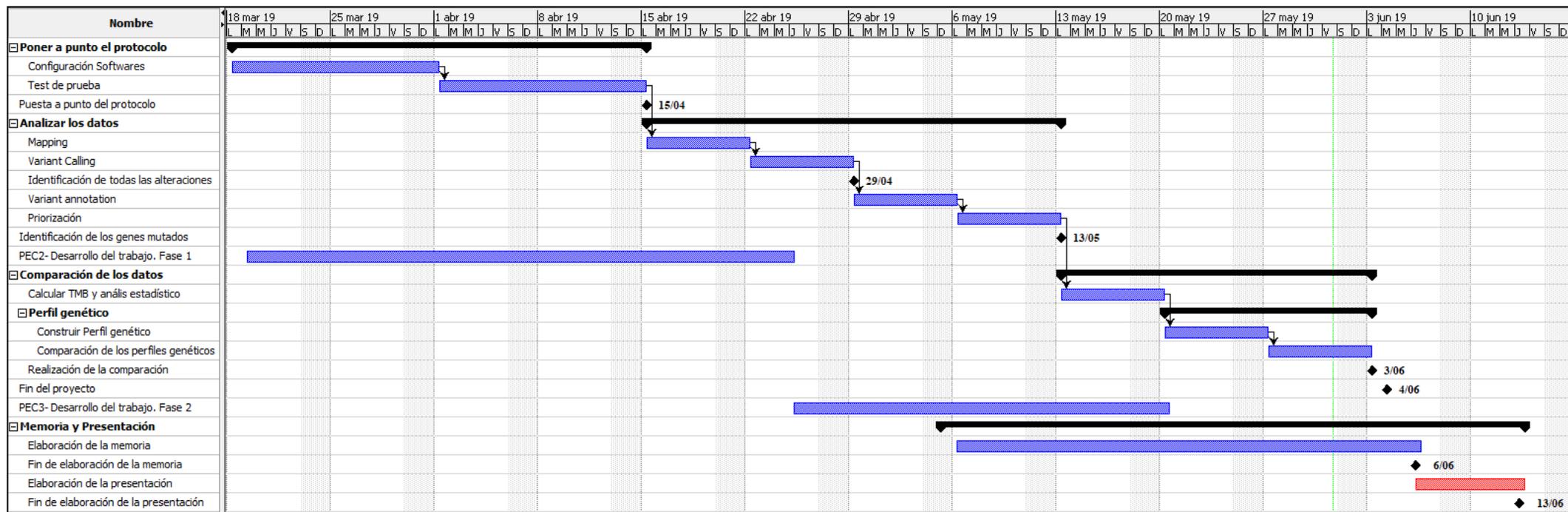


Figura 1. Diagrama de Gantt con la planificación temporal del proyecto.

1.6 Breve descripción de los otros capítulos de la memoria

- *Materiales y métodos*: descripción de los datos y herramientas bioinformáticas utilizadas para llevar a cabo el proyecto, así como el procedimiento que se ha seguido.
- *Resultados*: descripción detallada de los resultados obtenidos durante la realización del proyecto.
- *Discusión*: explicación de los resultados obtenidos durante el proyecto y comparación con estudios previos.

2. Materiales y métodos

2.1 Obtención de los datos

Para la realización del proyecto se utilizaron datos procedentes de la secuenciación del exoma completo (WES) de 22 schwannomas esporádicos (Illumina Truseq) y 9 schwannomas asociados a neurofibromatosis tipo 2 (Illumina Hiseq 2000). En ambos casos, la captura y amplificación estaba realizada con el método *SureSelect Human All Exon V5*. Los datos de schwannomas esporádicos proceden del estudio Agnihotri S. et al., 2016¹³ y fueron descargados desde el repositorio EGA⁴⁴ (*The European Genome-phenome Archive*).

2.2 Análisis

Utilizando los datos anteriores, se procedió a la realización del análisis de los mismos siguiendo las etapas descritas en el apartado 1.3 Enfoque y método seguido. Se utilizó un protocolo previamente diseñado por el grupo de investigación que incluye todos los pasos hasta *variant annotation*, que fue modificado para poder completar todas las etapas del análisis. En el apartado 8.Anexos, se puede ver todo el código utilizado, diferenciando entre las partes que ya estaban diseñadas previamente y las que fueron específicamente elaboradas para este trabajo.

Mapping

Las secuencias, en formato *fastq*, obtenidas de cada uno de los schwannomas se alinearon con el genoma de referencia (hs37d5) utilizando el algoritmo *bwa mem* del software BWA v. 0.7.13. El archivo obtenido después del alineamiento tiene formato *sam*, pero para utilizar el software Strelka, es necesario utilizar un archivo en formato *bam*. SAMtools v. 0.1.19 es una herramienta que, entre otras cosas, permite hacer la conversión entre ambos formatos.

Para mejorar y aumentar la velocidad del análisis posterior, las secuencias alineadas debieron ser ordenadas según su posición en el genoma de referencia. Además, se creó un índice de las mismas, con el objetivo de ser localizadas fácilmente por los diferentes softwares que se utilizaron en el análisis. Estos dos pasos, también fueron realizados con SAMtools.

Variant Calling

Para la identificación de las variantes se utilizó el algoritmo *configureStrelkaGermlineWorkflow.py* del software Strelka v. 2.9.3. El archivo obtenido en esta etapa está en formato *vcf*. Una vez que las variantes han sido identificadas es necesario realizar una normalización de las mismas para evitar ambigüedades, debido a que una misma variante puede ser representada de diferentes maneras⁴⁵. Este proceso consiste en:

1. Representar cada variante con el menor número posible de nucleótidos, eliminando los nucleótidos de los extremos que sean similares con el genoma de referencia.
2. Alinear hacia la izquierda cada variante añadiendo un nucleótido por este lado y eliminando otro por la derecha, hasta que la última base de cada alelo sea diferente.
3. En el caso de que alelos múltiples estén representados en la misma entrada, se separan en entradas diferentes.

Una vez finalizada la normalización, también se debe crear un índice de las variantes contenidas en el archivo vcf.

Variant Annotation

La identificación de las variantes se realizó con el software Annovar v. 2.1.1. Las variantes fueron anotadas utilizando las siguientes bases de datos:

- refGene: para identificar el gen, la región y el tipo de mutación (con sentido, sin sentido, desplazamiento del marco de lectura...).
- dbnsfp30a: para realizar una predicción funcional de las variantes. Incluye anotaciones de varios predictores *in silico* que predicen el efecto que tendrá una mutación en la secuencia final de la proteína.
- exac03: para obtener la frecuencia alélica de cada variante en diferentes poblaciones humanas.

Priorización

Debido a que, durante el proceso de *variant calling*, el número de variantes que se obtienen es extremadamente alto y que, la mayoría son neutras o falsos positivos, fue necesario realizar un proceso de priorización para trabajar únicamente con aquellas que tengan una alta probabilidad de ser dañinas y, por lo tanto, causantes del fenotipo. El proceso se realizó en varios pasos³⁴:

1. Calidad⁴⁶: sólo se consideró aquellas variantes que cumplían con los siguientes criterios de calidad:
 - Número de lecturas ≥ 10 .
 - Nivel de calidad Phred (probabilidad de que exista una alteración en la secuencia) ≥ 20 .
 - Calidad del genotipo (probabilidad de que el genotipo asignado sea el correcto) ≥ 20 .
2. Frecuencia alélica: si las mutaciones tienen una frecuencia alélica baja es más probable que estén relacionados con una enfermedad rara que aquellas que tienen

una frecuencia alélica alta. Por lo que se eliminaron todas las variantes con una frecuencia alélica mayor de 0.01.

3. Impacto biológico⁴⁷: el siguiente paso fue estudiar el impacto biológico de las variantes, es decir, el efecto que tienen sobre la secuencia proteica final. Las mutaciones que son sinónimas no alteran la secuencia de aminoácidos de la proteína, es decir, no tienen un efecto deletéreo, por lo que fueron eliminadas. Las mutaciones con cambio de sentido pueden tener o no un efecto en la secuencia proteica que puede ser deletéreo. Existen un gran número de predictores *in silico* que predicen el impacto biológico de una mutación. Las mutaciones sin sentido y con desplazamiento del marco de lectura tienen un gran efecto sobre la secuencia de la proteína, por lo que éstas tienen una alta probabilidad de estar implicadas en la patología. En este trabajo, se consideraron como dañinas todas las mutaciones sin sentido, con desplazamiento del marco de lectura y aquellas mutaciones con cambio de sentido que fueron detectadas como deletéreas por los predictores *Likelihood Ratio Test* (LRT) y *MutationTaster*.

MutationTaster utiliza información de diferentes bases de datos (conservación genética, cambios en los sitios de splicing, pérdida de función de la proteína, cambios que puedan afectar a la cantidad de ARNm) para predecir si una variante tiene una alta probabilidad de producir una patología utilizando un clasificador probabilístico Bayesiano⁴⁸.

LRT estima la probabilidad de que el codón que presenta la alteración haya evolucionado negativamente, con un mayor aumento del número de mutaciones no sinónimas⁴⁹.

4. Recurrencia: los genes que aparecen mutados en varias muestras independientes tienen mayor probabilidad de influir en el desarrollo de la patología, por lo que, para el resultado final, sólo se consideraron los genes que aparecen mutados en dos o más muestras dentro de cada grupo, independientemente del tipo de mutación que presenten.

Comparación entre los dos grupos de schwannomas

Una vez realizado el análisis anterior con cada muestra, se obtiene un conjunto de genes mutados que podrían influir en el desarrollo de la patología. Con este conjunto de genes se realizaron los siguientes estudios:

- Perfil genético: para cada grupo de schwannomas se construyó un perfil genético en el que se puede ver el número de mutaciones y los tipos de mutaciones que hay en cada muestra. Además, se elaboró una tabla en la que aparecen todos los genes mutados y su frecuencia dentro de cada grupo de schwannomas.
- Cálculo del número de mutaciones por megabase: para cada muestra, se calculó el número de mutaciones por megabase dividiendo el número de genes mutados entre 30 (tamaño aproximado del exoma en megabases).
- Comparación estadística: la comparación estadística entre el número medio de mutaciones por megabase por cada grupo de schwannomas se realizó con la prueba estadística no paramétrica U de Mann-Whitney.
- Rutas metabólicas enriquecidas: se utilizó la herramienta *Functional Annotation* de la aplicación web DAVID⁵⁰ v. 6.8. (*Database for Annotation, Visualization and Integrated Discovery*) para obtener el conjunto de rutas metabólicas que están más enriquecidas en cada tipo de schwannoma según las anotaciones de *Gene Ontology*^{51,52} (GO). Se consideraron como enriquecidas sólo aquellas rutas que tenían un p-valor (Prueba exacta de Fisher) igual o menor que 0.01.

3. Resultados

3.1. Número de mutaciones

El número medio de mutaciones por megabase (Figura 2 A.) fue significativamente mayor (U de Mann-Whitney, *p-valor* = 0.001) en los schwannomas esporádicos (3.42 mutaciones/Mb) comparado con los schwannomas asociados a neurofibromatosis tipo 2

(1.91 mutaciones/Mb). Además, de todos los genes que se han encontrado mutados en ambos grupos de schwannomas (Figura 2 B.), 153 se han visto únicamente en schwannomas asociados a neurofibromatosis tipo 2, mientras que 845 se han encontrado asociados con schwannomas esporádicos. Solamente, se han visto 59 genes (5.6% del total) presentes en ambos tipos de tumores.

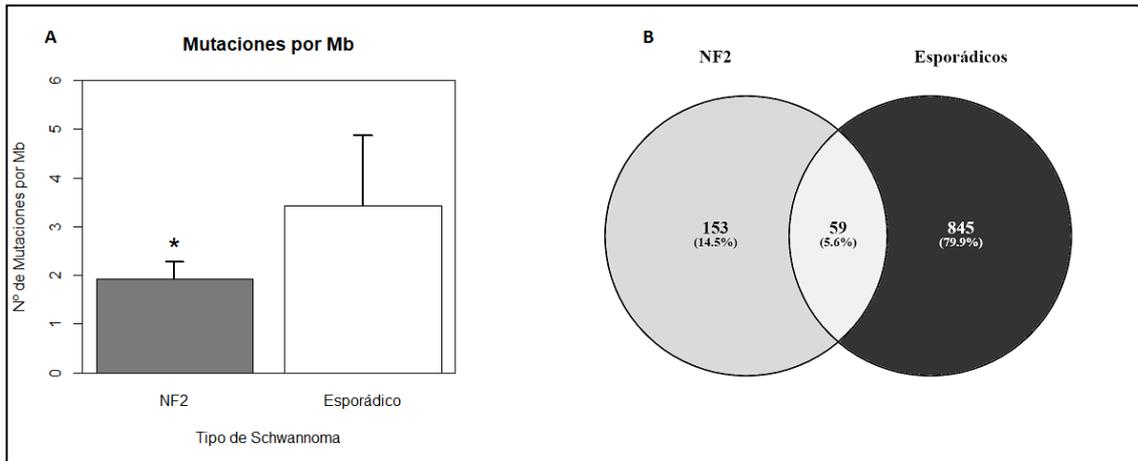


Figura 2. Número de mutaciones por cada grupo de Schwannoma.

3.2 Perfil genético

En la Figura 3., se puede ver el perfil genético de cada grupo de schwannoma. En ambos grupos, la mayoría de las mutaciones detectadas son de cambio de sentido y, unas pocas, sin sentido. No se han encontrado mutaciones de tipo desfase de la pauta de lectura. Como ya se ha comentado anteriormente, el número de mutaciones es, en general, bastante mayor en los schwannomas esporádicos, encontrándose una gran variabilidad entre las muestras (menos de 50 y más de 150 mutaciones). En el caso de los schwannomas asociados a neurofibromatosis tipo 2, el número de mutaciones por muestra es más homogéneo (± 50).

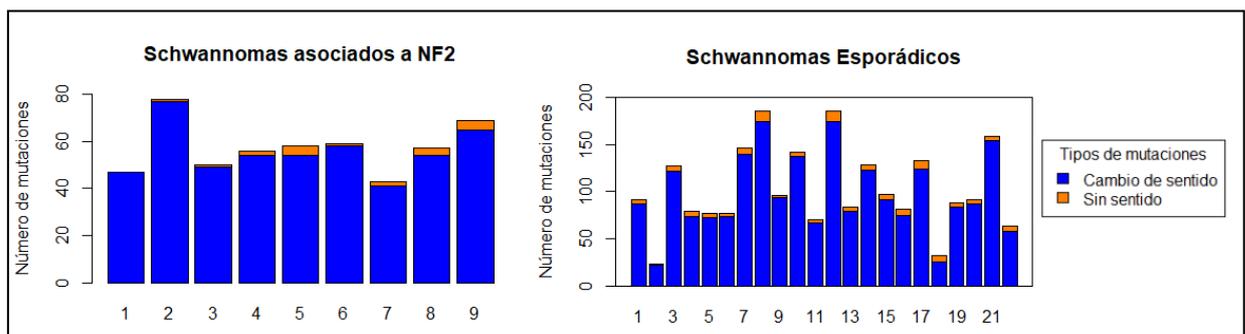


Figura 3. Perfil genético de los dos grupos de Schwannomas.

En la Tabla 1., se pueden ver los 10 genes más frecuentemente mutados en cada tipo de schwannoma. Como era de esperar, en base a estudios previos, el gen NF2 aparece mutado en ambos grupos (en un 44% de las muestras de schwannomas asociados a neurofibromatosis tipo 2 y en un 26% de las muestras de schwannomas esporádicos). Se puede ver que el gen PABPC3, también, aparece mutado con una alta frecuencia en ambos tipos de tumores (más del 50% de las muestras). En la Tabla 3. de los anexos se pueden ver todos los genes mutados en cada tipo de schwannoma con sus respectivas frecuencias.

Schwannomas asociados a NF2			Schwannomas Esporádicos		
Genes	Casos	Frecuencia (%)	Genes	Casos	Frecuencia (%)
HLA-DRB5	6	67	HLA-DRB1	12	52
PABPC3	5	56	PABPC3	12	52
CFTR	4	44	TTN	10	43
FLNC	4	44	BBS9	6	26
LRP2	4	44	COL6A3	6	26
NF2	4	44	DAAM2	6	26
SCUBE3	4	44	DCHS1	6	26
ZFHX3	4	44	HTT	6	26
ABCA13	3	33	LAMA5	6	26
AHNAK2	3	33	NF2	6	26

Tabla 1. Genes mutados más frecuentemente en cada grupo de schwannomas.

3.2 Rutas metabólicas enriquecidas

En la Figura 4 A. se pueden ver las rutas metabólicas más enriquecidas en los schwannomas asociados a neurofibromatosis tipo 2 y los procesos biológicos de los que forman parte estas rutas metabólicas (Figura 4 B.). La mayoría, están asociadas al desarrollo anatómico de órganos (con un 47% de los genes mutados), a la adhesión celular (14%) y a la proliferación celular (10%). En la Figura 5 A. se pueden ver las 10 rutas metabólicas más enriquecidas en los schwannomas esporádicos y los procesos biológicos de los que forman parte estas rutas metabólicas (Figura 5 B.). (en la Figura 6 de los anexos se puede ver una gráfica con todas las rutas metabólicas). La mayoría de

estas rutas están asociadas al desarrollo anatómico (41%), la organización de los componentes celulares (22%) y a la localización (8%).

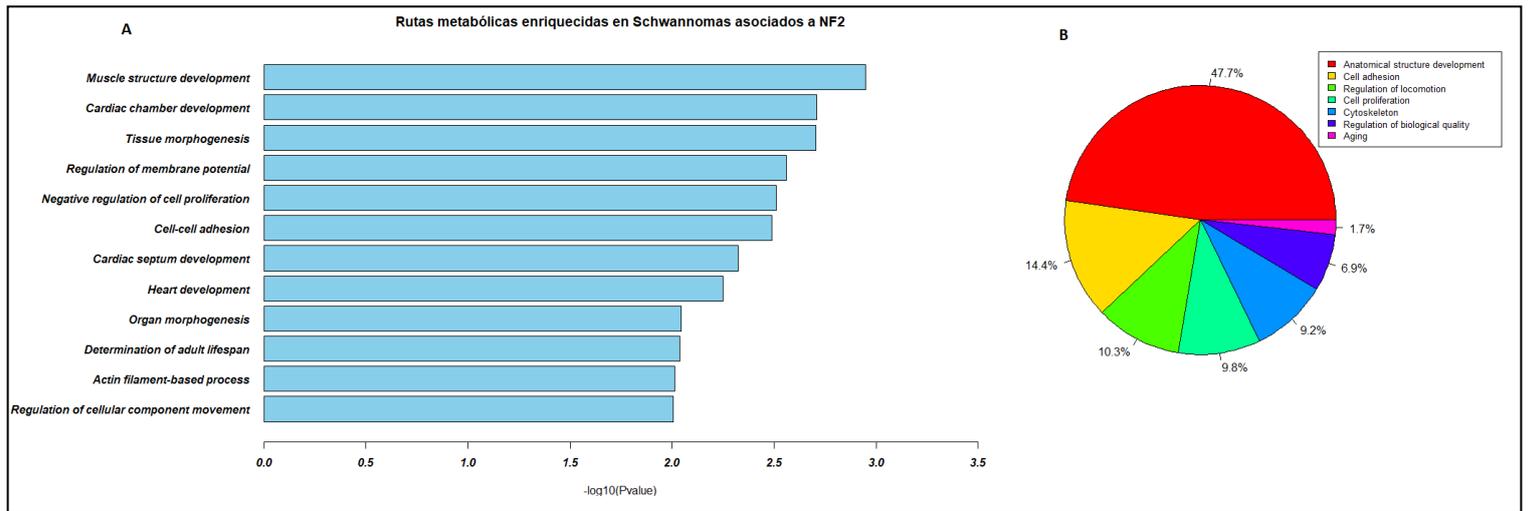


Figura 4. Rutas metabólicas enriquecidas en Schwannomas asociados a neurofibromatosis tipo 2.

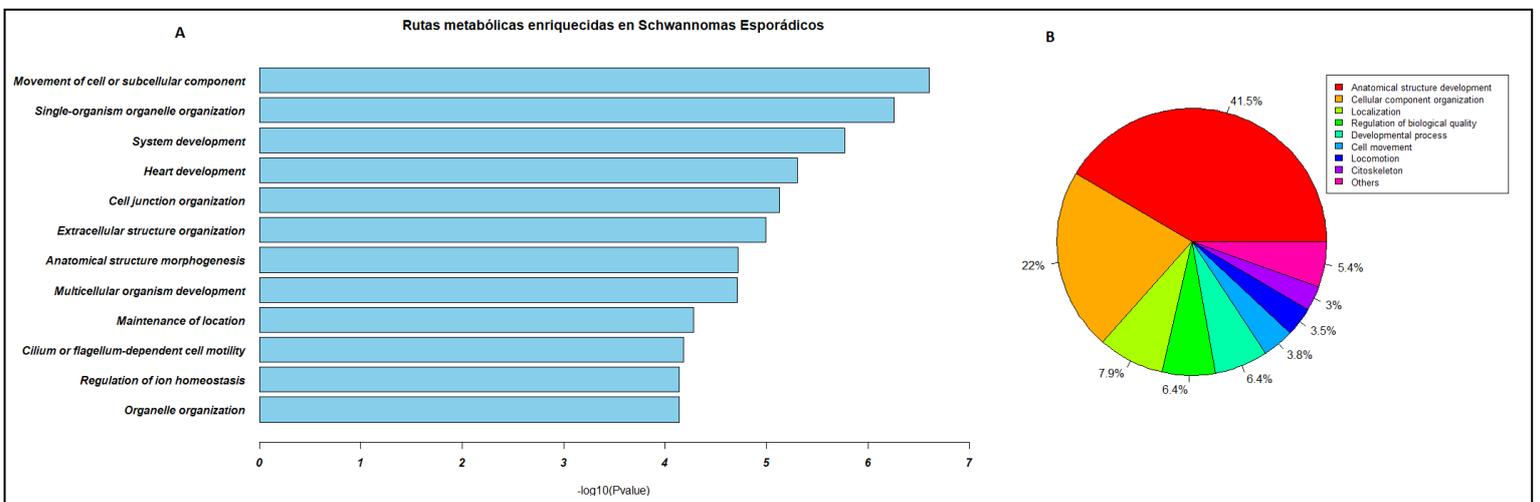


Figura 5. Diez rutas metabólicas más enriquecidas de los Schwannomas esporádicos.

4. Discusión

Como se puede ver en la Figura 2., en los schwannomas esporádicos aparece un número mayor de mutaciones y sólo se ha encontrado menos de un 6% de genes compartidos entre ambos tipos de schwannomas. Estos resultados sugieren que los schwannomas esporádicos y los schwannomas asociados a neurofibromatosis tipo 2 podrían desarrollarse debido a mecanismos moleculares diferentes. Además, el número de mutaciones por megabase en cada tipo de tumor es bajo comparado con otros tipos de tumores²⁵. Estos resultados van en la línea de otros estudios previos, los cuales no muestran evidencias de un alto número de mutaciones en los schwannomas⁵.

En cuanto a las rutas metabólicas que se encuentran más enriquecidas (Figura 4 y Figura 5), la mayoría de los genes, en ambos tipos de schwannomas, tienen funciones en el desarrollo anatómico (principalmente tejido muscular y cardíaco), con más de un 40% de los genes alterados dentro de cada grupo. En los schwannomas asociados a neurofibromatosis tipo 2, una parte importante de los genes (23%) están relacionados con la adhesión celular y el citoesqueleto y un 10% regulan la proliferación celular. En los schwannomas esporádicos, se ha encontrado que un 22% de los genes mutados participan en el control de la organización de los componentes celulares y un 6.4% están implicados, entre otras cosas, en la señalización celular (regulación del transporte de calcio y potencial de membrana, entre otros) y homeostasis (*Regulation of biological quality*). Estos hallazgos apoyan la idea de que los mecanismos moleculares que promueven el desarrollo de ambos tipos de tumores podrían ser diferentes.

En ambos tipos de tumores, se puede ver (Figura 3.) que la mayoría de las mutaciones implicadas son de cambio de sentido (cambio de un nucleótido que provoca un cambio en la secuencia de aminoácidos) y unas pocas son sin sentido (cambio de un nucleótido que provoca la aparición de un codón de *stop*). Estos resultados, para los schwannomas esporádicos, son similares a los que aparecen en el estudio Agnihotri S. et al., 2016¹³ (del cual se obtuvieron los datos), salvo que el número de mutaciones encontrado en este trabajo es mayor y no se han visto inserciones ni deleciones. Esto se podría explicar por las diferencias en el análisis de los datos entre un estudio y otro.

Como era de esperar, el gen NF2 está mutado en un alto porcentaje de las muestras (en un 44% de las muestras de schwannomas asociados a neurofibromatosis tipo 2 y en un 26% de las muestras de schwannomas esporádicos), lo cual coincide con estudios previos que colocan a NF2 como un gen clave en el desarrollo de los schwannomas^{13,31}. También se ha encontrado que el gen PABPC3 aparece mutado con una alta frecuencia (en un 56% de las muestras de schwannomas asociados a neurofibromatosis tipo 2 y en un 52% de las muestras de schwannomas esporádicos). Se trata de una proteína de unión a la cola poli A del ARNm que favorece la unión de los ribosomas, el inicio de la traducción y controla la degradación del ARNm. Previamente se ha relacionado con cáncer de tiroides⁵³, por lo que, junto con NF2 podría tener un papel central en el desarrollo de los schwannomas.

Otros genes, que se encuentran mutados con menos frecuencia exclusivamente en los schwannomas esporádicos han sido asociados previamente con estos tumores. De todos los genes encontrados únicamente en schwannomas asociados a neurofibromatosis tipo 2 ninguno ha sido asociado previamente con schwannomas (exceptuando NF2), hay que tener en cuenta que no existen estudios previos de este tipo. Aunque, algunos genes mutados en ambos tipos de tumores sí han sido relacionados con el desarrollo de schwannomas. En la Tabla 2.. se puede ver un resumen de todos estos genes con su frecuencia y el tipo de schwannoma en el que se han encontrado mutados en este trabajo. Algunos de estos genes (PKD1, INADL, TSC1) están relacionados con funciones de adhesión celular y otros (PLEC, ATM) controlan algunos aspectos del ciclo celular, por lo que podrían tener un papel importante en la regulación del crecimiento de los schwannomas.

Gen	Frecuencia (%)	Schwannomas	Publicación
TTN	43 / 33	Esporádicos / Asociados a Neurofibromatosis tipo 2	Håvik AL, 2018 ³¹
PLEC	26 / 22	Esporádicos / Asociados a Neurofibromatosis tipo 2	Håvik AL, 2018 ³¹
PKD1	13	Esporádicos	Håvik AL, 2018 ³¹
ATM	9 / 22	Esporádicos / Asociados a Neurofibromatosis tipo 2	Håvik AL, 2018 ³¹
EIF5B	9	Esporádicos	Håvik AL, 2018 ³¹
KALRN	9	Esporádicos	Håvik AL, 2018 ³¹
INADL	9	Esporádicos	Håvik AL, 2018 ³¹
TSC1	9	Esporádicos	Agnihotri S, 2016 ¹³

Tabla 2. Genes mutados previamente relacionados con Schwannomas.

En schwannomas esporádicos se han encontrado otros genes mutados, que no se han relacionado previamente con estos tumores, pero sí con otras patologías cancerosas. Los genes LAMA5 y COL6A3 han sido asociados con cáncer colorrectal^{54,55}. LAMA5 codifica para la cadena alfa de las lamininas, proteínas que se encuentran en la matriz extracelular y está asociadas con adhesión celular, migración, proliferación celular y metástasis y aparece en el 26% de las muestras. COL6A3 codifica para la cadena alfa-3 del colágeno tipo VI, que interviene en la organización de la matriz extracelular, también se encuentra mutado en el 26% de las muestras de schwannomas esporádicos, aunque también aparece, con algo menos frecuencia (22%) en los schwannomas asociados a neurofibromatosis tipo 2. El gen HLA-DRB1 aparece mutado en el 52% de las muestras de schwannomas esporádicos, codifica para una proteína con funciones en el sistema inmune, presentando péptidos derivados de la matriz extracelular, ha sido relacionado en estudios previos con cáncer cervical⁵⁶, por lo que podría tener un papel fundamental en el crecimiento de este tipo de tumores.

En los schwannomas asociados a neurofibromatosis tipo 2, también, se han encontrado algunos genes mutados, que en estudios previos se han relacionado con otros tipos de tumores y podrían influenciar en su crecimiento. El gen HLA-DRB5 se encuentra mutado en un 67% de las muestras y ha sido asociado previamente con leucemia linfocítica crónica⁵⁷. FLNC codifica para la filamina c, proteína que se encarga de

anclar las proteínas de membrana al citoesqueleto, aparece mutado en el 44% de las muestras y ha sido relacionado con linfoma esplénico marginal⁵⁸. ZFHX3 es un gen que, previamente, se ha encontrado mutado en cáncer de vejiga⁵⁹, se trata de un factor de transcripción que controla la diferenciación muscular y neuronal, que se encuentra mutado en el 44% de las muestras. También aparece mutado en los schwannomas esporádicos, aunque con menos frecuencia (13%).

5. Conclusiones

Como resultado de este trabajo, se han podido obtener una serie de genes que aparecen mutados en muestras de schwannomas esporádicos y schwannomas asociados a neurofibromatosis tipo 2. Aunque el número de genes mutados es bastante alto, a falta de más estudios que confirmen los resultados obtenidos y ayuden a disminuir esta lista, se pueden extraer las siguientes conclusiones:

- Los schwannomas esporádicos parecen necesitar un mayor número de mutaciones para su desarrollo que los schwannomas asociados a neurofibromatosis tipo 2.
- La mayoría de mutaciones que están implicadas en el crecimiento de ambos tipos de schwannomas son de cambio de sentido.
- El análisis de las rutas metabólicas enriquecidas muestra diferentes procesos biológicos en ambos grupos de schwannomas, por lo que los schwannomas esporádicos podrían seguir un mecanismo molecular de desarrollo diferente al de los asociados a neurofibromatosis tipo 2.
- Teniendo en cuenta los resultados encontrados en este trabajo y en estudios previos, NF2 puede tener un papel central en el desarrollo de ambos tipos de schwannomas.
- Estudios previos también muestran mutaciones en los genes PKD1, PLEC, TTN y ATM, por lo que pueden ser importantes para el crecimiento de los schwannomas.
- En este trabajo se proponen otros genes, como PABPC3, que pueden ser necesarios para la aparición de schwannomas esporádicos y schwannomas

asociados a neurofibromatosis tipo 2, aunque es necesario la realización de más estudios para confirmarlo.

- También se han encontrado otros genes mutados únicamente en schwannomas asociados neurofibromatosis tipo 2 como FLNC, HLA-DRB5 y ZFH3, que podrían estar relacionados con este tipo de tumores. Aunque no existen estudios previos y es necesario realizar una investigación más profunda para confirmarlo.
- Otros genes como LAMA5, HLA-DRB1 y COL6A3 podrían estar relacionados con schwannomas esporádicos, pero es necesaria la realización de más estudios para poder comprobarlo.

Para completar el estudio de los mecanismos moleculares que imperan en el desarrollo de schwannomas (concretamente esporádicos y asociados a neurofibromatosis tipo 2) se deben explorar, en un futuro, las siguientes líneas de trabajo:

- Estudiar las mutaciones en sitios de *splicing* que puedan estar involucradas en estas patologías.
- Investigar las mutaciones que sean únicamente somáticas (que aparecen *de novo* en el tumor o en células precursoras) comparando las mutaciones encontradas en el tumor con las mutaciones encontradas en una muestra de sangre del mismo paciente.
- Identificar genes que aparezcan mutados más frecuentemente de los esperado según el ratio de mutación del paciente utilizando herramientas bioinformáticas como MutSigCV⁶⁰, con el fin de detectar mutaciones conductoras (las que realmente influyen en desarrollo de los tumores).

Inicialmente, los objetivos y subobjetivos planteados en este estudio fueron los siguientes:

- A. Detectar posibles genes mutados en un conjunto de datos de 9 schwannomas de pacientes con neurofibromatosis tipo 2 y 22 schwannomas esporádicos obtenidos por WES.
 1. Poner a punto un protocolo para analizar los datos obtenidos por WES.
 2. Identificar las variantes de genes mutados utilizando el protocolo desarrollado en el objetivo anterior.

- B. Comparar los resultados de ambos tipos de schwannomas.
3. Comprobar estadísticamente si existen diferencias entre el número de mutaciones por megabase de los schwannomas esporádicos y los schwannomas asociados a neurofibromatosis tipo 2.
 4. Comparar los genes mutados y el tipo de mutación que aparecen más frecuentemente en cada tipo de schwannoma.
 5. Comprobar si existen diferencias en las rutas metabólicas en las que participan los genes mutados en cada grupo de schwannomas.

Finalmente, se ha obtenido un protocolo que puede ser utilizado para el análisis de datos obtenidos por WES de schwannomas (ver código en apartado Anexos). Por otro lado, se han conseguido analizar los datos para obtener genes mutados que pueden influir en el desarrollo de los schwannomas (Tabla 3. Genes mutados en cada grupo de Schwannomas.). También, se ha conseguido realizar la comparación del perfil genético, número de mutaciones por megabase y rutas metabólicas enriquecidas entre los schwannomas esporádicos y schwannomas asociados a neurofibromatosis tipo 2. Por lo que todos los objetivos propuestos han sido completados con éxito dentro del plazo previsto.

El desarrollo del trabajo se dividió en 3 fases, dedicando un tiempo total previsto para la ejecución del proyecto de 11 semanas:

1. Desarrollar el protocolo a utilizar. Duración 4 semanas.
2. Analizar los datos utilizando el protocolo desarrollado anteriormente. Duración 4 semanas.
3. Comparar los dos conjuntos de datos. Duración 3 semanas.

En la fecha prevista para la finalización del proyecto 3/06/19, se han conseguido cumplir con todos los objetivos planteados (algunos de ellos incluso antes del tiempo marcado), por lo que la planificación ha sido la adecuada, no siendo necesarios ningún tipo de cambio o modificación.

6. Glosario

- **Neurofibromatosis 2:** trastorno genético provocado por una mutación en el gen NF2 que produce una deficiencia en la proteína merlina. Como consecuencia existe una predisposición mayor al desarrollo de schwannomas y meningiomas (tumor que surge en las meninges).
- **Schwannomatosis:** predisposición genética a padecer múltiples schwannomas, normalmente localizados en las extremidades.
- **Complejo de Carney:** síndrome genético que predispone a la aparición de mixomas (tumores del tejido conjuntivo), lentiginosis (puntos de hiperpigmentación de la piel) y una sobreactividad endocrina.
- **Tinnitus:** percepción de zumbidos en los oídos.
- **Schwannomas esporádicos:** schwannomas que surgen de forma aislada sin estar relacionados con ningún síndrome genético predisponente.
- **Schwannomas asociados a neurofibromatosis tipo 2:** schwannomas cuyo desarrollo está relacionado a la deficiencia en la proteína merlina.
- **NGS:** técnicas de secuenciación de nueva generación. Conjunto de técnicas de secuenciación masiva que surgen después de la secuenciación de Sanger.
- **Fluoróforo:** componente de una molécula que emite flouorescencia.
- **Secuenciación de Sanger:** técnica de secuenciación diseñada por Frederick Sanger en 1975. Se basa en el uso de dideoxinucleótidos (nucleótidos que carecen del grupo hidroxilo del carbono 3') los cuales impiden que una cadena pueda seguir elongándose durante la replicación mediante la reacción en cadena de la polimerasa (PCR). Se utilizan 4 tubos de reacción con dideoxinucleótidos diferentes, una vez a terminado la PCR se realiza una electroforesis en gel de agarosa, donde se puede leer la secuencia de la cadena a partir de todos los fragmentos generados.
- **Paneles de genes:** se trata de tests genéticos que permiten la secuenciación de unos genes concretos (normalmente relacionados con una cierta patología o conjunto de patologías).
- **Exoma:** parte del genoma formada únicamente por los exones de los genes, es decir, las partes que se traducen a proteína.

- **Sondas:** pequeño fragmento de ADN complementario a una secuencia de interés que se quiere detectar.
- **Fenotipo:** características visibles de un individuo como resultado de la interacción de su genoma y el medio ambiente.
- **Mutación de cambio de sentido:** mutación puntual en la que se produce un cambio de nucleótido que conlleva la aparición de un nuevo codón con un cambio de aminoácido.
- **Mutación sin sentido:** mutación puntual que tienen como consecuencia la creación de un codón de *stop*, provocando que la secuencia de aminoácidos final sea más corta.
- **dNTP:** dinucleótidos trifosfato.
- **ARNm:** ARN mensajero.

7. Bibliografía

1. Louis DN, Ohgaki H, Wiestler OD, et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol (Berl)*. 2007;114(2):97-109. doi:10.1007/s00401-007-0243-4
2. Strowd RE, Blakeley JO. Common Histologically Benign Tumors of the Brain. *Contin Minneap Minn*. 2017;23(6, Neuro-oncology):1680-1708. doi:10.1212/CON.0000000000000541
3. Mankin HJ, Mankin KP. Schwannoma: a rare benign tumor of soft tissues. *Musculoskelet Surg*. 2014;98(2):95-99. doi:10.1007/s12306-014-0333-0
4. Hilton DA, Hanemann CO. Schwannomas and their pathogenesis. *Brain Pathol Zurich Switz*. 2014;24(3):205-220. doi:10.1111/bpa.12125
5. Benign Intracranial Tumors. - PubMed - NCBI. <https://www.ncbi.nlm.nih.gov/pubmed/?term=Benign+Intracranial+Tumors+Justin+T.+Jordan>. Accessed May 28, 2019.
6. Evans DGR. Neurofibromatosis type 2 (NF2): a clinical and molecular review. *Orphanet J Rare Dis*. 2009;4:16. doi:10.1186/1750-1172-4-16
7. Hanemann CO, Evans DG. News on the genetics, epidemiology, medical care and translational research of Schwannomas. *J Neurol*. 2006;253(12):1533-1541. doi:10.1007/s00415-006-0347-0
8. Warren C, James LA, Ramsden RT, et al. Identification of recurrent regions of chromosome loss and gain in vestibular schwannomas using comparative genomic hybridisation. *J Med Genet*. 2003;40(11):802-806. doi:10.1136/jmg.40.11.802
9. Paganini I, Capone GL, Vitte J, et al. Double somatic SMARCB1 and NF2 mutations in sporadic spinal schwannoma. *J Neurooncol*. 2018;137(1):33-38. doi:10.1007/s11060-017-2711-6
10. Sass H, Cayé-Thomasen P. Contemporary Molecular Biology of Sporadic Vestibular Schwannomas: A Systematic Review and Clinical Implications. *J Int Adv Otol*. 2018;14(2):322-329. doi:10.5152/iao.2018.4929
11. Sass HCR, Borup R, Alanin M, Nielsen FC, Cayé-Thomasen P. Gene expression, signal transduction pathways and functional networks associated with growth of sporadic vestibular schwannomas. *J Neurooncol*. 2017;131(2):283-292. doi:10.1007/s11060-016-2292-9
12. Woods R, Friedman JM, Evans DGR, Baser ME, Joe H. Exploring the "two-hit hypothesis" in NF2: tests of two-hit and three-hit models of vestibular schwannoma development. *Genet Epidemiol*. 2003;24(4):265-272. doi:10.1002/gepi.10238

13. Agnihotri S, Jalali S, Wilson MR, et al. The genomic landscape of schwannoma. *Nat Genet.* 2016;48(11):1339-1348. doi:10.1038/ng.3688
14. Schulz A, Zoch A, Morrison H. A neuronal function of the tumor suppressor protein merlin. *Acta Neuropathol Commun.* 2014;2:82. doi:10.1186/s40478-014-0082-1
15. NF2/merlin in hereditary neurofibromatosis 2 versus cancer: biologic mechanisms and clinical associations. - PubMed - NCBI. <https://www.ncbi.nlm.nih.gov/pubmed/?term=Nf2%2Fmerlin+in+hereditary+neurofibromatosis+2+versus+cancer%3A+biologic+mechanisms+and+clinical+associations>. Accessed May 28, 2019.
16. Blakeley JO, Plotkin SR. Therapeutic advances for the tumors associated with neurofibromatosis type 1, type 2, and schwannomatosis. *Neuro-Oncol.* 2016;18(5):624-638. doi:10.1093/neuonc/nov200
17. Petersen B-S, Fredrich B, Hoepfner MP, Ellinghaus D, Franke A. Opportunities and challenges of whole-genome and -exome sequencing. *BMC Genet.* 2017;18(1):14. doi:10.1186/s12863-017-0479-5
18. Adams DR, Eng CM. Next-Generation Sequencing to Diagnose Suspected Genetic Disorders. *N Engl J Med.* 2018;379(14):1353-1362. doi:10.1056/NEJMra1711801
19. Jamuar SS, Tan E-C. Clinical application of next-generation sequencing for Mendelian diseases. *Hum Genomics.* 2015;9:10. doi:10.1186/s40246-015-0031-5
20. Shabani Azim F, Houri H, Ghalavand Z, Nikmanesh B. Next Generation Sequencing in Clinical Oncology: Applications, Challenges and Promises: A Review Article. *Iran J Public Health.* 2018;47(10):1453-1457.
21. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet.* 2016;17(6):333-351. doi:10.1038/nrg.2016.49
22. Teer JK, Mullikin JC. Exome sequencing: the sweet spot before whole genomes. *Hum Mol Genet.* 2010;19(R2):R145-151. doi:10.1093/hmg/ddq333
23. Rabbani B, Tekin M, Mahdih N. The promise of whole-exome sequencing in medical genetics. *J Hum Genet.* 2014;59(1):5-15. doi:10.1038/jhg.2013.114
24. Majewski J, Schwartzenuber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? *J Med Genet.* 2011;48(9):580-589. doi:10.1136/jmedgenet-2011-100223
25. Chalmers ZR, Connelly CF, Fabrizio D, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* 2017;9(1):34. doi:10.1186/s13073-017-0424-2

26. Zhou Z, Zhu L, Niu X, et al. Comparison of genomic landscapes of large cell neuroendocrine carcinoma, small cell lung carcinoma, and large cell carcinoma. *Thorac Cancer*. 2019;10(4):839-847. doi:10.1111/1759-7714.13011
27. Fenizia F, Pasquale R, Roma C, Bergantino F, Iannaccone A, Normanno N. Measuring tumor mutation burden in non-small cell lung cancer: tissue versus liquid biopsy. *Transl Lung Cancer Res*. 2018;7(6):668-677. doi:10.21037/tlcr.2018.09.23
28. Tanskanen T, Gylfe AE, Katainen R, et al. Exome sequencing in diagnostic evaluation of colorectal cancer predisposition in young patients. *Scand J Gastroenterol*. 2013;48(6):672-678. doi:10.3109/00365521.2013.783102
29. Pareja F, Lee JY, Brown DN, et al. The Genomic Landscape of Mucinous Breast Cancer. *J Natl Cancer Inst*. January 2019. doi:10.1093/jnci/djy216
30. Parsons DW, Roy A, Yang Y, et al. Diagnostic Yield of Clinical Tumor and Germline Whole-Exome Sequencing for Children With Solid Tumors. *JAMA Oncol*. 2016;2(5):616-624. doi:10.1001/jamaoncol.2015.5699
31. Håvik AL, Bruland O, Myrseth E, et al. Genetic landscape of sporadic vestibular schwannoma. *J Neurosurg*. 2018;128(3):911-922. doi:10.3171/2016.10.JNS161384
32. Tetreault M, Bareke E, Nadaf J, Alirezaie N, Majewski J. Whole-exome sequencing as a diagnostic tool: current challenges and future opportunities. *Expert Rev Mol Diagn*. 2015;15(6):749-760. doi:10.1586/14737159.2015.1039516
33. Hintzsche JD, Robinson WA, Tan AC. A Survey of Computational Tools to Analyze and Interpret Whole Exome Sequencing Data. *Int J Genomics*. 2016;2016:7983236. doi:10.1155/2016/7983236
34. Bao R, Huang L, Andrade J, et al. Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. *Cancer Inform*. 2014;13(Suppl 2):67-82. doi:10.4137/CIN.S13779
35. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinforma Oxf Engl*. 2009;25(14):1754-1760. doi:10.1093/bioinformatics/btp324
36. Kim S, Scheffler K, Halpern AL, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods*. 2018;15(8):591-594. doi:10.1038/s41592-018-0051-x
37. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164. doi:10.1093/nar/gkq603

38. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinforma Oxf Engl.* 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352
39. Thankaswamy-Kosalai S, Sen P, Nookaew I. Evaluation and assessment of read-mapping by multiple next-generation sequencing aligners based on genome-wide characteristics. *Genomics.* 2017;109(3-4):186-191. doi:10.1016/j.ygeno.2017.03.001
40. Roberts ND, Kortschak RD, Parker WT, et al. A comparative analysis of algorithms for somatic SNV detection in cancer. *Bioinforma Oxf Engl.* 2013;29(18):2223-2230. doi:10.1093/bioinformatics/btt375
41. Cai L, Yuan W, Zhang Z, He L, Chou K-C. In-depth comparison of somatic point mutation callers based on different tumor next-generation sequencing depth data. *Sci Rep.* 2016;6:36540. doi:10.1038/srep36540
42. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinforma Oxf Engl.* 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352
43. VCFv4.2.pdf. <https://samtools.github.io/hts-specs/VCFv4.2.pdf>. Accessed May 24, 2019.
44. EGA European Genome-Phenome Archive. <https://ega-archive.org/>. Accessed April 12, 2019.
45. Tan A, Abecasis GR, Kang HM. Unified representation of genetic variants. *Bioinformatics.* 2015;31(13):2202-2204. doi:10.1093/bioinformatics/btv112
46. Carson AR, Smith EN, Matsui H, et al. Effective filtering strategies to improve data quality from population-based whole exome sequencing studies. *BMC Bioinformatics.* 2014;15:125. doi:10.1186/1471-2105-15-125
47. Jalali Sefid Dashti M, Gamiieldien J. A practical guide to filtering and prioritizing genetic variants. *BioTechniques.* 2017;62(1):18-30. doi:10.2144/000114492
48. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods.* 2010;7(8):575-576. doi:10.1038/nmeth0810-575
49. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. *Genome Res.* 2009;19(9):1553-1561. doi:10.1101/gr.092619.109
50. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44-57. doi:10.1038/nprot.2008.211

51. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25(1):25-29. doi:10.1038/75556
52. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* 2019;47(D1):D330-D338. doi:10.1093/nar/gky1055
53. Erinjeri NJ, Nicolson NG, Deyholos C, Korah R, Carling T. Whole-Exome Sequencing Identifies Two Discrete Druggable Signaling Pathways in Follicular Thyroid Cancer. *J Am Coll Surg.* 2018;226(6):950-959.e5. doi:10.1016/j.jamcollsurg.2018.01.059
54. Houlston RS, Cheadle J, Dobbins SE, et al. Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. *Nat Genet.* 2010;42(11):973-977. doi:10.1038/ng.670
55. Liu W, Li L, Ye H, Tao H, He H. Role of COL6A3 in colorectal cancer. *Oncol Rep.* 2018;39(6):2527-2536. doi:10.3892/or.2018.6331
56. Leo PJ, Madeleine MM, Wang S, et al. Defining the genetic susceptibility to cervical neoplasia-A genome-wide association study. *PLoS Genet.* 2017;13(8):e1006866. doi:10.1371/journal.pgen.1006866
57. Berndt SI, Camp NJ, Skibola CF, et al. Meta-analysis of genome-wide association studies discovers multiple loci for chronic lymphocytic leukemia. *Nat Commun.* 2016;7:10933. doi:10.1038/ncomms10933
58. Parry M, Rose-Zerilli MJ, Gibson J, et al. Whole exome sequencing identifies novel recurrently mutated genes in patients with splenic marginal zone lymphoma. *PLoS One.* 2013;8(12):e83244. doi:10.1371/journal.pone.0083244
59. Kawaguchi M, Hara N, Bilim V, et al. A diagnostic marker for superficial urothelial bladder carcinoma: lack of nuclear ATBF1 (ZFHX3) by immunohistochemistry suggests malignant progression. *BMC Cancer.* 2016;16(1):805. doi:10.1186/s12885-016-2845-5
60. Mutational heterogeneity in cancer and the search for new cancer-associated genes. - PubMed - NCBI. <https://www.ncbi.nlm.nih.gov/pubmed/23770567>. Accessed June 2, 2019.

8. Anexos

A continuación, se presenta el código que se ha utilizado para el análisis de los datos obtenidos por WES de los schwannomas esporádicos y asociados a neurofibromatosis tipo 2:

```
# CÓDIGO DISEÑADO POR EL GRUPO DE INVESTIGACIÓN #
##### Alineamiento #####
bwa mem # Alineamiento con el genoma de referencia.
samtools sort # Ordenar las secuencias.
samtools index # Crear índice de las secuencias.

##### Variant Calling #####
configureStrelkaGermlineWorkflow.py \ # Configuración del análisis de
Strelka.
    --exome \
    --archivo .bam ordenado \
    --genoma de referencia\
    --directorio del análisis
runWorkflow.py -m local -j # Ejecución de Strelka.

##### Normalización #####
bcftools norm # Normalización.
tabix # Crear índice de las variantes.

##### Variant Annotation #####
annovar .vcf.gz -protocol refGene,dbsfp30a,exac03nontcga -operation
g,f,f # Anotar variaciones, con las bases de datos y el tipo de
análisis (g: gene based (identificación del gen), f:filter based
(anotaciones según diferentes bases de datos))
tabix # Crear índice de las variantes

# CÓDIGO DISEÑADO PARA ESTE TRABAJO #
##### Filtrado y Priorización #####
echo "`date` INICIANDO FILTRADO Y PRIORIZACIÓN DE LA MUESTRA."
echo "`date` 1. Descomprimiendo el archivo .gz con las variantes
anotadas."
gunzip -k ${ruta del archivo}.annotated.vcf.gz

echo "`date` 2. Filtrado de calidad de las variantes encontradas."
# Filtrar las variantes que cumplen los criterios de calidad.
bcftools view -i 'FILTER="PASS" && GQ>=20 && DP>=10 && QUAL>=20'
${ruta del archivo}.annotated.vcf | bgzip > ${ruta de destino del
archivo}.calidad.vcf.gz

echo "`date` 3. Escoger variantes exónicas"
gunzip -k ${ruta del archivo}.calidad.vcf.gz
bcftools view -i 'Func.refGene="exonic"' ${ruta del
archivo}.calidad.vcf | bgzip > ${ruta de destino del
archivo}.exonic.vcf.gz

echo "`date` 4. Guardar variantes sin sentido y con pérdida del marco
de lectura"
gunzip -k ${ruta del archivo}.exonic.vcf.gz
bcftools view -i 'ExonicFunc.refGene = "stoploss" |
ExonicFunc.refGene="stopgain" |
```

```

ExonicFunc.refGene="frameshift_insertion" |
ExonicFunc.refGene="frameshift_deletion"' ${ruta del
archivo}.exonic.vcf | bgzip > ${ruta de destino del
archivo}.nonsense_indels.vcf.gz
tabix -f -p vcf ${ruta del archivo}. nonsense_indel.vcf.gz # Crear
índice de las variantes.

echo "`date` 5. Guardar variantes con cambio de sentido"
bcftools view -i 'ExonicFunc.refGene="nonsynonymous_SNV"' ${ruta del
archivo}.exonic.vcf | bgzip > ${ruta de destino del
archivo}.missense.vcf.gz

echo "`date` 6. Impacto biológico de las variantes con cambio de
sentido"
gunzip -k ${ruta del archivo}.missense.vcf.gz
$bin_bcftools view -i 'SIFT_pred="D" | LRT_pred="D" | MetaSVM_pred="D"
| Polyphen2_HDIV_pred="D" | Polyphen2_HVAR_pred="D" |
MutationTaster_pred="A"' ${ruta del archivo}.missense.vcf | bgzip >
${ruta de destino del archivo}.variantImpact.vcf.gz
# Guardar las variantes que son detectas como dañinas por, al menos,
un predictor in silico.
tabix -f -p vcf ${ruta del archivo}.variantImpact.vcf.gz # Crear
índice de las variantes.

echo "`date` 7. Eliminar archivos descomprimidos."
rm ${ruta del archivo}.annotated.vcf
rm ${ruta del archivo}.calidad.vcf
rm ${ruta del archivo}.exonic.vcf
rm ${ruta del archivo}.missense.vcf

## El resto del código está escrito en R ##
# Filtrado por frecuencia alélica e impacto biológico en las
mutaciones por cambio de sentido.
# Guardar las variantes que tienen una frecuencia alélica menor de
0.01 y son detectadas como deletéreas por los predictores LRT y
MutationTaster.
muestra_missense<- muestra_missense [muestra_missense $ExAC_ALL<0.01,]
muestra_missense <- muestra_missense [muestra_missense
$MutationTaster=="A" | muestra_missense $MutationTaster=="D" |
muestra_missense $LRT=="D",]

# Filtrado por frecuencia alélica en las mutaciones sin sentido y
pérdida del marco de lectura
muestra_nonsenseindel <- muestra_nonsenseindel[muestra_nonsenseindel
$ExAC_ALL<0.01,]

## Unir los archivos los dos archivos anteriores en un sólo
muestra <- rbind(muestra_missense, muestra_nonsenseindel)

## Obtener genes recurrentes ##
# Obtener la lista de genes mutados en cada muestra.
genes_muestra <- as.character(muestra[,muestra$Gene])

# Crear una lista con los genes mutados en todas las muestras de un
grupo.
listaGenesMutados <- c(genes_muestra1, genes_muestra2,
genes_muestra3...,)
listaGenesMutados <- sort(listaGenesMutados)
listaGenesMutados <- unique(listaGenesMutados)

```

```

# Obtener tabla con los genes mutados en cada muestra codificados con
1 (presencia del gen mutado) o 0 (ausencia del gen mutado).
genesMutados <- function(listaGenesMutados,muestra){
  vec = numeric()
  for(i in 1:length(listaGenes)){
    if(listaGenes[i] %in% muestra){
      vec[i] <- 1
    }else{
      vec[i] <- 0
    }
  }
  return(vec)
}

genes_cod_muestra <- genesMutados(listaGenesMutados,genes_muestra)
TablaGenesMutados <-
as.data.frame(cbind(genes_cod_muestral,genes_cod_muestra2...))
rownames(TablaGenesMutados <- listaGenesMutados

# Calcular la frecuencia con la que aparece cada gen mutado dentro de
cada grupo.
n<- número de muestras en cada grupo
for(i in 1:nrow(TablaGenesMutados)){
  TablaGenesMutados[i,n + 1] <-
round((sum(TablaGenesMutados[i,1:n])/n)*100)
}
colnames(TablaGenesMutados) <- c(seq(1,n), "Frecuencias (%)")

## Filtrar los genes que aparecen en más de 2 muestras.
frecuencia <- (2*100)/n
TablaGenesMutados2 <- TablaGenesMutados[TablaGenesMutados$`Frecuencias
(%)`>=frecuencia,]
genesRecurrentes <- rownames(TablaGenesMutados2)

muestra <- muestra[muestra$Gene %in% genesRecurrentes,]

```

Rutas metabólicas enriquecidas en Schwannomas Esporádicos

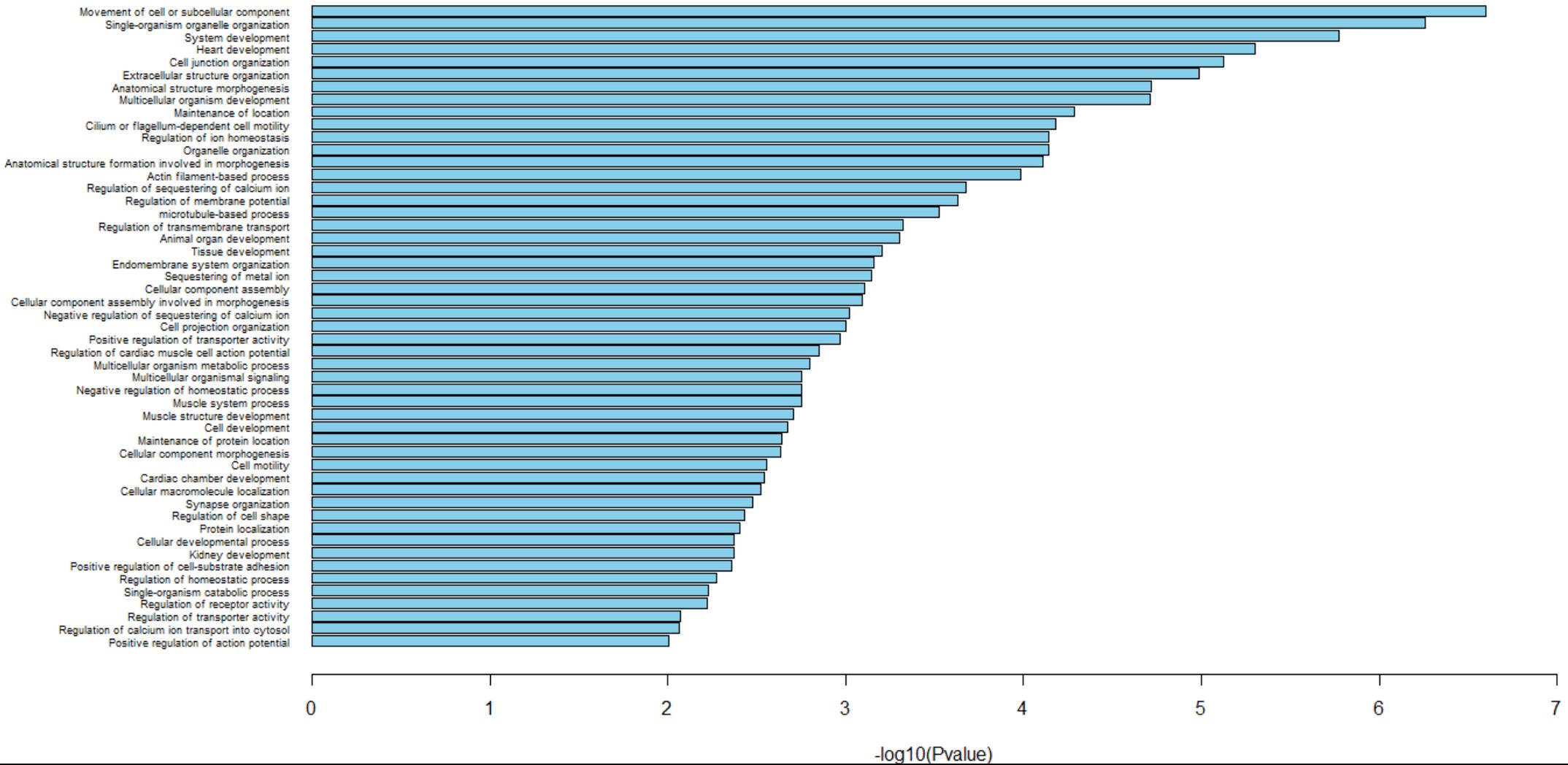


Figura 6. Rutas metabólicas enriquecidas en los Schwannomas esporádicos.

Tabla 3. Genes mutados en cada grupo de Schwannomas.

Schwannomas asociados a NF2			Schwannomas Esporádicos		
Genes	Casos	Frecuencia (%)	Genes	Casos	Frecuencia (%)
HLA-DRB5	6	67	HLA-DRB1	12	52
PABPC3	5	56	PABPC3	12	52
CFTR	4	44	TTN	10	43
FLNC	4	44	BBS9	6	26
LRP2	4	44	COL6A3	6	26
NF2	4	44	DAAM2	6	26
SCUBE3	4	44	DCHS1	6	26
ZFHX3	4	44	HTT	6	26
ABCA13	3	33	LAMA5	6	26
AHNAK2	3	33	NF2	6	26
ARHGAP8	3	33	PLEC	6	26
CACNA1S	3	33	SYNE1	6	26
CBS	3	33	BIRC6	5	22
CEP131	3	33	CAD	5	22
CYR61	3	33	CEP112	5	22
DFNB31	3	33	COL11A2	5	22
DNAH7	3	33	DNAH1	5	22
FREM3	3	33	DNAH5	5	22
HSPG2	3	33	DNAH7	5	22
LYST	3	33	FAT4	5	22
NRAP	3	33	LOXHD1	5	22
NUBP2	3	33	OBSCN	5	22
NUMA1	3	33	PRMT3	5	22
PCDHGA11	3	33	TTC30B	5	22
RFX5	3	33	ADAMTS8	4	17
RFX8	3	33	AGBL5	4	17
RYR1	3	33	AGL	4	17
SYNE2	3	33	ANK2	4	17
TGFBR2	3	33	CAPN9	4	17
TTN	3	33	CBS	4	17
UNC13B	3	33	EFCAB7	4	17
ABCA12	2	22	EIF3A	4	17
ABCB6	2	22	EXOC3L1	4	17
ADAM17	2	22	FREM2	4	17
ADAMTS8	2	22	HEATR1	4	17
AFF3	2	22	IGFALS	4	17
ALDH1L1	2	22	KIAA1217	4	17
ALG8	2	22	LAMA3	4	17
ANK3	2	22	MEGF8	4	17
ANKRD2	2	22	MLH3	4	17
ANO5	2	22	MYO1H	4	17

ARHGEF1	2	22	NEB	4	17
ARHGEF10L	2	22	NWD1	4	17
ARSB	2	22	OTOF	4	17
ARVCF	2	22	PIBF1	4	17
ATM	2	22	PKHD1	4	17
BIRC6	2	22	PKHD1L1	4	17
C2	2	22	PSMD5	4	17
CBFA2T3	2	22	RYR1	4	17
CCDC191	2	22	SMC2	4	17
CD101	2	22	SPATA45	4	17
CDH23	2	22	TNC	4	17
CELSR2	2	22	TTC21A	4	17
CGN	2	22	UNC13B	4	17
CILP	2	22	UTP20	4	17
CLCN1	2	22	UTRN	4	17
CNTN6	2	22	WFS1	4	17
COL6A2	2	22	ZC3H14	4	17
COL6A3	2	22	ZRANB3	4	17
CSMD2	2	22	AARS	3	13
CYB5D2	2	22	ACAP3	3	13
CYP39A1	2	22	ACVR2B	3	13
DCDC2B	2	22	ADAMTSL4	3	13
DGKA	2	22	ADGRV1	3	13
DHX38	2	22	AFF3	3	13
DIP2B	2	22	AHDC1	3	13
DLG5	2	22	ALDH3B2	3	13
DNAH1	2	22	AMMECR1L	3	13
DNAH3	2	22	AP1M2	3	13
DPT	2	22	APPL1	3	13
DPYSL5	2	22	AQP12B	3	13
DRC7	2	22	ARHGEF37	3	13
DSP	2	22	ASTN2	3	13
EMC1	2	22	ATG2B	3	13
ENTHD2	2	22	ATP8B4	3	13
EPG5	2	22	ATP9B	3	13
EXOC1	2	22	BAIAP3	3	13
EXOC4	2	22	CAPZA3	3	13
FAM129A	2	22	CARS	3	13
FAM135A	2	22	CDC20	3	13
FAM160A2	2	22	CEP350	3	13
FANCC	2	22	CLCN3	3	13
FAT1	2	22	CNTN3	3	13
FAT4	2	22	COL11A1	3	13
FLNB	2	22	COL5A2	3	13
FNDC3B	2	22	CORO7	3	13

FOCAD	2	22	CRYGA	3	13
FOXM1	2	22	CUL7	3	13
FRAS1	2	22	DCHS2	3	13
GALC	2	22	DDX58	3	13
GPRC5A	2	22	DIAPH1	3	13
GPRC6A	2	22	DNAH10	3	13
GRIK5	2	22	DNAH11	3	13
GRWD1	2	22	DNAH3	3	13
HELZ2	2	22	DNAH6	3	13
HPS5	2	22	DOCK5	3	13
HS6ST3	2	22	DOPEY2	3	13
ICE2	2	22	DSP	3	13
IFT80	2	22	DUS2	3	13
IL16	2	22	DVL2	3	13
INSL3	2	22	ECEL1	3	13
IP6K3	2	22	EDNRB	3	13
IQGAP2	2	22	EIF3B	3	13
ISL1	2	22	EML5	3	13
ISM1	2	22	EPG5	3	13
KCNK7	2	22	FAM126A	3	13
KIAA0100	2	22	FAM83A	3	13
KIAA0196	2	22	FARP1	3	13
KIAA0391	2	22	FBN3	3	13
KIF1B	2	22	FGFR1	3	13
KLF11	2	22	FLNB	3	13
KMT5C	2	22	FN1	3	13
KRT75	2	22	FUT2	3	13
LAMC2	2	22	GGA3	3	13
LGR4	2	22	GPATCH2L	3	13
LGR6	2	22	GPD2	3	13
LMF1	2	22	HEATR5B	3	13
LMF2	2	22	HIP1R	3	13
LOXHD1	2	22	HLA-DRB5	3	13
LRBA	2	22	IFNA10	3	13
LRRC36	2	22	IFT122	3	13
MAP3K12	2	22	IGSF21	3	13
MAP4K2	2	22	INO80	3	13
MBOAT2	2	22	IQGAP3	3	13
MFSD12	2	22	JAK2	3	13
MSH2	2	22	KCTD3	3	13
MTF1	2	22	KIF16B	3	13
MYBPC3	2	22	KIFC2	3	13
MYH15	2	22	KMT2A	3	13
MYO3B	2	22	KRT6B	3	13
MYOM1	2	22	LAMA1	3	13

MYOM3	2	22	LAMB4	3	13
MYPN	2	22	LOXL2	3	13
NBEAL2	2	22	LPIN3	3	13
NCF2	2	22	LRP2	3	13
NDUFS1	2	22	MACF1	3	13
NEB	2	22	MARC1	3	13
NEDD4	2	22	MED12	3	13
OBSCN	2	22	MGAT4B	3	13
OTOF	2	22	MMP27	3	13
OTOG	2	22	MMRN1	3	13
OTUD7A	2	22	MROH2A	3	13
P3H1	2	22	MROH7	3	13
PADI4	2	22	MYH1	3	13
PCCA	2	22	MYH6	3	13
PCF11	2	22	MYH7B	3	13
PCNXL3	2	22	MYH8	3	13
PDIA5	2	22	MYO15A	3	13
PIGU	2	22	MYO3A	3	13
PKHD1	2	22	MYOM2	3	13
PKHD1L1	2	22	MYOM3	3	13
PKP3	2	22	NCAPD3	3	13
PLAA	2	22	NFATC2	3	13
PLEC	2	22	NID1	3	13
PLXNA4	2	22	NOS1	3	13
POLR1C	2	22	NOTCH4	3	13
PPP2R3A	2	22	NR0B2	3	13
PTGIR	2	22	NRAP	3	13
PTPN13	2	22	NSUN5	3	13
PTPRF	2	22	NUMA1	3	13
PTPRT	2	22	PCCB	3	13
PUM2	2	22	PCDHA11	3	13
PUS3	2	22	PCDHGA6	3	13
PYGL	2	22	PCF11	3	13
RAD54L	2	22	PDE11A	3	13
RALGAPA1	2	22	PDIA2	3	13
RANBP6	2	22	PKD1	3	13
RNF217	2	22	PLB1	3	13
ROBO1	2	22	PLEKHA7	3	13
RPS6KC1	2	22	POLE	3	13
RYR3	2	22	PSME4	3	13
SCN4A	2	22	PTPRK	3	13
SCRN2	2	22	PTPRQ	3	13
SDHD	2	22	PUM1	3	13
SEL1L3	2	22	QSER1	3	13
SI	2	22	RAI1	3	13

SLC10A2	2	22
SLC26A1	2	22
SLC2A4	2	22
SLC9A3R1	2	22
SMG1	2	22
SNCAIP	2	22
SP1	2	22
SPECC1L	2	22
STEAP3	2	22
SVOPL	2	22
SYNDIG1	2	22
TBX5	2	22
TIAM1	2	22
TIMM44	2	22
TKT	2	22
TMEM104	2	22
TMEM143	2	22
TMEM184A	2	22
TMPRSS11F	2	22
TRAP1	2	22
TRIM24	2	22
TRIM29	2	22
TRIOBP	2	22
TRIP11	2	22
TTC22	2	22
TTC28	2	22
TTYH1	2	22
TYR	2	22
UTP20	2	22
VILL	2	22
VPS13D	2	22
VWF	2	22
WDR64	2	22
XAB2	2	22
ZIC5	2	22
ZNF248	2	22

RANBP17	3	13
RELN	3	13
REXO1	3	13
ROBO3	3	13
RPL3L	3	13
RYR3	3	13
SACS	3	13
SCAPER	3	13
SCARA3	3	13
SCN7A	3	13
SCN9A	3	13
SLC5A4	3	13
SLC9A4	3	13
SMYD1	3	13
SORBS2	3	13
STEAP4	3	13
SVEP1	3	13
SYNE2	3	13
TAOK3	3	13
TBC1D1	3	13
TDRD3	3	13
TGM1	3	13
TGM7	3	13
THSD7A	3	13
THSD7B	3	13
TJP1	3	13
TLL1	3	13
TMC2	3	13
TNFAIP3	3	13
TOP3A	3	13
TPR	3	13
TRANK1	3	13
TSFM	3	13
TTC17	3	13
TTLL1	3	13
TULP4	3	13
TYRP1	3	13
UNC13C	3	13
USP34	3	13
VAV3	3	13
VPS13A	3	13
VPS13B	3	13
VPS39	3	13
VWA3A	3	13
WDR66	3	13

WNT10A	3	13
ZFHX3	3	13
ZGRF1	3	13
ZHX2	3	13
ZNF292	3	13
ZNF318	3	13
ZNF543	3	13
AAMP	2	9
ABCA7	2	9
ABCC12	2	9
ABCC2	2	9
ABCD2	2	9
ABHD16A	2	9
ABHD6	2	9
ACO2	2	9
ACOT8	2	9
ACTN2	2	9
ACTR5	2	9
ADAM15	2	9
ADAMTS1	2	9
ADAMTSL3	2	9
ADARB2	2	9
ADGRF4	2	9
ADGRL1	2	9
AFF2	2	9
AGBL4	2	9
AGER	2	9
AGRN	2	9
AGTR1	2	9
AHNAK2	2	9
AIP	2	9
AK8	2	9
AKAP9	2	9
AKR1B15	2	9
ALDH1L1	2	9
ALKBH8	2	9
ALPK1	2	9
ALS2	2	9
ALX1	2	9
ANGPTL7	2	9
ANK3	2	9
AOC1	2	9
AP4B1	2	9
APBA3	2	9
AR	2	9

ARAP3	2	9
ARFGEF3	2	9
ARHGAP11A	2	9
ARHGAP24	2	9
ARHGAP31	2	9
ARHGAP8	2	9
ARHGEF16	2	9
ARID4B	2	9
ARPC3	2	9
ARRDC2	2	9
ART3	2	9
ASAH2	2	9
ASPM	2	9
ASPN	2	9
ASUN	2	9
ASXL1	2	9
ATAD3C	2	9
ATG4B	2	9
ATM	2	9
ATP7A	2	9
ATXN7L1	2	9
AURKB	2	9
AVIL	2	9
BCL9	2	9
BFSP1	2	9
BICD2	2	9
BOC	2	9
BPIFB3	2	9
BRINP3	2	9
BRIP1	2	9
BSN	2	9
BYSL	2	9
C18orf25	2	9
C1orf100	2	9
C1orf87	2	9
C1QTNF9	2	9
C21orf33	2	9
C21orf58	2	9
C9orf135	2	9
CACNA1I	2	9
CALHM1	2	9
CAMSAP2	2	9
CAND2	2	9
CAPN10	2	9
CAPN12	2	9

CASQ1	2	9
CAT	2	9
CBLC	2	9
CBR1	2	9
CCBE1	2	9
CCDC146	2	9
CCDC150	2	9
CCDC74B	2	9
CCDC86	2	9
CCDC89	2	9
CCDC92	2	9
CCHCR1	2	9
CCT7	2	9
CDADC1	2	9
CDH18	2	9
CDH22	2	9
CDH23	2	9
CDH26	2	9
CDK18	2	9
CDK2	2	9
CEP295NL	2	9
CES5A	2	9
CFAP43	2	9
CFAP44	2	9
CFAP58	2	9
CFAP61	2	9
CFHR2	2	9
CGNL1	2	9
CHD6	2	9
CHD8	2	9
CHODL	2	9
CHP2	2	9
CHST10	2	9
CHST4	2	9
CHTF18	2	9
CIITA	2	9
CLCA1	2	9
CLCNKA	2	9
CLDN18	2	9
CLEC16A	2	9
CLHC1	2	9
CLIP2	2	9
CLIP4	2	9
CLMN	2	9
CLSTN1	2	9

CNKS3	2	9
CNNM1	2	9
CNTRL	2	9
COG7	2	9
COL12A1	2	9
COL16A1	2	9
COL27A1	2	9
COL6A2	2	9
COL6A5	2	9
COL7A1	2	9
CRHR2	2	9
CRTC2	2	9
CRY2	2	9
CSMD1	2	9
CSMD2	2	9
CSMD3	2	9
CSNK1G1	2	9
CSNK2A3	2	9
CSPG4	2	9
CSPP1	2	9
CSTF2	2	9
CTH	2	9
CTSB	2	9
CYP2C9	2	9
CYP4A22	2	9
CYP4F11	2	9
DAPK3	2	9
DDX24	2	9
DDX56	2	9
DENND5A	2	9
DFNB31	2	9
DGKG	2	9
DGKZ	2	9
DHCR24	2	9
DHX16	2	9
DHX8	2	9
DIP2B	2	9
DLG5	2	9
DMGDH	2	9
DMRT2	2	9
DNAH17	2	9
DNAH2	2	9
DNAH8	2	9
DNAH9	2	9
DNAJC10	2	9

DNAJC14	2	9
DNAJC17	2	9
DNER	2	9
DOCK1	2	9
DOCK8	2	9
DPCD	2	9
DPP10	2	9
DST	2	9
DUS3L	2	9
DYNC1H1	2	9
DYNC2H1	2	9
EDEM3	2	9
EEF2K	2	9
EFEMP1	2	9
EGF	2	9
EGFLAM	2	9
EIF3I	2	9
EIF5B	2	9
ELMO3	2	9
EPHB6	2	9
EPM2A	2	9
EPN1	2	9
ERAP2	2	9
ERBB3	2	9
ERN2	2	9
ESRRB	2	9
EVC2	2	9
EVI5	2	9
EXOC1	2	9
F5	2	9
FAM186B	2	9
FAM188B	2	9
FAM198A	2	9
FAM20A	2	9
FAM214A	2	9
FAM45A	2	9
FAM81B	2	9
FANCC	2	9
FANCI	2	9
FANK1	2	9
FAT2	2	9
FBN1	2	9
FBN2	2	9
FBXO31	2	9
FEM1A	2	9

FGD5	2	9
FGD6	2	9
FGGY	2	9
FIBCD1	2	9
FIGNL1	2	9
FKBP15	2	9
FLG	2	9
FMNL3	2	9
FNDC1	2	9
FOCAD	2	9
FOXBI	2	9
FOXRED2	2	9
FRAS1	2	9
FRS3	2	9
FRY	2	9
FTSJ3	2	9
FZD3	2	9
GBA	2	9
GCC2	2	9
GCFC2	2	9
GFAP	2	9
GFI1B	2	9
GFM1	2	9
GGA2	2	9
GGCT	2	9
GIT2	2	9
GJA5	2	9
GK2	2	9
GLCE	2	9
GLI3	2	9
GLT8D2	2	9
GLTSCR2	2	9
GNA13	2	9
GNAQ	2	9
GOLGA3	2	9
GOLGA4	2	9
GPR156	2	9
GPR84	2	9
GRIN2B	2	9
GRK4	2	9
GRM7	2	9
GSE1	2	9
GSTO1	2	9
HAAO	2	9
HAP1	2	9

HCFC1	2	9
HCN2	2	9
HCRTR1	2	9
HDAC4	2	9
HDGFRP2	2	9
HDX	2	9
HECTD4	2	9
HIC2	2	9
HIVEP2	2	9
HMGCLL1	2	9
HOOK2	2	9
HOXB5	2	9
HPR	2	9
HRASLS	2	9
HSD17B14	2	9
HTRA4	2	9
IDNK	2	9
IFRD2	2	9
IFT46	2	9
IGSF3	2	9
IL34	2	9
INADL	2	9
INPP4B	2	9
IQSEC3	2	9
ITGA11	2	9
ITGB8	2	9
ITM2C	2	9
ITPR3	2	9
IZUMO3	2	9
JADE1	2	9
JAKMIP2	2	9
JMJD1C	2	9
PLA2G4B	2	9
JPH3	2	9
KALRN	2	9
KANK2	2	9
KAT6A	2	9
KCNA3	2	9
KCNH6	2	9
KCNK13	2	9
KCNK7	2	9
KDM5A	2	9
KDR	2	9
KIAA0319L	2	9
KIAA0368	2	9

KIAA0408	2	9
KIAA0586	2	9
KIAA1522	2	9
KIAA1549	2	9
KIAA2026	2	9
KIF21A	2	9
KIF21B	2	9
KLC2	2	9
KLHL1	2	9
KLHL28	2	9
KLRK1	2	9
KMT2C	2	9
KMT2D	2	9
KPNA3	2	9
KRT28	2	9
KRT38	2	9
KRT76	2	9
KRT82	2	9
KRT84	2	9
LAMA4	2	9
LAMB2	2	9
LDB3	2	9
LOXL4	2	9
LPA	2	9
LRMP	2	9
LRP12	2	9
LRP1B	2	9
LRRC16A	2	9
LRRC8E	2	9
LYSMD2	2	9
LYST	2	9
MAATS1	2	9
MAD1L1	2	9
MAGEC1	2	9
MAML1	2	9
MAN2B1	2	9
MAP3K3	2	9
MAP7D1	2	9
MAPK15	2	9
MAPK8IP1	2	9
MARK3	2	9
MARS2	2	9
MARVELD2	2	9
MASP2	2	9
MAST3	2	9

MCEE	2	9
MCM2	2	9
MCPH1	2	9
MDN1	2	9
MEP1A	2	9
MFGE8	2	9
MICAL2	2	9
MICAL3	2	9
MKX	2	9
MLH1	2	9
MMP7	2	9
MPI	2	9
MPPED1	2	9
MSH2	2	9
MT1E	2	9
MTERF3	2	9
MTFR1	2	9
MTMR9	2	9
MYCBP2	2	9
MYH11	2	9
MYH13	2	9
MYH14	2	9
MYH7	2	9
MYO1F	2	9
MYO5A	2	9
MYO5B	2	9
MYOF	2	9
NBEAL2	2	9
NCBP3	2	9
NCKAP5	2	9
NCL	2	9
NDUFS3	2	9
NEFM	2	9
NEK11	2	9
NEK5	2	9
NEO1	2	9
NEURL4	2	9
NFAT5	2	9
NHSL1	2	9
NKAIN3	2	9
NLRC3	2	9
NME6	2	9
NOD2	2	9
NOS2	2	9
NPHP3	2	9

NPHP4	2	9
NPL	2	9
NPR1	2	9
NTSR1	2	9
NUFIP1	2	9
NUPL2	2	9
ODF3	2	9
OGDH	2	9
OGFOD2	2	9
OLAH	2	9
OLFML2B	2	9
OR10G3	2	9
OR1J1	2	9
OR4A15	2	9
OR4A47	2	9
OR4D1	2	9
OR51I1	2	9
OR5M3	2	9
OR6C2	2	9
OR6C75	2	9
OR6K2	2	9
OR8B8	2	9
OR8U1	2	9
OTOGL	2	9
OXR1	2	9
P4HA2	2	9
P4HA3	2	9
PADI2	2	9
PAK7	2	9
PANK2	2	9
PAPPA	2	9
PARP8	2	9
PARS2	2	9
PATL1	2	9
PBK	2	9
PBXIP1	2	9
PCDHA12	2	9
PCDHA4	2	9
PCDHB11	2	9
PCDHGA2	2	9
PCDHGA8	2	9
PCLO	2	9
PCM1	2	9
PCSK5	2	9
PDCD11	2	9

PDE1A	2	9
PDE6A	2	9
PDE6B	2	9
PEPD	2	9
PHLDB3	2	9
PIF1	2	9
PIGF	2	9
PIWIL2	2	9
PIWIL4	2	9
PKP1	2	9
PKP2	2	9
PKP3	2	9
PLA2G7	2	9
PLCH2	2	9
PLD2	2	9
PLEKHG1	2	9
PLEKHG5	2	9
PLP2	2	9
PLPP2	2	9
PLRG1	2	9
PMP2	2	9
PNPLA7	2	9
POLG	2	9
POLH	2	9
POLR3A	2	9
POMGNT2	2	9
PPARGC1A	2	9
PPM1H	2	9
PPP1R9A	2	9
PPP2R1B	2	9
PPRC1	2	9
PREX1	2	9
PRKD3	2	9
PRR30	2	9
PRRC2A	2	9
PRRX1	2	9
PRX	2	9
PSD3	2	9
PTBP3	2	9
PTGIS	2	9
PTK2B	2	9
PTPN4	2	9
PTPRJ	2	9
PUS7L	2	9
R3HDM1	2	9

RAB27A	2	9
RABGAP1	2	9
RAD21L1	2	9
RAD50	2	9
RAD54L2	2	9
RANBP6	2	9
RAPGEF3	2	9
RASA3	2	9
RB1	2	9
RBMXL2	2	9
RECK	2	9
RECQL	2	9
RGS8	2	9
RHOBTB2	2	9
RIMKLA	2	9
RIOK2	2	9
RLTPR	2	9
RNF113B	2	9
RNF207	2	9
RNF44	2	9
RPAP3	2	9
RPF1	2	9
RPL26L1	2	9
RRP12	2	9
RSF1	2	9
RSL1D1	2	9
RUSC1	2	9
RUSC2	2	9
RWDD2B	2	9
SAV1	2	9
SBF2	2	9
SCN4A	2	9
SCN5A	2	9
SCNN1A	2	9
SCNN1G	2	9
SCUBE2	2	9
SDK2	2	9
SDR42E1	2	9
SEC16B	2	9
SEMA4D	2	9
SEMA5B	2	9
SERPINA10	2	9
SERPINA4	2	9
SERPINB12	2	9
SEZ6	2	9

SFMBT2	2	9
SH2D3A	2	9
SH3TC1	2	9
SH3TC2	2	9
SHPRH	2	9
SI	2	9
SLA2	2	9
SLAMF8	2	9
SLC10A1	2	9
SLC10A2	2	9
SLC12A5	2	9
SLC17A6	2	9
SLC25A26	2	9
SLC25A40	2	9
SLC26A10	2	9
SLC26A9	2	9
SLC27A5	2	9
SLC2A2	2	9
SLC30A6	2	9
SLC39A2	2	9
SLC4A5	2	9
SLC5A2	2	9
SLC6A3	2	9
SLF1	2	9
SMPD1	2	9
SMTN	2	9
SNTG2	2	9
SPAG17	2	9
SPATA20	2	9
SPATA5L1	2	9
SPATA6L	2	9
SPDEF	2	9
SPHKAP	2	9
SPTA1	2	9
SPTBN1	2	9
SPTBN4	2	9
SRCAP	2	9
SSFA2	2	9
STAB1	2	9
STAB2	2	9
STOX1	2	9
SUCLA2	2	9
SULF2	2	9
SUSD1	2	9
SYCE2	2	9

SYNE3	2	9
SYNJ2	2	9
SYT10	2	9
SYT16	2	9
TAB2	2	9
TAF1C	2	9
TAS1R1	2	9
TAS2R10	2	9
TATDN1	2	9
TAX1BP3	2	9
TBX1	2	9
TBXAS1	2	9
TEAD3	2	9
TECPR1	2	9
TEKT4	2	9
TENM4	2	9
TEP1	2	9
TET2	2	9
TFAP2D	2	9
TJP3	2	9
TLK1	2	9
TLN2	2	9
TMEM154	2	9
TMEM168	2	9
TMEM247	2	9
TMEM41A	2	9
TMEM43	2	9
TMPRSS2	2	9
TMPRSS7	2	9
TNFRSF21	2	9
TNRC6A	2	9
TNS2	2	9
TPO	2	9
TRAF3IP2	2	9
TRAPPC9	2	9
TRIM42	2	9
TRIM50	2	9
TRIP10	2	9
TRIP11	2	9
TRIP6	2	9
TRMT11	2	9
TRMT1L	2	9
TRPM2	2	9
TRRAP	2	9
TSC1	2	9

TSC22D1	2	9
TSHZ2	2	9
TTC37	2	9
UBXN11	2	9
UMODL1	2	9
UPF2	2	9
USP44	2	9
USP48	2	9
USP6	2	9
VARS	2	9
VPS13C	2	9
VPS16	2	9
VPS41	2	9
VPS8	2	9
VWA2	2	9
WDFY4	2	9
WDR1	2	9
WDR12	2	9
WDR36	2	9
WDR45	2	9
WSB2	2	9
WWC1	2	9
WVOX	2	9
XIRP2	2	9
XPNPEP1	2	9
XPO4	2	9
XXYL1	2	9
XYLT2	2	9
YAE1D1	2	9
YWHAE	2	9
YY1AP1	2	9
ZBTB16	2	9
ZC3H4	2	9
ZFYVE26	2	9
ZFYVE28	2	9
ZIC5	2	9
ZMYND15	2	9
ZNF143	2	9
ZNF180	2	9
ZNF319	2	9
ZNF462	2	9
ZNF513	2	9
ZNF534	2	9
ZNF570	2	9
ZNF654	2	9

ZNF750	2	9
ZNF816	2	9
ZNF862	2	9
ZNF91	2	9
ZSCAN29	2	9
ZSWIM5	2	9
ZW10	2	9
ZYG11A	2	9